Chapter 2

Review of literature
Rheumatic diseases including RA and OA are diverse group of disorders affecting the musculoskeletal system. The etiopathogenesis of these diseases is poorly understood. Hence, one of the biggest challenges in arthritis research is to identify molecular markers that are suitable for diagnosis, monitoring of disease activity and prediction of disease outcome. The advent of high-throughput technologies have accelerated research in the search for new potential biomarkers that can aid in the early diagnosis as well as prognosis of RA and OA. These studies can be categorized as genomics, transcriptomics and proteomics. In this section an overview of transcriptomic and proteomics studies in RA and OA that have been reported in the context of this study have been discussed.

2.1 Transcriptomic alterations in RA and OA

Discovery of microarray technology has revolutionized the study of gene expression profiles for a variety of diseases including arthritis. It has generated new perspectives for the high-throughput analysis of biological systems. Further it is possible to monitor the expression of thousands of genes in a single experiment by using cDNA and oligonucleotide microarray platforms (Toonen et al., 2008).

One of the first gene expression studies in RA was carried out using subtractive hybridization and cDNA arrays. This study revealed increased expression of genes associated with chronic inflammation such as immunoglobulins and HLA-DR in the synovium of RA patients when compared to the synovium of normal individuals (Zanders et al., 2000). In another study, it was demonstrated that 63 genes were found to be differentially expressed between the synovial tissues of RA and OA patients. Some of the
genes overexpressed in RA include, cathepsin L (CTSL), guanylate binding protein 1, interferon-inducible (GBP1) and Glyoxalase 1 (GLO1) (Devauchelle et al., 2004).

A large scale expression profiling study of erosive RA synovial tissues by cDNA microarray technology demonstrated considerable heterogeneity between different patients. This study revealed the existence of two distinct types of RA synovial tissues (RAI and RAIi). RAI tissues showed an elevated expression of genes pertaining to inflammation and adaptive immune response. Detailed analyses of the genes involved in the high inflammation tissues indicated their involvement in interferon/ signal transducer and activator of transcription (STAT)-1 pathway. In contrast, RAIi tissues showed increased expression of genes associated with fibroblast dedifferentiation. This study suggested the heterogeneity nature of RA (van der Pouw Kraan et al., 2003). A study by Tsubaki and his co-authors, also demonstrated the existence of heterogeneity in RA patients using synovial lining tissues by cDNA microarray consisting of 23,040 cDNAs (Tsubaki et al., 2005). A cDNA microarray platform consisting of 14,519 genes were used to compare the synovial lining cells of RA and OA patients. The expression of signal transducer and activator of transcription 1 (STAT1), interferon regulatory factor 1 (IRF1), and the chemokines CXCL9, CXCL10, and CCL5 were significantly higher in the synovium of RA than that of OA and indicated a significant role of synovial lining cells in inflammatory and proliferative processes in RA (Yoshida et al., 2012).

In addition to synovial tissues, gene expression studies were also carried out using synovial fibroblasts. Fibroblasts like synoviocytes (FLS) are key players in joint destruction in RA and studying the mRNA expression profile of FLS may provide vital clues to the disease pathogenesis. Transcriptomic profiling was carried out using
synoviocyte cultures obtained from RA patients and control cases using a cDNA array consisting of 588 cancer related genes. Platelet-derived growth factor receptor alpha (PDGFRA), plasminogen activator inhibitor-1 (PAPI), and stromal cell derived factor 1A (SDF1A) were found to be upregulated in RA synoviocytes than its normal counterpart from this study (Watanabe et al., 2002). In another study, it has been shown that the heterogeneity between RA synovial tissues was reflected in FLS as well, thereby providing a possible association between FLS behaviour and inflammation status of RA synovium (Kasperkovitz et al., 2005). An mRNA expression profiling study between cultured OA, RA and healthy synovial fibroblasts illustrated that OA synovial fibroblasts displayed a homogeneous transcriptomic profile than RA synovial fibroblasts (Del Rey et al., 2012).

Bone marrow derived mononuclear cells (BMMCs) from RA and OA were also subjected to microarray analyses in a study by Nakamura and his colleagues. Out of the 103 genes that were identified from this study, 10 genes including FK506-binding protein 5 (FKBP5), C-type lectin superfamily member 9 (CLECSF9), tyrosylprotein sulfotransferase 1 (TPST1), chemokine receptor 4 (CXCR4) and nuclear factor-kappa B (NF-kappaB) were found to be overexpressed in RA when compared to OA patients (Nakamura et al., 2006).

