



A Comprehensive Gene Expression Meta-analysis Identifies Novel Immune Signatures in Rheumatoid Arthritis Patients

Sumbul Afroz^{1†}, Jeevan Giddaluru^{1†}, Sandeep Vishwakarma², Saima Naz², Aleem Ahmed Khan² and Nooruddin Khan^{1*}

¹Department of Biotechnology and Bioinformatics, School of Life Sciences, University of Hyderabad, Hyderabad, India,

²Centre for Liver Research and Diagnostics, Central Laboratory for Stem Cell Research and Translational Medicine, Deccan College of Medical Sciences, Kanchanbagh, Hyderabad, India

OPEN ACCESS

Edited by:

Heiko Mühl,
Goethe University Frankfurt, Germany

Reviewed by:

Paul Proost,
K.U. Leuven, Belgium
John Isaacs,
Newcastle University, UK
João Eurico Fonseca,
Lisbon University, Portugal

***Correspondence:**

Nooruddin Khan
noor@uohyd.ac.in

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted
to Inflammation,
a section of the journal
Frontiers in Immunology

Received: 25 October 2016

Accepted: 17 January 2017

Published: 02 February 2017

Citation:

Afroz S, Giddaluru J, Vishwakarma S,
Naz S, Khan AA and Khan N (2017)
A Comprehensive Gene Expression
Meta-analysis Identifies Novel
Immune Signatures in Rheumatoid
Arthritis Patients.
Front. Immunol. 8:74.
doi: 10.3389/fimmu.2017.00074

Rheumatoid arthritis (RA), a symmetric polyarticular arthritis, has long been feared as one of the most disabling forms of arthritis. Identification of gene signatures associated with RA onset and progression would lead toward development of novel diagnostics and therapeutic interventions. This study was undertaken to identify unique gene signatures of RA patients through large-scale meta-profiling of a diverse collection of gene expression data sets. We carried out a meta-analysis of 8 publicly available RA patients' (107 RA patients and 76 healthy controls) gene expression data sets and further validated a few meta-signatures in RA patients through quantitative real-time PCR (RT-qPCR). We identified a robust meta-profile comprising 33 differentially expressed genes, which were consistently and significantly expressed across all the data sets. Our meta-analysis unearthed upregulation of a few novel gene signatures including *PLCG2*, *HLA-DOB*, *HLA-F*, *EIF4E2*, and *CYFIP2*, which were validated in peripheral blood mononuclear cell samples of RA patients. Further, functional and pathway enrichment analysis reveals perturbation of several meta-genes involved in signaling pathways pertaining to inflammation, antigen presentation, hypoxia, and apoptosis during RA. Additionally, *PLCG2* (phospholipase Cγ2) popped out as a novel meta-gene involved in most of the pathways relevant to RA including inflammasome activation, platelet aggregation, and activation, thereby suggesting *PLCG2* as a potential therapeutic target for controlling excessive inflammation during RA. In conclusion, these findings highlight the utility of meta-analysis approach in identifying novel gene signatures that might provide mechanistic insights into disease onset, progression and possibly lead toward the development of better diagnostic and therapeutic interventions against RA.

Keywords: rheumatoid arthritis, microarrays, meta-analysis, synovial inflammation, autoimmunity

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, progressive, and inflammatory autoimmune disease, which continues to cause global disability having a worldwide prevalence of 1% (1). The disease is classified primarily by clinical phenotype and predominantly associated with articular, extra-articular, and systemic complications including anemia, cardiovascular diseases, osteoporosis, fatigue, and

depression (2, 3). There exists a considerable geographical and temporal variability as well as unpredictability based on severity and manifestations in the occurrence of RA within populations (4). RA is characterized by synovial inflammation [excessive infiltration and activation of neutrophils, mononuclear cells including T cells, B cells, plasma cells, and mast cells (5), as well as complement factors such as C5a at rheumatoid joints (6)], autoantibody generation (rheumatoid factor and anti-citrullinated protein antibody) (3) and degradation of bones and cartilage leading to bone deformity (5). The etiology or cause of RA remains elusive. However, various studies have illustrated a convoluted interplay between genetic factors and environmental exposures in most of the disease cases (3, 7). Unprecedented efforts have been made to identify multitude of factors that have foremost allusions in RA pathogenesis including pro-inflammatory cytokines (TNF- α , IL-6, IL-17, and IL-1) (8) as the orchestrator of vital processes during RA progression involving synovitis (9), angiogenesis (production of pro-angiogenic factors particularly VEGF) (10, 11), osteoclastogenesis (increased production of macrophage colony-stimulating factor and RANK ligand) (12), cartilage degradation (enhanced secretion of MMPs into synovial fluid at the joints) (9, 13), and acute-phase proteins production such as CRP (exacerbates RA-related tissue damage) (14).

On the genetic front, the strongest link in RA has been ascribed to the third hypervariable region of the HLA-DR β chains (aa. 70–74) within MHC. The active epitope within the area, glutamine-leucine-arginine-alanine-alanine (QKRAA) or QRRAA has been found to be associated with multiple RA allied DR genes including DR1, DR4, and DR14 (15). The location of the epitope within HLA-DR determines the specificity of the peptides presented to CD4 $^{+}$ T-cells. However, the specific peptides that bind to DR proteins in RA patients have not yet been identified (6). The abundance of effector T cells and IL-17-producing Th17 cells in the synovial milieu emphasize the critical role of these cells in RA pathogenesis. The activation of T-cells through HLA-DR4 (present on APCs) mediated presentation of specific peptides during RA initiates an array of complex immune responses including T-cell oligoclonality, germinal center reactions, and B-cell hypermutation highlighting local antigen-specific T-cell-mediated B-cell responses (3). The appropriate interconversion between Th1/2/17/Treg phenotypes has crucial implication toward RA outcome (16). In particular, an imbalance between Th17 and Treg cells has an impact on local TNF levels in the synovial fluid that blocks the differentiation of Treg cells and along with IL-17, macrophage, and dendritic cells derived TGF- β and IL-1 β , shifts the T-cell homeostasis toward inflammation (17). The importance of B-cells in RA have been highlighted in autoantigen presentation and cytokine production (IL-6, TNF, and lymphotoxin- β) (3). Apart from acting as the precursor for autoantibody-generating plasma cells, B-cells are the only antigen-presenting cells supporting the activation of autoreactive T-cells during RA (18).

Although several factors have been identified, yet there is a dearth of knowledge about the complex gene networks associated with the disease. Over the last two decades, DNA microarray technology allowed us to interrogate thousands of

genes simultaneously enrooting the discovery of disease-relevant genes. Multiple microarray experiments have been conducted to identify potential gene signatures that are responsible for the pathogenesis of various diseases. The analyses of these experiments usually generate hundreds of differentially expressed genes (DEGs), making it difficult to generate a convincing transcriptome profile of a particular disease condition. To address this issue, we implemented a meta-analysis method to extract a global transcriptional profile of RA patients available on public databases. Our gene expression meta-analysis revealed some credible gene signatures that have shown relevance to RA condition. Finally, we went a step further to validate these gene signatures in RA patient samples obtained clinically. In conclusion, this meta-analysis approach of publicly available microarray data suggests us to generate a significant gene expression profile that might represent an early step toward enhancement in the detection, treatment, and prevention of RA.

MATERIALS AND METHODS

Data Sets Review

A comprehensive search was administered for RA gene expression data sets available on Gene Expression Omnibus (GEO) database. Independent of age, gender, race, and region, two sample groups, i.e., RA patients and healthy controls (HCs) were considered for this study. After a thorough search, we identified eight data sets, out of which two data sets (GSE25160, GSE42296) (19, 20) had a sample source of peripheral blood mononuclear cells (PBMCs) and six data sets (GSE1919, GSE12021, GSE48780, GSE55235, GSE55457, GSE55584) (21–24) had a sample source of synovial tissue. These eight data sets were subjected to independent differential gene expression analysis followed by a meta-analysis. The GEO accession IDs and samples information in each data set is summarized in Table 1.

