Chapter 4
Methodology
4.1. Identification of differentially expressed transcripts/mRNAs between RA and OA by DNA microarrays

4.1.1. Sample collection

Approximately 2-4 ml of the synovial fluid samples were aspirated from the affected knees of RA and OA patients in vacutainers (Becton, Dickinson and Company, New Jersey, USA) coated with sodium heparin and stored at 4°C. Only samples collected without any blood contamination during aspiration were included in this study. All the samples were collected from patients who satisfied the American College of Rheumatology criteria for RA and OA (Arnett et al., 1988 and Altman et al., 1986). The samples were obtained after getting informed consent from the patients and approval from the ethical committees of the Armed Forces Medical College, Pune, India and Fortis Hospitals, Bangalore, India. Normal synovial fluid samples were not collected due to ethical concerns.

4.1.2 RNA isolation and cDNA synthesis

Synovial fluid mononuclear cells were isolated from 7 RA synovial fluid samples using standard Ficoll-Paque density gradient centrifugation method (GE Healthcare Inc., Piscataway, NJ). Isolated cells were stored overnight in RNAlater (Ambion, UK) initially at 4 °C and then transferred to -80 °C until use. Using the RNeasy lipid tissue kit (Qiagen, Valencia, CA), RNA was isolated from ~5 X 10⁶ cells per sample. Quality and integrity of total RNA were assessed by determining the RNA Integrity Number (RIN) value using the 2,100 Bioanalyzer platform (Agilent Technologies, Palo Alto, CA). Subsequently, the low RNA input linear amplification kit (Agilent Technologies, Palo Alto, CA) was used to synthesize cDNA from the total RNA input from each sample. The cDNA products were
then used in an *in vitro* transcription reaction in the presence of Cy3-CTP to generate Cy3-labeled cRNAs to perform a one-color microarray experiment (Agilent Technologies, Palo Alto, CA). The labeled cRNA molecules were then purified using RNeasy spin columns (Qiagen, Valencia, CA) to remove excess free nucleotides. All samples with specific activity >8.0 were considered suitable for hybridization.

### 4.1.3 Hybridization and scanning

The Cy3-labeled cRNAs in appropriate buffers were hybridized onto 44K whole human genome oligonucleotide microarrays (Agilent Technologies, Palo Alto, CA) for 17 hrs at 65°C. Subsequently, the arrays were washed according to manufacturer’s instructions followed by scanning with an Agilent microarray scanner (G2505B) (Agilent Technologies, Palo Alto, CA). The images were processed and analyzed using the feature extraction software (version 10).
4.2. Identification and validation of proteins present in the synovial fluid of osteoarthritis patients by high resolution mass spectrometry

4.2.1 Sample collection and processing

Synovial fluid samples were collected from the affected joints of 10 OA and 10 RA patients, clinically diagnosed as per the criteria of American College of Rheumatology (Arnett et al., 1988 and Arman et al., 1986). 10 samples each from OA and RA patients included 7 females and 3 males with an average age of 65 and 52 years respectively. Out of these 10, 5 samples were used for OA cataloging and iTRAQ experiments and the rest were used for the validation experiments. Around 5 ml of synovial fluid was aspirated from each patient in heparin containing BD vacutainers (Becton, Dickinson and Company, New Jersey). The synovial fluid was then centrifuged at 1,500g for 15 minutes and the supernatants were then filtered by using 0.22μm filters (Catalog number: SLGV033RS Millipore, Massachusetts, USA) and stored at -80°C until further processing. Twelve mg of protein isolated from five OA synovial fluid samples was pooled and depleted using Human 6-Multiple Affinity Removal LC Column (MARS-6) (Agilent Technologies, Santa Clara, USA) as per manufacturer’s instructions. The depleted synovial fluid samples from each round were pooled and their protein concentration was estimated by Lowry’s method (Lowry et al., 1951). Protein from the depleted and pooled protein sample was subsequently fractionated by SDS-PAGE at protein level and by, strong cation exchange (SCX) chromatography and pl-based OFFGEL electrophoresis at peptide level.
4.2.2 SDS-PAGE and in-gel digestion

300 µg of OA synovial fluid protein depleted of abundant proteins was resolved on a 10% SDS-PAGE (16X18cm). The gel was then stained using colloidal Coomassie blue. Twenty eight gel bands were excised and destained using 40 mM ammonium bicarbonate in 40% acetonitrile (ACN). The sample was subjected to reduction using 5 mM DTT (60°C for 45 minutes) followed by alkylation using 20 mM iodoacetamide (room temperature for 10 min in dark). Trypsin digestion was carried out at 37°C for 12-16 hrs (Catalog number: V5111 Sequencing grade, Promega, Madison, WI, US). Peptides were extracted from gel pieces sequentially using 0.4% formic acid in 3% ACN twice, once using 0.4% formic acid in 50% ACN and once using 100% ACN. The extracted peptides were dried and stored at -80°C until LC-MS/MS analysis. SDS-PAGE analysis of the depleted OA synovial fluid is provided in Figure 6.