Due to systemic manifestations in RA, gene expression profiling studies were also carried out using peripheral blood mononuclear cells (PBMCs). 81 genes were found to be differentially expressed in the PBMCs of RA patients when compared to healthy control individuals in a study by Batliwalla and his group. These genes include IL-1 receptor antagonist (IL1RN), S100 calcium binding protein A12 (S100A12) and Grb2-associated binding protein (GAB1) (Batliwalla et al., 2005). Olsen and colleagues in 2004, have shown that genes involved in cell cycle and immune regulation were found to be
differentially expressed in early arthritis (disease duration <2 years) when compared to established arthritis (with an average disease duration of 10 years) using PBMCs. This study suggested that early arthritis signature might partly reflect the response to unknown infectious agents (Olsen et al., 2004).

In order to predict the most beneficial treatment for RA patients at an early stage, transcriptomic studies are being carried out to gain insight into the pharmacodynamics and to predict genes responsive to various drugs used for RA treatment. Gene expression profiling of RA synovial tissues from responders and non-responders of infliximab revealed the differential expression of 279 genes. Out of these 279 genes, matrix metalloproteinase 3 (MMP3) was found to be significantly overexpressed in good responders when compared to non-responders suggesting its potential to be used as a marker for good therapeutic response. Further, this study also demonstrated that tumor necrosis factor alpha (TNF-α) could be an important biomarker for successful infliximab treatment (Lindberg et al., 2006). In a recent study, pharmacogenomic analyses by transcriptomic profiling of PBMCs from RA patients showed marked interindividual variations in the pharmacological responses at 3 and 6 months after rituximab treatment (Vosslander et al., 2011). It has also been shown that baseline expression levels of interferon type I response genes are useful as a predictive biomarker for non-response to rituximab in RA patients (Raterman et al., 2012).

A recent study has shown that microarray technology can be effectively used to predict the response to anti-TNF treatment in RA patients (Toonen et al., 2012). Sanayama and his colleagues illustrated that genome-wide DNA microarray can be used to identify candidate biomarkers to predict the therapeutic response to tocilizumab in RA patients and
demonstrated the association of interferon signaling and metallothioneins in RA pathophysiology (Sanayama et al., 2014).

A whole genome expression profiling of subchondral bone from OA patients using Agilent microarray platform has been recently reported. A total of 972 genes were differentially expressed by ≥2-fold in the subchondral bone of OA patients when compared to that of non-OA patients. Novel pathways implicated in bone remodeling by osteoclasts such as leptin and periostin were identified in this study. (Chou et al., 2013).

Meta analysis of gene expression profiles of RA and OA patients especially from synovial tissues was carried out to investigate the expression signatures specific for these two diseases. It was found from the analysis that immunity, inflammation and apoptosis related pathways were commonly associated with the development of RA and OA. Regulation of autophagy, endocytosis, calcium transport and endoplasmic reticulum stress and vasopressin related pathways showed significant differences between RA and OA thereby providing insights in to the molecular mechanism of these two diseases (Li et al., 2014). In another meta analysis study, gene expression profiles in BMMCs from 9 RA and 10 OA obtained from Gene Expression Omnibus (GEO) were compared. It was found that 2581 genes were upregulated and 649 genes were downregulated in RA compared to that of OA. Differentially expressed genes including SP1 transcription factor (SP1), retinoic acid receptor, alpha (RARA), v-ets avian erythroblastosis virus E26 oncogene homolog 1 (ETS1), FBJ murine osteosarcoma viral oncogene homolog (FOS) and estrogen receptor 1 (ESR1) were found to be associated with RA through cancer and immunity related pathways. It was also suggested that these genes might also serve as therapeutic targets for RA (Yin et al., 2013).
The above mentioned gene expression studies have been performed using synovial tissues, synovial fibroblasts, PBMCs and BMMCs obtained from RA and/or OA. Further only limited number of genes were analysed at a time. A microarray study using the SFMCs from RA and OA patients has not been reported yet.

2.2 Proteomic alterations in RA and OA

Proteomics can be defined as the global characterization of the protein complement, the proteome of a cell type or tissue. The advantage of proteomics is that the actual functional molecules which will provide a true picture of events happening at the site of disease activity are being studied. Proteomic technologies have provided impetus for biomarker discovery for a wide range of diseases including arthritis. The diagnostic tools that are currently in use have their own limitations and provide an inaccurate assessment of disease progression.