Data Sets Analyses

The raw data (.CEL files) of each data set was downloaded from GEO accession links. As the eight data sets were of Affymetrix platform, a single normalization method (RMA) was performed on the raw data. Samples of each data set were categorized into two groups; RA patients and HCs. The differential expression of genes between the two groups (RA vs. HC) was calculated using a

TABLE 1 | GEO data sets and samples summary.

Data set ID	GEO accession	Sample source	Rheumatoid arthritis samples	Healthy samples
1	GSE1919	Synovial tissue	5	5
2	GSE12021	Synovial tissue	13	24
3	GSE25160	Peripheral blood mononuclear cells (PBMCs)	13	4
4	GSE42296	PBMCs	19	4
5	GSE48780	Synovial tissue	27	6
6	GSE55235	Synovial tissue	10	10
7	GSE55457	Synovial tissue	10	13
8	GSE55584	Synovial tissue	10	10

moderated *T*-test, assigning a specific threshold (*p*-value < 0.05). To bypass the multiple testing problems, we adjusted the *p*-values using an optimized false discovery rate (25). Genes with *q*-values (adjusted *p*-values) less than 0.05 were considered as DEGs. The DEGs of each data set were annotated with Entrez IDs, official gene symbols, and gene names. All the analyses were performed in R using various Bioconductor packages (26, 27).

Meta-analysis

Consequent to the individual data sets analyses, eight sets of DEGs were obtained and assigned to the meta-analysis study. The primary task of our meta-analysis method was to extract the intersected DEGs among the set of DEGs. The step by step automated meta-analysis method was as follows: (i) extract the intersected genes (meta-genes), (ii) exclude genes with inconsistent expression, and (iii) check for significance by combining adjusted *p*-values of each gene from all the data sets. The weighted Z-method was used to combine the individual *q*-values of each gene, where each test was assigned as weight (w_i) and implemented using *survcomp*, an R package (28).

$$\text{Weighted } Z\text{-method, } Zw = \frac{\sum_{i=1}^k w_i Z_i}{\sqrt{\sum_{i=1}^k w_i^2}}$$

The meta-analysis algorithm was implemented in R.

GO Functional and Pathway Enrichment

The obtained 33 meta-genes were subjected to GO functional and pathway enrichment analysis. Biological processes and molecular functions of meta-genes were generated using Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>), a web-based online tool (29). Pathway enrichment analysis of meta-genes was performed using ReactomeFIViz (30), a Cytoscape plug-in based on Reactome pathway database (31).

Specimen Collection

Blood samples were collected from the individuals clinically diagnosed with RA reporting to Owaisi Hospital and Research Centre, Hyderabad, India. The procedure of the entire study was carried out by the protocols approved by the Institute Ethics Committee, University of Hyderabad. Written informed consent was taken from each patient, and the complete clinical information and medical history was well documented.

PBMCs Isolation

For PBMCs isolation, blood samples were collected in the sterile EDTA Vacutainer blood collection tubes. PBMCs from the whole blood of HCs and RA patients were isolated using Ficoll Histopaque (Sigma Aldrich Company, UK) density gradient method as described earlier (32). Briefly, 1.5 ml of whole blood was diluted with 1× PBS-EDTA in a 1:1 ratio, was layered upon 1 ml Ficoll, and centrifuged at 400 × g for 30 min. The buffy layer was carefully removed and washed with 1× PBS-EDTA twice at 200 × g for the removal of blood platelets. Isolated PBMCs were suspended in TRIZOL (Invitrogen, Life Technologies) reagent for RNA extraction.

Quantitative Real-Time PCR (RT-qPCR)

Total RNA from isolated PBMCs was extracted using TRIZOL reagent. RNA was reverse transcribed into cDNA using VERSO cDNA synthesis kit (Thermo Scientific) according to manufacturer's instructions. As template, approximately 30 ng cDNA was used for RT-qPCR and the signals were detected using real-time PCR system. Quantitative real-time PCR was performed using Mastercycler ep realplex (Eppendorf). The cDNA was amplified using SYBR Green Mix (Kapa Biosystems) with gene-specific primers (Table S4 in Supplementary Material). The thermal cycler parameters followed are as follows: one cycle of 94°C for 2 min followed by 40 cycles of 30 s at 94°C, 30 s annealing at 56°C and 40 s at 68°C. The relative mRNA expression of each gene was calculated using housekeeping gene β-actin as the reference gene (33).

RESULTS

Computational Analysis

Preliminary Analyses

Followed by a complete systemic review of RA data sets available on GEO database, we downloaded raw data of eight data sets. Sample source of six data sets was of synovial tissue, and the remaining two data sets were of PBMCs (Table 1). In total, the study included 107 RA patients and 76 HCs samples. Before the meta-analysis study, we analyzed and generated the results of individual data sets to obtain eight set of DEGs. The raw data of each data set were normalized using RMA method. Later, the differential expression of each gene in each study was calculated for patients group with respect to the healthy group using a *T*-test statistic. Genes with *p*-values less than 0.05 were assigned as DEGs. Adjusted *p*-values (*q*-values) were calculated using Benjamini Hochberg's method (25) to eliminate the false positives that arise due to multiple testing. The DEGs of each study were annotated with Entrez IDs, official gene symbols, and gene names. A total of 245,672 were measured, out of which 84,436 have shown a significant change in expression. We obtained eight sets of DEGs that were assigned for our meta-analysis study. The schematic representation of the analyses method is shown in Figure 1.

Meta-analysis

The purpose of performing meta-analysis was to elicit the intersection of significant gene signatures (meta-profile) from the sets of DEGs obtained during our preliminary analyses. Despite the availability of various established microarray meta-analysis methods for the detection of common DEGs (meta-genes) (34), we sought to use a specific meta-analysis approach, where we combined the statistical parameters (*q*-values) obtained from differential gene expression analysis of individual data sets. After our preliminary analyses, we extracted DEGs that were commonly expressed across all data sets. We identified 78 genes that were commonly expressed in all eight data sets (Figure 2A), out of which 33 genes were consistently expressed (either upregulated or downregulated) in at least seven out of eight data sets (Figure 2B). In contrast, 34 out of remaining 45 genes have shown a consistent reverse expression between PBMCs and synovial tissue data sets