![Figure 6. SDS-PAGE analysis of depleted OA synovial fluid](image)
4.2.3 In-solution digestion

Five hundred µg of depleted synovial fluid protein was reconstituted in 40 mM ammonium bicarbonate. It was then reduced (5mM DTT), alkylated (20mM iodoacetamide (IAA)) and digested overnight using trypsin as mentioned above.

4.2.4 Strong cation exchange (SCX) chromatography

200µg of digested peptide mixture was acidified using 1 M phosphoric acid and equilibrated with 10 mM potassium phosphate buffer containing 25% acetonitrile, pH 2.85 (solvent A) and fractionated using SCX on a Polysulfoethyl A column (PolyLC, Columbia, MD) (300 Å, 5 µm, 100 × 2.1 mm) using an Agilent 1200 HPLC system (Agilent Technologies, Santa Clara, USA) containing a binary pump, UV detector and a fraction collector. The peptides were eluted using a salt gradient (0 to 100%) between solvent A and solvent B (10 mM potassium phosphate buffer containing 25% acetonitrile, 350 mM KCl, pH 2.85). Fractions were collected every minute and the elution profile was obtained by recording the absorbance at 214 nm. 89 fractions were collected in total which were then pooled into 26 fractions based on the absorbance. The pooled fractionations were then completely dried, reconstituted in 0.1% trifluoroacetic acid, and further desalted using stage-tips packed with C18 material (Rappsilber et al., 2007). Desalted fractions were dried in speedvac and reconstituted in 10 µl of 0.1% TFA prior to reversed-phase (RP) liquid chromatography based tandem mass spectrometry (LC-MS/MS) analysis. A chromatographic profile of depleted OA synovial fluid by SCX-based fractionation is shown in Figure 7.
4.2.5 OFFGEL fractionation

300 μg of in-solution digested depleted tryptic peptides was used for isoelectric point based fractionation using Agilent’s 3100 OFFGEL fractionator (Agilent Technologies, Santa Clara, USA). As per the manufacturer’s protocol, peptides were separated using pH 3-10 IPG strip. The peptides were focused for 50kVh with maximum current of 50μA and maximum voltage set to 4000V. Twelve fractions were collected after fractionation and then acidified using 1% TFA prior to sample cleaning using stage-tips as described earlier (Rappsilber et al., 2007).

4.2.6 Lectin affinity enrichment

Approximately 10 mg of the total protein pooled from five OA samples was diluted in 10 mM phosphate buffer, pH 7.8. For glycoprotein enrichment, the samples were incubated with a mixture of three agarose conjugated lectins- concanavalin A (Con A), wheat germ
agglutinin and jacalin (Vector labs, USA) for 12 h at 4°C. The beads were then washed three times using wash buffer (10 mM phosphate buffer, pH 7.8) and the bound proteins were eluted using a mixture of carbohydrates (100 mM each of N-acetylglucosamine, melibiose and galactose). The eluate was dialyzed to remove free sugars and then concentrated using 3 kDa cut-off filters. The protein concentration was estimated by Lowry's method (Lowry et al., 1951). Two hundred and fifty μg of the enriched protein fraction was then resolved by SDS-PAGE. Twenty six gel bands were excised and subjected to in-gel trypsin digestion procedure as described in the previous section. Two hundred and fifty μg of the enriched glycoprotein was also subjected to SCX fractionation as described earlier in section. Twenty fractions were collected and desalted using stage tips as mentioned above.

4.2.7 LC-MS/MS analysis

Tandem mass spectrometric analysis of 112 fractions obtained from depleted total proteome and enriched glycoproteome was carried out using LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) interfaced with Agilent 1200 (Agilent technologies, Santa Clara, CA, USA) nano liquid chromatography system. The LC system consisted of an enrichment column (3 cm × 75μm, C18 material 5μ particle size, 100 Å pore size) and an analytical column (10 cm × 75μm, C18 material C18 material 5μ particle size, 100 Å pore size) packed using pressure injection cell. Electrospray ionization source was fitted with an emitter tip 8 μm (New Objective, Woburn, MA) and maintained at 2000 V ion spray voltage. Peptide samples were loaded onto an enrichment column in 0.1% formic acid, 5% ACN for 15 min and peptide separation carried out using a linear gradient of 7-35% solvent B (90% ACN in 0.1%
formic acid) for 60 minutes at a constant flow rate of 350 nl/min. Data was acquired using Xcalibur 2.1 (Thermo Scientific, Bremen, Germany). The MS spectra were acquired in a data-dependent manner in the m/z range of 350 to 1800 and survey scans were acquired in Orbitrap mass analyzer at a mass resolution of 60,000 at 400 m/z. The MS/MS data was acquired in Orbitrap mass analyzer at a resolution of 15,000 at 400 m/z by targeting top 20 most abundant precursor ions for fragmentation using higher energy collisional dissociation activation at 39% normalised collision energy. Single and unassigned charge state precursor ions were rejected. The dynamic exclusion option was enabled during data acquisition with exclusion duration of 60 seconds. Lock mass option was enabled for real time calibration using polycyclodimethylsioxane (m/z, 445.12) ions (Olsen et al., 2005).