Studying the synovial fluid proteome should yield a higher concentration of potential biomarkers than serum or plasma, as the synovial fluid is in direct physical contact with the synovium, ligament, meniscus, joint capsule and bone (Ruiz-Romero et al., 2010). Alterations in the structure and metabolism of any of these tissues during disease should be reflected as alterations in the composition of the synovial fluid proteome. Therefore, the synovial fluid proteome has the potential to indicate the severity and progression of the disease (Hui et al., 2012). Advances in proteomic technologies have facilitated extensive proteomic characterization of several body fluids including plasma, urine, bile, tear and hemodialysis fluid (Liu et al., 2006; Marimuthu et al., 2011; Barbhuiya et al., 2011; de
Souza et al., 2006 and Molina et al., 2005). A detailed molecular characterization of the synovial fluid could identify proteins associated with pathogenesis, which can be developed as markers for evaluation of the disease in early stages and its progression.

2.2.1 Mass spectrometry-based studies of OA and RA synovial fluid proteome

Different proteomic approaches have been used for the characterization of human synovial fluid proteome. The dynamic range of protein concentration in synovial fluid ranges from mg/mL for abundant proteins including albumin and immunoglobulins to pg/mL for molecules such as TNF-alpha (Cretu et al., 2013). The currently available proteomic analytical methods have a concentration range within the order of $10^2$-$10^5$ magnitude, resulting in less identification of low abundant proteins. In order to reduce the protein complexity of synovial fluid several strategies to deplete the abundant proteins by using depletion columns followed by deep fractionation methodologies including strong cation exchange chromatography (SCX), isoelectric focussing (IEF) or basic reverse phase liquid chromatography have been adopted (Cretu et al., 2013).

Mass spectrometry-based quantitative proteomics approach can achieve relative and absolute quantitation of proteins between different samples, resulting in the identification of differentially expressed proteins that can be potential biomarkers. Several quantitative proteomic approaches such as label free and chemical labeling strategies are available. Label free approach is semi-quantitative and is based on the peak intensity of peptides in the MS scan or on the observed spectra per peptide across the samples. As an alternative approach, chemical in vitro labeling methods including isotope coded affinity tagging (ICAT), isobaric tags for relative and absolute quantitation (iTRAQ) and tandem mass tags
(TMT) have also been used for quantitation (Cretu et al., 2013). These methods are highly suitable for the proteomic analysis of tissues and body fluids.

iTRAQ labeling is an amine group specific labeling method. The reporter ions produced during MS/MS fragmentation are detected in low mass range and allows for relative quantitation. iTRAQ labeling method is being chosen for the following reasons, primary amines in the samples are uniformly labeled, allows multiplexing of samples, unlike other labeling strategies, iTRAQ labeling all peptides co-elute and thus the signals from peptides originating from different samples are additive allowing relative quantitation.

A five-fold increase in the expression of 18 protein spots including haptoglobin was demonstrated among different synovial fluid samples from OA patients using two dimensional gel electrophoresis (2D-PAGE) platform (Yamagiwa et al., 2003). However, there were many pitfalls in 2D-PAGE analysis due to low resolution and variability between the 2-D gels (Cretu et al., 2013).

In another study, 135 proteins were identified from synovial fluid and 18 of them were shown to be differentially expressed in OA patients by using electrospray ionization tandem mass spectrometry. Proteins identified to be elevated in OA included alpha 1-microglobulin (AMBP), apolipoprotein E (APOE), complement component 3 (C3), haptoglobin (HP), orosomucoid 1 (ORM1) and group specific component (vitamin D binding protein) (GC) (Gobezie et al., 2007).

A combination of ultrafiltration and solid phase extraction methods of sample preparation was employed to achieve elimination of high concentrations of hyaluronic acid. The
samples were separated and identified by using reverse phase Nano LC in combination with LTQ-OrbitrapXL mass spectrometer. With this method, 501 peptides originating from 40 proteins were identified from healthy controls and OA patients. Proteins which were previously found to be associated with OA such as proteoglycan 4 (PRG4), collagen II (COL2A1), vimentin (VIM) and matrix Gla (MGP) were also identified in this study validating the approach used in this study (Kamphorst et al., 2007). In another study, biomarkers for OA including vitronectin V65 fragment and C3f peptide were detected in the synovial fluid by using surface enhanced laser desorption/ionization (SELDI) (de Seny et al., 2011).

In a recent study, abnormally high levels of complement components were shown in OA synovial fluid. The authors also suggest that targeting the complement components can be used a disease modifying therapy for OA patients (Wang et al., 2011). Sohn and his colleagues identified 108 proteins from OA synovial fluid and found that only 36% of them were known to be in the plasma/serum (Sohn et al., 2012). In a very recent study 66 proteins, involved in acute phase response, complement and coagulation pathways were reported to be differentially expressed between healthy and OA synovial fluid in a recent study (Ritter et al., 2013). A summary of previous proteomic studies on OA synovial fluid is provided in Table. 4.