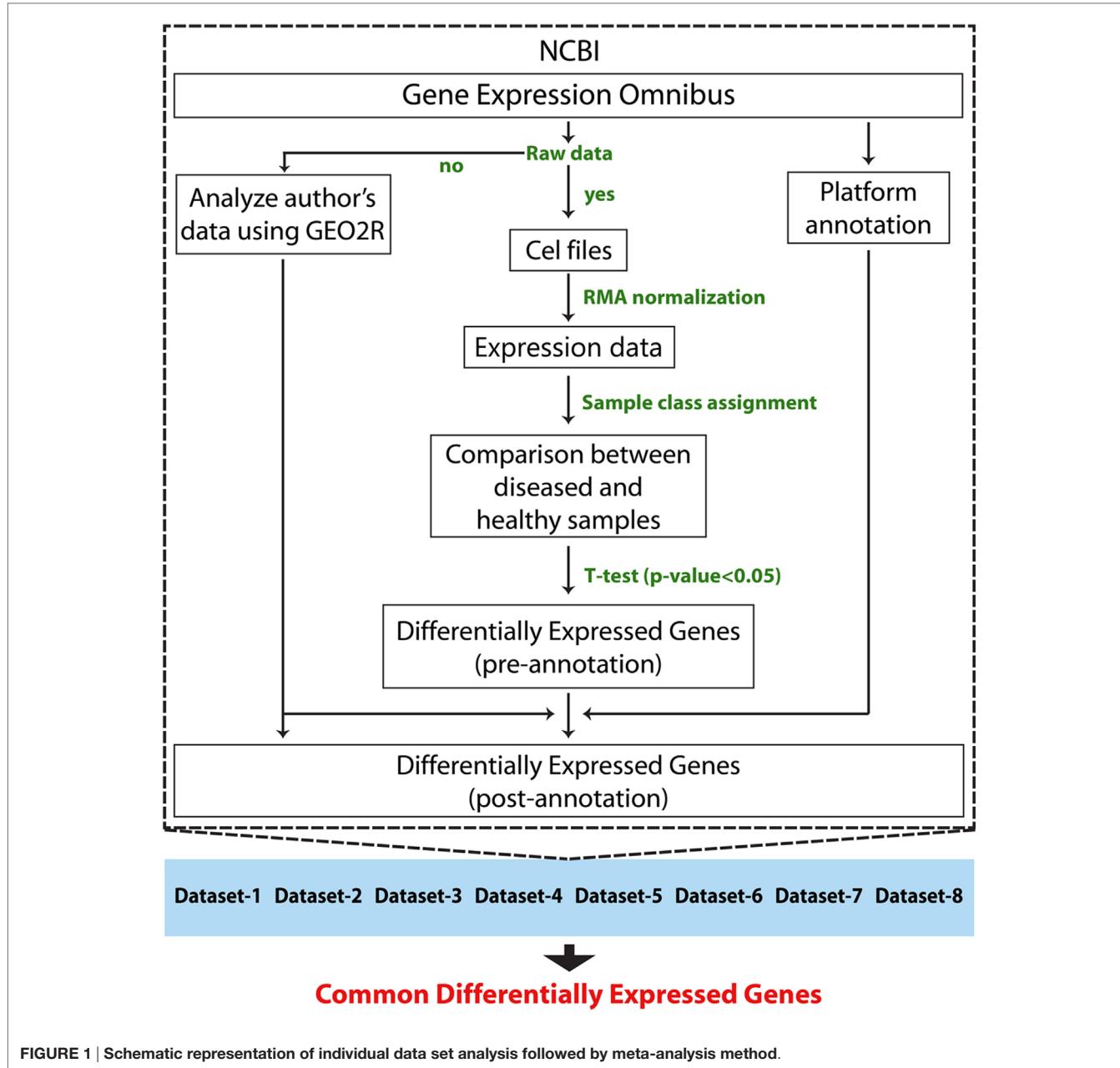


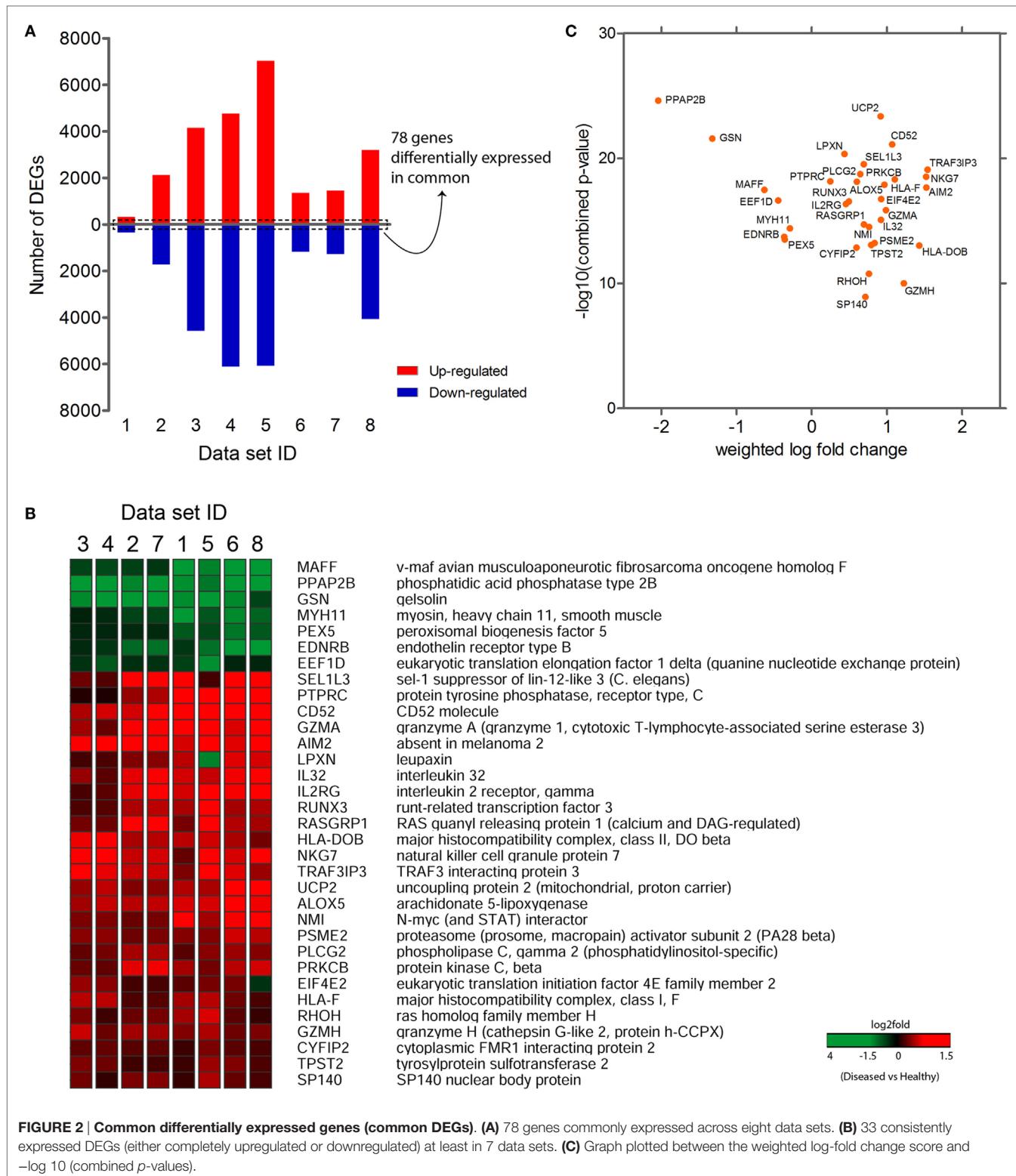
FIGURE 1 | Schematic representation of individual data set analysis followed by meta-analysis method.

(Figure S1 in Supplementary Material). This variation might be possible due to the difference in cellular tropism and secretion of various kinds of chemokines and cytokines in the synovium as compared to PBMCs. Also, the overall systemic responses occurring in blood against the localized responses generated in synovium might influence the expression level of a particular gene in both these areas. In the next step, we computed combined p -values of consistently expressed genes using weighted Z-method from the obtained independent q -values of each data set. The genes were ranked based on the combined p -values, thus evaluating the significance level of each gene. Also, a weighted log-fold change score was calculated for the list of genes and were plotted against the combined p -values (Figure 2C). In total, we

observed 7 downregulated and 26 upregulated genes in the list of consistently expressed genes that were finally assigned as our candidate meta-signatures. As an adjunct analysis, a meta-analysis was performed for seven data sets (excluding the data set with least number of DEGs, i.e., data set 1) to obtain additional number of common DEGs (Figure S2 in Supplementary Material).