4.2.8 Data analysis

Mass spectrometry data was analyzed using multiple search engines to maximize the peptide identifications. Proteome Discoverer 1.3 (Thermo Scientific, Bremen, Germany) was used to carry out the peak list generation and database searches. Precursor mass range of 500 to 8,000 Da and signal to noise ratio of 1.5 were used as the criteria for generation of peak list files. NCBI RefSeq 49 human protein database with known contaminants (32,967 entries) was used as a reference database. Sequest and Mascot algorithms were used to carry out database searches. The parameters used for database searches include trypsin as a protease with allowed one missed cleavage, carbamidomethyl cysteine as a fixed modification, and oxidation of methionine as a dynamic modification. Precursor ion mass error window of 20 ppm and fragment ion mass error window of 0.1 Da were allowed. The raw data obtained were searched against decoy database to calculate 1% false discovery rate cut-off score (Kandasamy et al., 2009). Spectra that matched to the
contaminants and those that did not pass the 1% FDR threshold were not considered for analysis.

4.2.9 Bioinformatics analysis

Gene Ontology (GO) (Ashburner et al., 2000) analysis was done to identify the biological processes and the molecular function associated with the identified proteins. Sub-cellular localization, PTMs, transmembrane domain and signal peptide information of the identified proteins were obtained from Human Protein Reference Database (HPRD) (http://www.hprd.org), which is a GO compliant database (Prasad et al., 2009).

4.2.10 Multiple reaction monitoring (MRM)

Multiple reaction monitoring-based assays were developed to validate the results of LC-MS/MS analysis for three target proteins. Skyline 2.1 (Mac Lean et al., 2010) was used for method development, data analysis and interpretation of the MRM results. Proteotypic peptides for each protein were selected from the discovery phase LC-MS/MS data. Preference was given to proteotypic peptides with precursor charge +2 that did not contain cysteine or methionine. A minimum of four transitions were monitored for each peptide. Equal protein amounts from 10 OA synovial fluid samples were subjected to trypsin digestion as described earlier. MRM of each sample was carried out in triplicates on TSQ Quantum Ultra (Thermo, San Jose, CA) interfaced with Easy nanoLC II (previously Proxeon, Thermo Scientific, Bremen, Germany). Peptides were enriched on a trap column (5μm, 75μm × 2cm) for 5 minutes with solvent A (5%ACN in 0.1% formic acid). The peptides were separated on analytical column (3μm, 75μm × 10cm) with a linear gradient
of 7-35% solvent B (95%ACN in 0.1% formic acid) for 60 min at a constant flow rate of 300 nl/min. Both columns were packed in-house using Magic C18 AQ (Michrom Bioresources). Spray voltage of 2.5 kV was applied and ion transfer tube was maintained at 275 °C. MRM data was acquired with Q1 and Q3 set at resolution of 0.4 and 0.7 respectively. The collision energy for each transition was optimized in Skyline based on the preliminary results.
4.3. Identification and validation of synovial fluid proteins differentially expressed between RA and OA patients by high resolution mass spectrometry

4.3.1 Depletion of synovial fluid proteome and iTRAQ labeling

Protein estimation of the synovial fluid samples was carried out using Lowry’s method [89]. Equal amounts of protein from five samples each of OA and RA synovial fluid were pooled separately. The pooled samples were then depleted to remove the 14 most abundant proteins (Albumin, haptoglobin, transferrin, IgA, IgG, alpha 1-antitrypsin, alpha 2-antitrypsin, alpha 1-acid glycoprotein, apolipoprotein A1, apolipoprotein A2, complement C3, IgM, transthyretin and fibrinogen) by using Human 14 multiple affinity removal spin cartridge (Agilent Technologies, Santa Clara, California, USA). The depleted protein was then washed and concentrated using 3kDa MWCO filters (Amicon, Millipore, Ireland). Approximately 65μg equivalent of the depleted synovial fluid protein from each group was subjected to trypsin digestion and iTRAQ labeling as described earlier (Pawar et al., 2011). Briefly, denaturation of the protein was done using 2% SDS followed by reduction and alkylation with reducing agent tris (2-carboxyethyl) phosphine (TCEP) and cysteine blocking agent, methyl methanethiosulfonate (MMTS), respectively. Subsequently the samples were digested with sequencing grade trypsin (Promega, Madison, WI, USA) at 37°C overnight. The tryptic peptides from each group were then labeled with 4-plex iTRAQ reagents (iTRAQ Reagents Multiplex kit, Applied Biosystems, California, USA) as per the manufacturer’s instructions. OA and RA synovial fluid derived tryptic peptides were labeled with 116 and 117 iTRAQ labels respectively. The labeled peptides were then pooled, vacuum-dried and reconstituted in 10mM KH₂PO₄, 20% acetonitrile (pH 2.8) (solvent A) and fractionated by strong cation exchange (SCX) chromatography. SCX chromatography and LC-MS/MS analysis were carried out as described earlier. A typical
SCX-based chromatographic profile of the iTRAQ labeled RA and OA synovial fluid is shown in Figure 8.