In a proteomic study between RA and OA, SELDI mass spectrometer was used to identify proteins that were present in RA synovial fluid but not in OA. Myeloid related protein 8 was found to be specifically expressed in RA synovial fluid and was also validated by enzyme immunoassays (Uchida et al., 2002). Subsequently, in another proteomic study, 2-D gel electrophoresis followed by matrix assisted laser desorption/ionization-time of flight
(MALDI-TOF) analysis was carried out to identify synovial fluid and plasma proteins that could differentiate between RA and reactive arthritis (ReA) or OA. S100A9, S100A12 and serum amyloid A (SAA) were found to be present only in RA but not in ReA and OA in that study (Sinz et al., 2002).

A study aimed at the identification of protein biomarkers in serum and SF of RA using mass spectrometry was carried out by Liao and his group. Protein expression profiles in synovial fluid samples were compared between RA patients with erosive and non erosive disease in order to identify prognostic markers of severity of this disease. C-reactive protein (CRP) and 6 members of the S100 protein family of calcium-binding proteins were elevated in the SF of RA patients with erosive disease. Selected candidate biomarkers were verified in the sera of RA patients with erosive disease and were compared with that of healthy controls using multiple reaction monitoring mass spectrometry. CRP, S100A8 S100A9, and S100A12 proteins were found to be elevated in the serum of RA patients with erosive disease (Liao et al., 2004).

A label free quantitative analysis of RA and OA synovial fluid proteome resulted in the identification of 135 proteins using MALDI-TOF/TOF mass spectrometer. Proteins involved in complement activation, inflammation and immune response were found to be relatively more abundant in RA and those that participated in the extracellular matrix formation and remodeling were more abundant in OA synovial fluid (Mateos et al., 2012).

In another proteomic study using SELDI-TOF mass spectrometer, S100A12 was found to be overexpressed in RA synovial fluid when compared to OA synovial fluid (Han et al., 2012). An immunoproteomics study has also been reported to identify autoantigens in RA. In this study, the expression of vimentin (VIM), gelsolin (GSN), alpha-2-HS-glycoprotein
(AHSG), glial fibrillary acidic protein (GFAP) and alpha-1-B glycoprotein (A1BG) was found to be reported to be higher in RA synovial fluid than that of OA and were also validated by western blot (Biswas et al., 2013).

Most of these investigations that have been described above were performed using low resolution mass spectrometers and with minimal fractionation of the samples, which limited the depth of coverage.

Despite significant advances towards the understanding of the pathophysiology of RA and OA, early diagnosis and therapeutic intervention remain a challenge (da Mota et al., 2009). Thus, there is a need for continued discovery efforts to identify novel biomarkers with the desired sensitivity and specificity for RA and OA.
<table>
<thead>
<tr>
<th>Synovial fluid used (Healthy/Osteoarthritis)</th>
<th>Method</th>
<th>Mass spectrometer used</th>
<th>Number of proteins identified</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Healthy/Osteoarthritis</td>
<td>In-gel digestion</td>
<td>LCQ DECA XP</td>
<td>135</td>
<td>Gobezie, R <em>et al.</em>, 2007</td>
</tr>
<tr>
<td>2 Osteoarthritis</td>
<td>Depletion of albumin &amp; IgG, IEF, In-gel digestion</td>
<td>XCT Ultra Ion Trap</td>
<td>108</td>
<td>Sohn, DH <em>et al.</em>, 2012</td>
</tr>
<tr>
<td>3 Healthy/Osteoarthritis</td>
<td>2D-DIGE</td>
<td>-</td>
<td>66</td>
<td>Ritter, SY <em>et al.</em>, 2013</td>
</tr>
<tr>
<td>4 Healthy/Osteoarthritis</td>
<td>Ultrafiltration and Solid phase extraction</td>
<td>LTQ XL-Orbitrap</td>
<td>40</td>
<td>Kamphorst, JJ <em>et al.</em>, 2007</td>
</tr>
<tr>
<td>5 Osteoarthritis</td>
<td>Depletion of albumin &amp; IgG, 2-DE</td>
<td>-</td>
<td>18</td>
<td>Yamagiwa, H <em>et al.</em>, 2003</td>
</tr>
<tr>
<td>6 Healthy/Osteoarthritis</td>
<td>IEF, 2D-DIGE</td>
<td>-</td>
<td>12</td>
<td>Wang, Q <em>et al.</em>, 2012</td>
</tr>
<tr>
<td>7 Osteoarthritis</td>
<td>Protein chip array</td>
<td>SELDI-TOF-MS</td>
<td>4</td>
<td>de Seny, D <em>et al.</em>, 2011</td>
</tr>
</tbody>
</table>

Abbreviations: IEF: Isoelectric focussing; DIGE: Differential in gel electrophoresis