GO Enrichment Analysis

We performed a GO enrichment analysis to the 33 meta-signatures using Enrichr (29), an online tool based on Gene Ontology. Enrichment analysis was performed for two GO categories: biological process and molecular function. The top enriched GO terms for biological process were: regulation of B cell receptor



(BCR) signaling pathway (GO: 0050855), cellular calcium ion homeostasis (GO: 0006874), mast cell activation (GO: 0045576), leukocyte activation (GO: 0045321), positive regulation of T cell-mediated cytotoxicity (GO: 0001916), cytolysis (GO:

0019835), immune response-regulating cell surface receptor signaling pathway (GO: 0002768), and regulation of protein oligomerization (GO: 0032459). The top enriched GO terms for molecular functions included peptide binding (GO: 0042277),

amide binding (GO: 0033218), translation factor activity, nucleic acid binding (GO: 0008135), type 1 angiotensin receptor binding (GO: 0031702), angiotensin receptor binding (GO: 0031701), MHC class II receptor activity (GO: 0032395), and RNA cap binding (GO: 0000339). Detailed enrichment tables are provided in Tables S1 and S2 in Supplementary Material.

Pathway Enrichment Analysis

We performed pathway enrichment using ReactomeFiViz plugin based on Reactome pathways database. A total of 20 out of 33 genes were involved in various pathways. The hit genes list comprised AIM2, ALOX5, CYFIP2, EDNRB, EEF1D, EIF4E2, GSN, HLA-DOB, HLA-F, IL2RG, MAFF, MYH11, PLCG2, PPAP2B, PRKCB, PSME2, PTPRC, RASGRP1, RHOH, and ST6GAL1. Various immune system pathways such as platelet activation, signaling and aggregation, the AIM2 inflammasome, downstream signaling events of BCR, signaling by the BCR, ER-phagosome pathway, and antigen processing cross presentation were enriched. The other distinct pathways include platelet activation, signaling and aggregation, GPVI-mediated activation cascade, hemostasis, synthesis of lipoxins (LX), synthesis of 5-eicosatetraenoic acids, disinhibition of SNARE formation, and signaling by VEGF. Top enriched pathway table is provided in Table S3 in Supplementary Material.

Validation of Meta-analysis Data

RT-qPCR Analysis of Selected Genes

Pathway enrichment analysis of the meta-gene profile revealed a total of 20 genes involved in distinct pathways. Eight of the pathways associated genes (PLCG2, AIM2, ALOX-5, HLA-DOB, HLA-F, EIF4E2, PRKCB, and CYFIP2) were further studied by quantitative real-time PCR to validate the results obtained through the meta-analysis of publicly available RA patient's microarray data sets. The genes for this in-depth study were selected by their diverse cellular functions, the magnitude of their differential expression, novelty (genes with unknown or less known functions), and for their potential relevance to disease. Quantitative RT-PCR for the genes quoted above was carried out using 19 patients clinically diagnosed with RA, and eight healthy volunteers were taken as HCs. Although the number of samples analyzed in the meta-analysis study and real-time PCR are limited, our results demonstrate that the relative expression levels analyzed through RT-qPCR for all the eight genes mentioned above were significantly upregulated in RA patient's PBMCs as compared to HCs (Figures 3A–H). Of note, AIM2 (Figure 3B) and CYFIP2 (Figure 3H) were the two genes that were most significantly upregulated in PBMC samples of RA patients as compared to HCs. Overall, the RT-qPCR results were concordant with our meta-analysis study, suggesting the role of meta-analysis approach as a stupendous tool to identify differentially expressed gene signatures that might act as novel candidates to develop new therapeutic interventions against RA. Further, pathway networks of these validated candidate signatures were generated using Pathway Commons (31, 35–58) (www.pathwaycommons.org), a web resource for biological pathway data, and visualized using Cytoscape software (59) (Figures 4 and 5).

DISCUSSION

The etiology of RA like most of the immune-mediated disorders is multifactorial, which is influenced by genetic, environmental, hormonal, and immunological factors. However, genes, gene networks associated with RA onset and progression are not clearly defined. Our meta-analysis of RA gene expression data identified a robust meta-gene profile comprising 33 DEGs, which showed a consistent and significant up- or downregulation across all the data sets. To further understand the significance of these meta-genes in RA pathogenesis, we performed a GO functional and pathway enrichment analysis, which led to the identification of novel gene signatures that went unnoticed in previous RA studies. These genes include *PLCG2*, *HLA-DOB*, *HLA-F*, *EIF4E2*, and *CYFIP2*, which have been previously implicated in inflammation, antigen presentation, hypoxia, and apoptosis. Interestingly, expressions of these genes were also found to be significantly upregulated in PBMCs of clinical RA patients. It has been well established that the severity of RA is majorly dictated by the production of an array of pro-inflammatory cytokines in the synovial joint particularly cytokines like TNF and IL-1 β , which triggers multiple cellular and immunological interactions as a result of excessive inflammation. Although biological therapies against TNF and other pro-inflammatory cytokines offer momentary clinical benefits, they are often associated with severe side effects (60). Therefore, it is decisive to identify and quantify combination of other genes, which might be helpful in controlling the massive inflammatory responses occurring during RA progression. Our meta-analysis study exemplifies the involvement of few inflammatory genes (*PLCG2*, *AIM2*, and *ALOX-5*), which might act as novel candidates for controlling inflammation during the disease condition (Figure 2B). Additionally, we have identified a previously unnoticed inflammatory gene, *PLCG2* (phospholipase C-gamma 2), which has been found to be upregulated in all the microarray data sets as well in the clinical RA patients (Figures 2B and 3A). *PLCG2*, a transmembrane signaling enzyme, catalyzes the conversion of membrane phospholipid, phosphatidylinositol 4,5 bisphosphate into second messenger molecules inositol triphosphate (IP3) and diacylglycerol (DAG) using calcium as the cofactor, thereby activating NLRP3 inflammasome. A gain of function mutation in *PLCG2* gene is associated with phospholipase $C\gamma 2$ (*PLC γ 2*)-associated antibody deficiency and immune dysregulation (APLAID), dominant inherited disorder, and auto-inflammation (61), thereby suggesting that *PLCG2* might be a potential therapeutic target for controlling inflammation during RA pathogenesis. In consent with previous studies that have highlighted the role of *AIM2* (absent in melanoma 2) and *ALOX-5* (arachidonate 5-lipoxygenase) genes in RA pathogenesis, our study also demonstrated the enhanced expression of these genes in RA patients as compared to HCs (Figures 3B,C). *AIM2* is known to induce the formation of caspase-1-activating inflammasome, thereby controlling the proteolytic maturation of pro-inflammatory cytokines IL-1 β and IL-18 (62). Conditional deletion of *AIM2* in RA mice model shows reduced inflammatory responses suggesting the role of *AIM2* in controlling inflammation during RA (62). *ALOX-5* is expressed basically in mature myeloid cells and developing/memory B-lymphocytes and is