### 4.3.2 Data analysis

Proteome Discoverer Beta Version 1.3 (Thermo Fisher Scientific Inc., Bremen, Germany) was used for database searches. A precursor mass range of 350-8000 Da and a signal to noise of 1.5 were used. A combined Mascot (Mascot version 2.2, Matrix Science) and SEQUEST search was done using the Proteome Discoverer suite (Version 1.3.339, Thermo Scientific, Bremen, Germany) against the NCBI Human RefSeq database 50 containing 33,947 entries with known contaminants. Search parameters included trypsin as the enzyme with maximum 1 missed cleavage allowed; oxidation of methionine was set as a dynamic modification while alkylation at cysteine and iTRAQ modification at N-terminus of the peptide and lysine were set as static modifications. Precursor and fragment mass tolerance were set to 20 ppm and 0.1 Da, respectively. Peptide and protein data were fetched using high peptide confidence and rank one peptide match filters. Reporter ion quantitation node was used for relative expression pattern of proteins based on the relative intensities of reporter ions for the corresponding peptides. The raw data obtained was searched against decoy database to calculate 1% false discovery rate cut-off score (Kandasamy et al., 2009). Spectra that matched to the contaminants and those that did not pass the 1% FDR threshold were not considered for analysis. Bioinformatics analysis was carried out as described in the earlier section.
4.3.3 GeneSpring analysis

Pathway Architect module from GeneSpring GX12 was used to identify significantly enriched pathways. Using ‘single experiment analysis’ tool in the pathway architect module the differentially expressed list of genes was searched against publicly available pathways.

4.3.4 Multiple reaction monitoring (MRM)

MRM assays (Lange et al., 2008) were designed to validate the differentially expressed protein, CAPG in RA and OA synovial fluid. Skyline version 2.1 (MacLean et al., 2010) was used for method development and optimization. The target peptide selected for CAPG was QAALQVAEGFISR (z = +2, m/z = 695.38) and top four transitions monitored included, y8→ 878.47, y7→779.40, y6→708.37 and y5→579.32. The samples were
subjected to in-solution digestion as described earlier. All samples were analyzed in triplicate on TSQ Quantum Ultra (Thermo Scientific, San Jose, CA) interfaced with Easy nanoLC II (previously Proxeon, Thermo Scientific, Bremen, Germany). The peptides were enriched on a trap column (5 μm, 75 μm × 2 cm.) with 0.1% formic acid and 5% ACN for 5 minutes and separated on an analytical column (3 μm, 75 μm × 10 cm) with an increasing linear gradient from 5-35% of solvent B (90% ACN in 0.1% formic acid) for 60 min at a constant flow rate of 300 nl/min. Both columns were packed in-house using Magic AQ C18 material (Michrom Bioresources). Spray voltage of 2.5 kV was applied and ion transfer tube was maintained at 275 °C. The data was acquired with Q1 and Q3 set at resolution of 0.4 and 0.7 respectively. The collision energy for each transition was optimized with the help of Skyline based on the preliminary results (MacLean et al., 2010).

4.3.5 Western blot analysis

40 μg of protein from each individual samples of RA (n=10) and OA (n=10) synovial fluid was used for Western blot analysis. The primary antibody used was a rabbit anti-human polyclonal antibody for CAPG (Proteintech Group, Inc, Chicago, USA). SDS-PAGE gels were electroblotted on nitrocellulose membrane (Whatman Inc., Maine, USA) at 250 mA for 2 hours using TE-70 ECL semi-dry transfer unit (GE Healthcare, Pittsburgh, USA). The membranes were then blocked with 5% non-fat dry milk and washed with phosphate buffered saline containing 0.05% tween (PBST). The membranes were incubated with primary antibody for 2 hours, washed with PBST and incubated at room temperature for 1 hour in diluted (1:2500) anti-rabbit IgG antibody conjugated with horseradish peroxidase (GE Healthcare, UK).