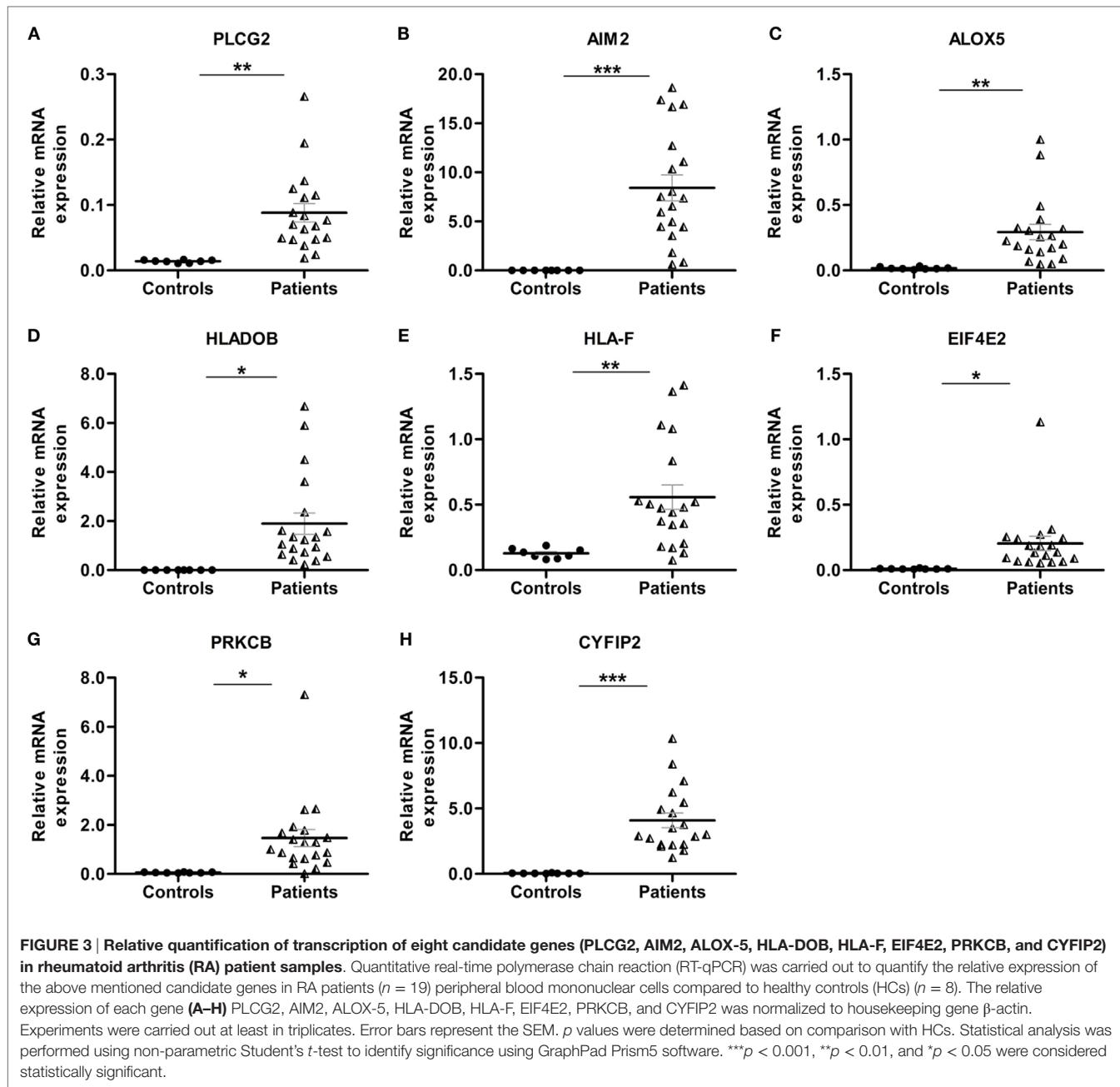


FIGURE 3 | Relative quantification of transcription of eight candidate genes (PLCG2, AIM2, ALOX-5, HLA-DOB, HLA-F, EIF4E2, PRKCB, and CYFIP2) in rheumatoid arthritis (RA) patient samples. Quantitative real-time polymerase chain reaction (RT-qPCR) was carried out to quantify the relative expression of the above mentioned candidate genes in RA patients ($n = 19$) peripheral blood mononuclear cells compared to healthy controls (HCs) ($n = 8$). The relative expression of each gene (A–H) PLCG2, AIM2, ALOX-5, HLA-DOB, HLA-F, EIF4E2, PRKCB, and CYFIP2 was normalized to housekeeping gene β -actin. Experiments were carried out at least in triplicates. Error bars represent the SEM. p values were determined based on comparison with HCs. Statistical analysis was performed using non-parametric Student's t -test to identify significance using GraphPad Prism5 software. *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$ were considered statistically significant.

responsible for the generation of excessive leukotrienes (LTB4, LTC4, LTD4, and LTE4) in the synovial fluid of RA patients (63). Inhibitors of ALOX5 gene dampen TNF- α -induced NF- κ B, thereby abrogating synovial inflammation (63). ALOX5 in coordination with 5-LO-activating protein (FLAP) leads to the synthesis of LTA4, one of the critical intermediates of leukotrienes synthesis pathway (64). The long association of human leukocytic antigen (HLA) is confirmed in patients positive for rheumatoid factor. Appropriate antigen presentation or alteration in peptide affinity play a crucial role in promoting autoreactive adaptive immune responses (3). Intriguingly, our study for the first time has documented gene signatures HLA-DOB and HLA-F, which

were found to be upregulated across all RA patient's microarray data sets as well as in RA patients' PBMCs (Figures 3D,E). HLA-DOB, a B-cell lineage, MHC-II-related molecule has been reported to have strong immunogenicity for human T-cells, one of its identified CTL epitopes HLA-DOB_{232–240} acts as a resilient immunotherapeutic candidate for targeting multiple myeloma (65). HLA-F, a non-classical MHC-I molecule, is currently enigmatic of all the HLA molecules; hence, its precise function remains elusive. Recent studies have demonstrated enhanced HLA-F expression in cancer stroma in the case of both breast and gastric cancers (66). HLA-F on B-cells was found to induce immune tolerance in tumor cells by interacting with inhibitory

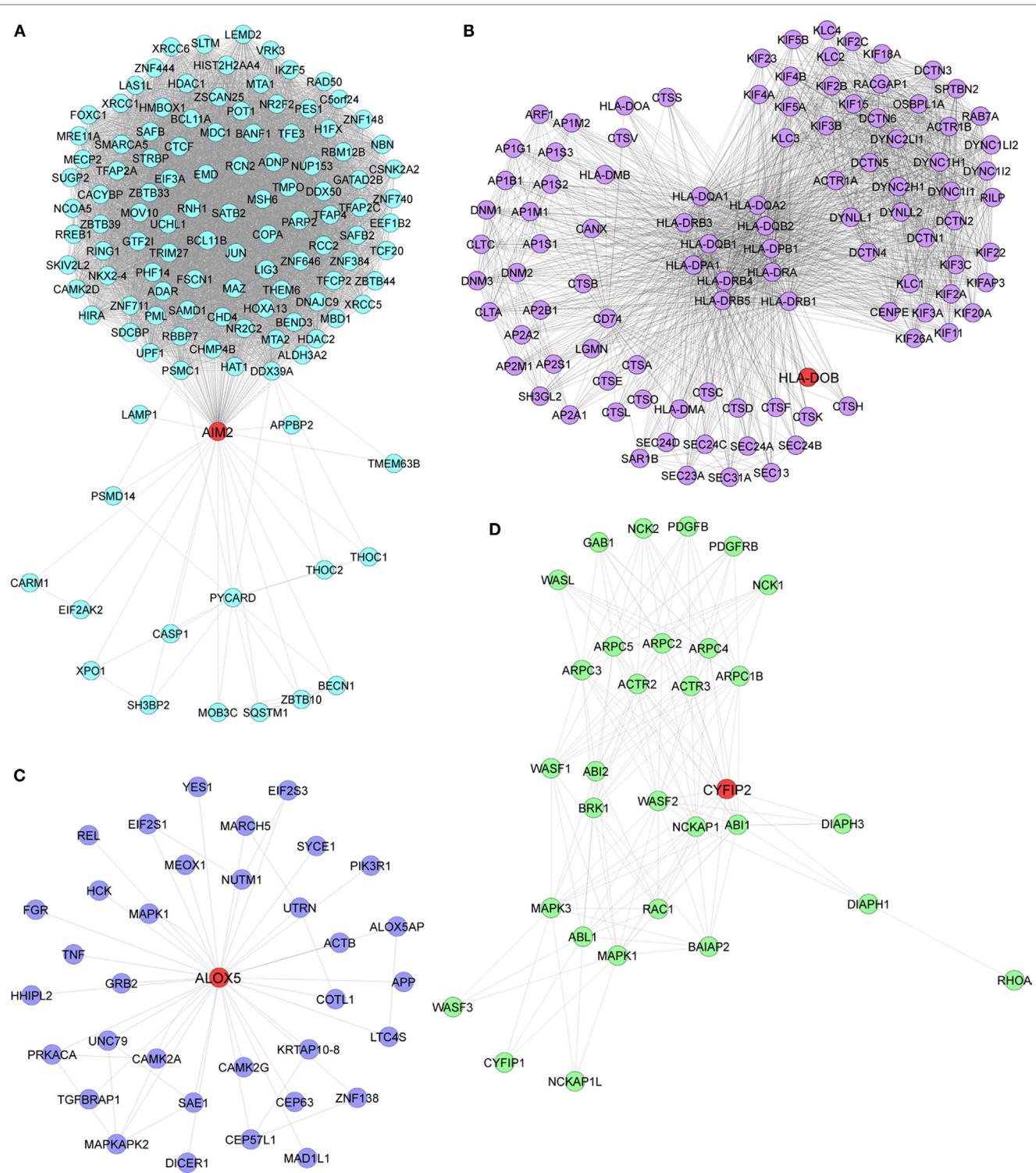


FIGURE 4 | Biological pathway networks of (A) AIM2, (B) HLA-DOB, (C) ALOX5, and (D) CYFIP2, generated using pathway commons and visualized on Cytoscape.

receptor ILT-2 and ILT-4 of the natural killer or CTLs, thereby blocking their cytotoxicity for the tumor cells (67). These findings categorically depict the immunological relevance of HLA gene

signatures in multiple cancers; however, their abundance in RA patients and correlation with RA onset and progression needs to be completely deciphered.

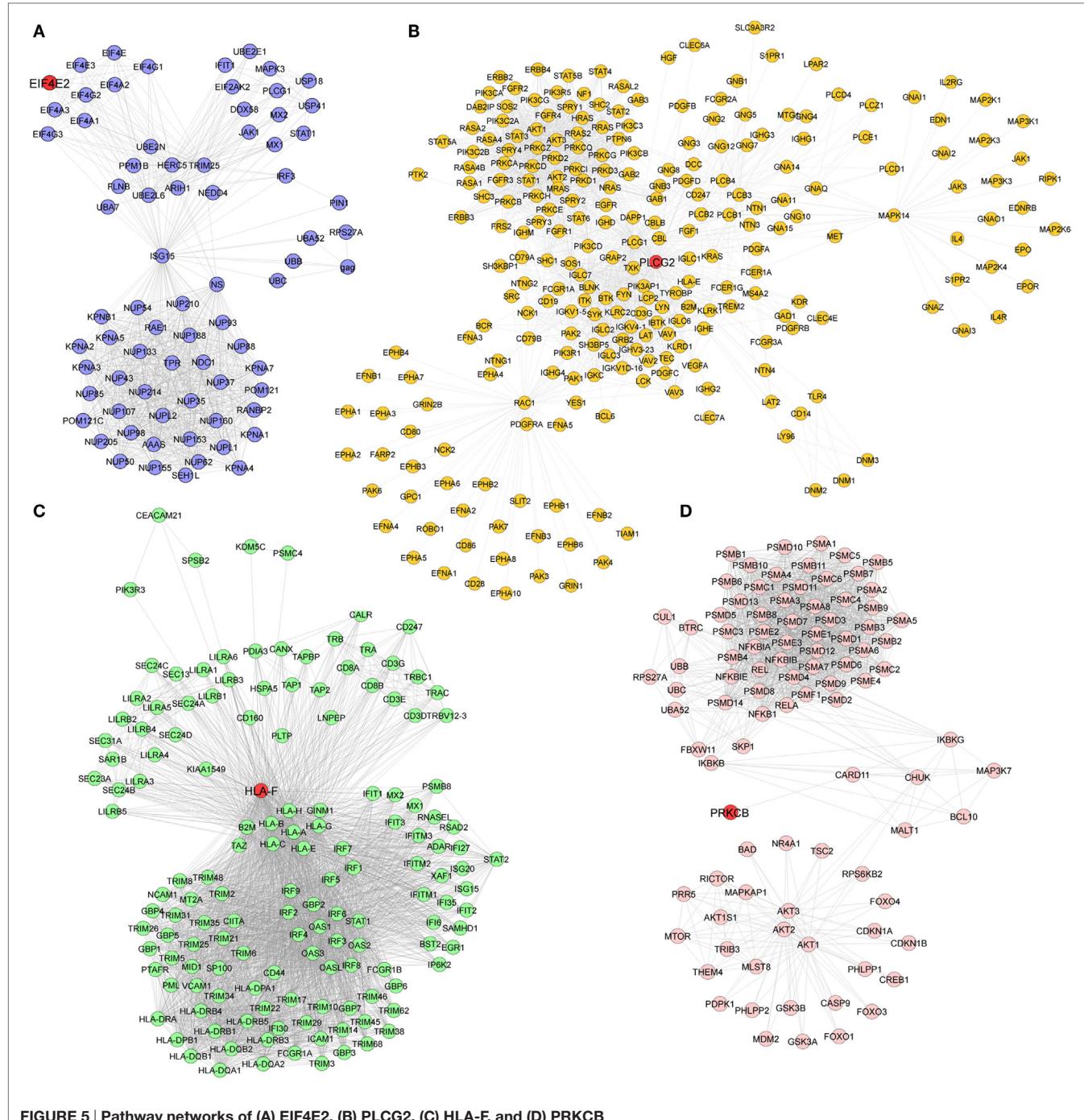


FIGURE 5 | Pathway networks of (A) EIF4E2, (B) PLCG2, (C) HLA-F, and (D) PRKCB.

Angiogenesis is one of the critical events in the perpetuation of RA (68). One of the important signals required for triggering angiogenesis during RA is hypoxia, which is created majorly as a result of synovial hyperplasia (10). Our meta-analysis study for the first time has identified EIF4E2 gene (eukaryotic initiation factor 4E2) encoding eIF4E2 protein, an mRNA cap binding homolog of eIF4E, which was found to be upregulated across all RA microarray data sets considered in our meta-analysis study as well as through RT-qPCR in RA patients (**Figure 3F**).

eIF4E2 was found to be an inhibitor of translation in normal cells (69), while it binds with HIF-2 α -RBM4 complex and drives the translation of proteins under hypoxic (low oxygen tension) conditions. Cancer cells exploit the eIF4E2-mediated protein synthesis to sustain hypoxic conditions and consequently grow to significant sizes (70). Perhaps, eIF4E2 might play a critical role in driving hypoxia-induced angiogenesis during RA, hence can act as a valuable pharmacological candidate for controlling RA. Further, our study also depicts the involvement of previously

unnoticed genes such as protein kinase C- β , PRKCB (a negative regulator of BCR signaling) and cytoplasmic FMR1-interacting protein 2, CYFIP2 (inducer of p53 mediated apoptosis), which were also upregulated in all data sets and clinical RA patients (**Figures 3G,H**). CYFIP2 gene has a p53 responsive element that confers binding of p53 as well as the activation of some heterologous reporter. Inducible expression of CYFIP2 is responsible for caspase activation and cellular apoptosis (71). Although these novel immune signatures are abundantly expressed in RA patients, which categorically suggests the involvement of these gene/pathways in RA pathogenesis, still their functional relevance in terms of gene expression and disease outcome need to be deciphered. Nevertheless, identification of these novel gene signatures might provide a unique prospective to define novel molecular diagnostic candidates and ascertain potential pharmacological targets for the development of therapeutic interventions against RA.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content. All the authors read and approved the final version of the manuscript to be published.

REFERENCES

- Rudan I, Sidhu S, Papana A, Meng SJ, Xin-Wei Y, Wang W, et al. Prevalence of rheumatoid arthritis in low- and middle-income countries: a systematic review and analysis. *J Glob Health* (2015) 5(1). doi:10.1159/jogh.05.010409
- Pollard L, Choy EH, Scott DL. The consequences of rheumatoid arthritis: quality of life measures in the individual patient. *Clin Exp Rheumatol* (2005) 23(5 Suppl 39):S43–52.
- McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med* (2011) 365(23):2205–19. doi:10.1056/NEJMra1004965
- Abdel-Nasser AM, Rasker JJ, Vaikenburg HA. Epidemiological and clinical aspects relating to the variability of rheumatoid arthritis. *Semin Arthritis Rheum* (1997) 27(2):123–40. doi:10.1016/S0049-0172(97)80012-1
- Choy E. Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. *Rheumatology* (2012) 51(Suppl 5):v3–11. doi:10.1093/rheumatology/kes113
- Firestein GS. Evolving concepts of rheumatoid arthritis. *Nature* (2003) 423(6937):356–61. doi:10.1038/nature01661
- MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* (2000) 43(1):30–7. doi:10.1002/1529-0131(200001)43:1<30::AID-ANR5>3.0.CO;2-B
- McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. *Nat Rev Immunol* (2007) 7(6):429–42. doi:10.1038/nri2094
- Smolen JS, Steiner G. Therapeutic strategies for rheumatoid arthritis. *Nat Rev Drug Discov* (2003) 2(6):473–88. doi:10.1038/nrd1109
- Paleolog EM. Angiogenesis in rheumatoid arthritis. *Arthritis Res* (2002) 4(Suppl 3):S81–90. doi:10.1186/ar527
- Nakahara H, Song J, Sugimoto M, Hagihara K, Kishimoto T, Yoshizaki K, et al. Anti-interleukin-6 receptor antibody therapy reduces vascular endothelial growth factor production in rheumatoid arthritis. *Arthritis Rheum* (2003) 48(6):1521–9. doi:10.1002/art.11143
- Teitelbaum SL. Bone resorption by osteoclasts. *Science* (2000) 289(5484):1504–8. doi:10.1126/science.289.5484.1504
- Smolen JS, Aletaha D, Koeller M, Weisman MH, Emery P. New therapies for treatment of rheumatoid arthritis. *Lancet* (2007) 370(9602):1861–74. doi:10.1016/S0140-6736(07)60784-3
- Panichi V, Migliori M, De Pietro S, Taccolla D, Andreini B, Metelli MR, et al. The link of biocompatibility to cytokine production. *Kidney Int Suppl* (2000) 76:S96–103. doi:10.1046/j.1523-1755.2000.07612.x
- Weyand CM, Hicok KC, Conn DL, Goronzy JJ. The influence of HLA-DRB1 genes on disease severity in rheumatoid arthritis. *Ann Intern Med* (1992) 117(10):801–6. doi:10.7326/0003-4819-117-10-801
- Gaffen SL. Role of IL-17 in the pathogenesis of rheumatoid arthritis. *Curr Rheumatol Rep* (2009) 11(5):365–70. doi:10.1007/s11926-009-0052-y
- Nadkarni S, Mauri C, Ehrenstein MR. Anti-TNF-alpha therapy induces a distinct regulatory T cell population in patients with rheumatoid arthritis via TGF-beta. *J Exp Med* (2007) 204(1):33–9. doi:10.1084/jem.20061531
- Silverman GJ, Carson DA. Roles of B cells in rheumatoid arthritis. *Arthritis Res Ther* (2003) 5(4):S1. doi:10.1186/ar1010
- Mesko B, Poliska S, Szamosi S, Szekanecz Z, Podani J, Varadi C, et al. Peripheral blood gene expression and IgG glycosylation profiles as markers of tocilizumab treatment in rheumatoid arthritis. *J Rheumatol* (2012) 39(5):916–28. doi:10.3899/jrheum.110961
- Mesko B, Poliska S, Váncsa A, Szekanecz Z, Palatka K, Hollo Z, et al. Peripheral blood derived gene panels predict response to infliximab in rheumatoid arthritis and Crohn's disease. *Genome Med* (2013) 5(6):1. doi:10.1186/gm463
- Ungephuam U, Haeupl T, Witt H, Koczan D, Krenn V, Huber H, et al. Molecular signatures and new candidates to target the pathogenesis of rheumatoid arthritis. *Physiol Genomics* (2010) 42(4):267–82. doi:10.1152/physiogenomics.00004.2010
- Huber R, Hummert C, Gausmann U, Pohlers D, Koczan D, Guthke R, et al. Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. *Arthritis Res Ther* (2008) 10(4):1. doi:10.1186/ar2485
- Sun Y, Caplazi P, Zhang J, Mazloom A, Kummerfeld S, Quinones G, et al. PILR α negatively regulates mouse inflammatory arthritis. *J Immunol* (2014) 193(2):860–70. doi:10.4049/jimmunol.1400045
- Woetzel D, Huber R, Kupfer P, Pohlers D, Pfaff M, Driesch D, et al. Identification of rheumatoid arthritis and osteoarthritis patients by transcriptome-based rule set generation. *Arthritis Res Ther* (2014) 16(2):1. doi:10.1186/ar4526
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* (1995) 57(1):289–300.

NK had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. NK, SA, and JG participated in study conception and design. SA and JG were involved in the acquisition of data. SV and SN were responsible for recruitment and clinical evaluation of RA patients. NK, SA, and JG participated in analysis and interpretation of data and manuscript writing.

FUNDING

This study was financially supported by Department of Biotechnology, Government of India [BT/PR8624/MED/29/798/2013], Department of Science and Technology, Government of India [DST No: SB/YS/LS-163/2013], Department of Science and Technology, Nano Mission, Government of India [DST No. SR/NM/NS-1040/2013(G)], and University Grants Commission, Government of India [UGC No. MRP-MAJOR-BIOT-2013-40689].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fimmu.2017.00074/full#supplementary-material>.

26. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* (2004) 20(3):307–15. doi:10.1093/bioinformatics/btg405
27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi:10.1093/nar/gkv007
28. Whitlock M. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J Evol Biol* (2005) 18(5):1368–73. doi:10.1111/j.1420-9101.2005.00917.x
29. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* (2013) 14(1):128. doi:10.1186/1471-2105-14-128
30. Wu G, Feng X, Stein L. Research a human functional protein interaction network and its application to cancer data analysis. *Genome Biol* (2010) 11:R53. doi:10.1186/gb-2010-11-5-r53
31. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. *Nucleic Acids Res* (2014) 42(D1):D472–7. doi:10.1093/nar/gkt1102
32. Bausinger J, Speit G. The impact of lymphocyte isolation on induced DNA damage in human blood samples measured by the comet assay. *Mutagenesis* (2016) 31(5):567–72. doi:10.1093/mutage/gew021
33. Ravindran R, Khan N, Nakaya HI, Li S, Loebbermann J, Maddur MS, et al. Vaccine activation of the nutrient sensor GCN2 in dendritic cells enhances antigen presentation. *Science* (2014) 343(6168):313–7. doi:10.1126/science.1246829
34. Hong F, Breitling R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics* (2008) 24(3):374–82. doi:10.1093/bioinformatics/btm620
35. Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, et al. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res* (2011) 39(Database issue):D685–90. doi:10.1093/nar/gkq1039
36. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the pathway interaction database. *Nucleic Acids Res* (2009) 37(Suppl 1):D674–9. doi:10.1093/nar/gkn653
37. Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrypek E, Murray B, et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* (2011) 40(D1):D261–70. doi:10.1093/nar/gkr1122
38. Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol* (2004) 6(1):1. doi:10.1186/gb-2004-6-1-p1
39. Prasad TK, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, et al. Human protein reference database – 2009 update. *Nucleic Acids Res* (2009) 37(Suppl 1):D767–72. doi:10.1093/nar/gkn892
40. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* (2013) 41(D1):D377–86. doi:10.1093/nar/gks1118
41. Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D. The database of interacting proteins: 2004 update. *Nucleic Acids Res* (2004) 32(Suppl 1):D449–51. doi:10.1093/nar/gkh086
42. Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* (2006) 34(Suppl 1):D535–9. doi:10.1093/nar/gkj109
43. Orchard S, Ammari M, Aranda B, Breuza L, Brigandt I, Broackes-Carter F, et al. The MINTAct project – IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res* (2013) 42(D1):D358–63. doi:10.1093/nar/gkt1115
44. Bader GD, Betel D, Hogue CW. BIND: the biomolecular interaction network database. *Nucleic Acids Res* (2003) 31(1):248–50. doi:10.1093/nar/gkg056
45. Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, et al. CORUM: the comprehensive resource of mammalian protein complexes – 2009. *Nucleic Acids Res* (2010) 38(Suppl 1):D497–501. doi:10.1093/nar/gkp914
46. Wingender E. The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform* (2008) 9(4):326–32. doi:10.1093/bib/bbn016
47. Hsu S-D, Tseng Y-T, Shrestha S, Lin Y-L, Khaleel A, Chou C-H, et al. miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* (2014) 42(D1):D78–85. doi:10.1093/nar/gkt1266
48. Law V, Knox C, Djoumou Y, Jewison T, Guo AC, Liu Y, et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* (2014) 42(D1):D1091–7. doi:10.1093/nar/gkt1068
49. Thiele I, Swainston N, Fleming RM, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. *Nat Biotechnol* (2013) 31(5):419–25. doi:10.1038/nbt.2488
50. Davis AP, Grondin CJ, Lennon-Hopkins K, Saraceni-Richards C, Sciaky D, King BL, et al. The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. *Nucleic Acids Res* (2015) 43(D1):D914–20. doi:10.1093/nar/gku935
51. Wrzodek C, Büchel F, Ruff M, Dräger A, Zell A. Precise generation of systems biology models from KEGG pathways. *BMC Syst Biol* (2013) 7(1):1. doi:10.1186/1752-0509-7-15
52. Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, et al. SMPDB 2.0: big improvements to the small molecule pathway database. *Nucleic Acids Res* (2013) 42(D1):D478–84. doi:10.1093/nar/gkt1067
53. Yamamoto S, Sakai N, Nakamura H, Fukagawa H, Fukuda K, Takagi T. INOH: ontology-based highly structured database of signal transduction pathways. *Database* (2011) 2011:bar052. doi:10.1093/database/bar052
54. Kandasamy K, Mohan SS, Raju R, Keerthikumar S, Kumar GSS, Venugopal AK, et al. NetPath: a public resource of curated signal transduction pathways. *Genome Biol* (2010) 11(1):1. doi:10.1186/gb-2010-11-1-r3
55. Pico AR, Kelder T, Van Iersel MP, Hanspers K, Conklin BR, Evelo C. WikiPathways: pathway editing for the people. *PLoS Biol* (2008) 6(7):e184. doi:10.1371/journal.pbio.0060184
56. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al. The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Res* (2013) 41(D1):D456–63. doi:10.1093/nar/gks1146
57. Consortium U. Activities at the universal protein resource (UniProt). *Nucleic Acids Res* (2014) 42(D1):D191–8. doi:10.1093/nar/gkt1140
58. Chambers J, Davies M, Gaulton A, Hersey A, Velankar S, Petryszak R, et al. UniChem: a unified chemical structure cross-referencing and identifier tracking system. *J Cheminform* (2013) 5(1):1. doi:10.1186/1758-2946-5-3
59. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* (2003) 13(11):2498–504. doi:10.1101/gr.1239303
60. Devauchelle V, Marion S, Cagnard N, Mistou S, Falgarone G, Breban M, et al. DNA microarray allows molecular profiling of rheumatoid arthritis and identification of pathophysiological targets. *Genes Immun* (2004) 5(8):597–608. doi:10.1038/sj.gene.6364132
61. Chae JJ, Park YH, Park C, Hwang IY, Hoffmann P, Kehrl JH, et al. Connecting two pathways through Ca²⁺ signaling: NLRP3 inflammasome activation induced by a hypermorphic PLCG2 mutation. *Arthritis Rheumatol* (2015) 67(2):563–7. doi:10.1002/art.38961
62. Baum R, Sharma S, Carpenter S, Li QZ, Bustos P, Fitzgerald KA, et al. Cutting edge: AIM2 and endosomal TLRs differentially regulate arthritis and autoantibody production in DNase II-deficient mice. *J Immunol* (2015) 194(3):873–7. doi:10.4049/jimmunol.1402573
63. Lin HC, Lin TH, Wu MY, Chiu YC, Tang CH, Hour MJ, et al. 5-Lipoxygenase inhibitors attenuate TNF-alpha-induced inflammation in human synovial fibroblasts. *PLoS One* (2014) 9(9):e107890. doi:10.1371/journal.pone.0107890
64. Mahshid Y, Markoutsas S, Dincbas-Renqvist V, Surun D, Christensson B, Sander B, et al. Phosphorylation of serine 523 on 5-lipoxygenase in human B lymphocytes. *Prostaglandins Leukot Essent Fatty Acids* (2015) 100:33–40. doi:10.1016/j.plefa.2015.06.003
65. Kang YJ, Zeng W, Song W, Reinhold B, Choi J, Brusic V, et al. Identification of human leucocyte antigen (HLA)-A*0201-restricted cytotoxic T lymphocyte epitopes derived from HLA-DObeta as a novel target for multiple myeloma. *Br J Haematol* (2013) 163(3):343–51. doi:10.1111/bjh.12544
66. Harada A, Ishigami S, Kijima Y, Nakajo A, Arigami T, Kurahara H, et al. Clinical implication of human leukocyte antigen (HLA)-F expression in breast cancer. *Pathol Int* (2015) 65(11):569–74. doi:10.1111/pin.12343

67. Lepin EJM, Bastin JM, Allan DSJ, Roncador G, Braud VM, Mason DY, et al. Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. *Eur J Immunol* (2000) 30(12):3552–61. doi:10.1002/1521-4141(200012)30:12<3552::AID-IMMU3552>3.0.CO;2-L
68. Koch AE. Angiogenesis as a target in rheumatoid arthritis. *Ann Rheum Dis* (2003) 62(Suppl 2):ii60–7. doi:10.1136/ard.62.suppl_2.ii60
69. Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, et al. A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol* (2012) 32(17):3585–93. doi:10.1128/Mcb.00455-12
70. Uniacke J, Perera JK, Lachance G, Francisco CB, Lee S. Cancer cells exploit eIF4E2-directed synthesis of hypoxia response proteins to drive tumor progression. *Cancer Res* (2014) 74(5):1379–89. doi:10.1158/0008-5472.CAN-13-2278
71. Jackson RS, Cho YJ, Stein S, Liang P. CYFIP2, a direct p53 target, is leptomyctin-B sensitive. *Cell Cycle* (2007) 6(1):95–103. doi:10.4161/Cc.6.1.3665

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Afroz, Giddaluru, Vishwakarma, Naz, Khan and Khan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.