UNIVERSITY^{OF} BIRMINGHAM

University of Birmingham Research at Birmingham

Stromal transcriptional profiles reveal hierarchies of anatomical site, serum response and disease and identify disease specific pathways

Filer, Andrew; Antczak, Philipp; Parsonage, Gregory; Legault, Holly M; O'Toole, Margot; Pearson, Mark; Thomas, Andrew M; Scheel-Toellner, Dagmar; Raza, Karim; Buckley, Christopher D; Falciani, Francesco

DOI:

10.1371/journal.pone.0120917

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Filer, A, Antczak, P, Parsonage, G, Legault, HM, O'Toole, M, Pearson, M, Thomas, AM, Scheel-Toellner, D, Raza, K, Buckley, CD & Falciani, F 2015, 'Stromal transcriptional profiles reveal hierarchies of anatomical site, serum response and disease and identify disease specific pathways', *PLoS ONE*, vol. 10, no. 3, e0120917. https://doi.org/10.1371/journal.pone.0120917

Link to publication on Research at Birmingham portal

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- $\bullet \textsc{Users}$ may freely distribute the URL that is used to identify this publication.
- •Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- •User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
 •Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 06. May. 2024





GOPEN ACCESS

Citation: Filer A, Antczak P, Parsonage GN, Legault HM, O'Toole M, Pearson MJ, et al. (2015) Stromal Transcriptional Profiles Reveal Hierarchies of Anatomical Site, Serum Response and Disease and Identify Disease Specific Pathways. PLoS ONE 10 (3): e0120917. doi:10.1371/journal.pone.0120917

Academic Editor: Lars Kaderali, Technische Universität Dresden, Medical Faculty, GERMANY

Received: May 13, 2014

Accepted: February 9, 2015

Published: March 25, 2015

Copyright: © 2015 Filer et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are

Data Availability Statement: Microarray data is available from Gene Expression Omnibus (Accession number GSE56409).

Funding: Supported by research grants from the Medical Research Council and Arthritis Research UK (Grant 18547). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 305549. AF, CB and KR have received research grant funding from Wyeth. Partially funded by the BBSRC/EraSysBio Linknet BB/I004556/1 grant to FF.

RESEARCH ARTICLE

Stromal Transcriptional Profiles Reveal Hierarchies of Anatomical Site, Serum Response and Disease and Identify Disease Specific Pathways

Andrew Filer^{1,2©‡}*, Philipp Antczak^{3©}, Greg N. Parsonage⁴, Holly M. Legault⁵, Margot O'Toole⁵, Mark J. Pearson⁶, Andrew M. Thomas⁷, Dagmar Scheel-Toellner¹, Karim Raza^{1,8}, Christopher D. Buckley^{1,8}, Francesco Falciani^{3‡}*

- 1 Rheumatology Research Group, Centre for Muscoloskeletal Ageing Research, School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Queen Elizabeth Hospital, Birmingham, B15 2WD, UK, 2 University Hospitals Birmingham NHS Foundation Trust, Mindelsohn Way, Birmingham, B15 2WB, UK, 3 Centre of Computational Biology and Modelling (CCBM), Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool, UK, 4 School of Cancer Sciences, College of Medical and Dental Sciences, The University of Birmingham, B15 2TT, UK, 5 Biological Technologies, Wyeth Research, Cambridge, Massachusetts 02140, USA, 6 MRC-ARUK Centre for Musculoskeletal Ageing Research, School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Queen Elizabeth Hospital, Birmingham, B15 2WD, UK, 7 The Royal Orthopaedic Hospital NHS Foundation Trust, Birmingham, UK, 8 Sandwell and West Birmingham Hospitals NHS Trust, Dudley Road, Birmingham, B18 7QH, UK
- These authors contributed equally to this work.
- ‡ These authors would like to be considered Joint Corresponding Authors.
- * f.falciani@liv.ac.uk (FF), a.filer@bham.ac.uk (AF)

Abstract

Synovial fibroblasts in persistent inflammatory arthritis have been suggested to have parallels with cancer growth and wound healing, both of which involve a stereotypical serum response programme. We tested the hypothesis that a serum response programme can be used to classify diseased tissues, and investigated the serum response programme in fibroblasts from multiple anatomical sites and two diseases. To test our hypothesis we utilized a bioinformatics approach to explore a publicly available microarray dataset including rheumatoid arthritis (RA), osteoarthritis (OA) and normal synovial tissue, then extended those findings in a new microarray dataset representing matched synovial, bone marrow and skin fibroblasts cultured from RA and OA patients undergoing arthroplasty. The classical fibroblast serum response programme discretely classified RA, OA and normal synovial tissues. Analysis of low and high serum treated fibroblast microarray data revealed a hierarchy of control, with anatomical site the most powerful classifier followed by response to serum and then disease. In contrast to skin and bone marrow fibroblasts, exposure of synovial fibroblasts to serum led to convergence of RA and OA expression profiles. Pathway analysis revealed three inter-linked gene networks characterising OA synovial fibroblasts: Cell remodelling through insulin-like growth factors, differentiation and angiogenesis through _3 integrin, and regulation of apoptosis through CD44. We have demonstrated that Fibroblast serum response signatures define disease at the tissue level, and that an OA specific,



Competing Interests: The authors would also like to confirm that at the time of the original submission Francesco Falciani was a PLOS ONE Editorial Board member. However, Dr. Falciani is no longer serving in that position. Moreover, the authors confirm that the fact that some of the co-authors are employed in a company does not alter their adherence to all PLOS ONE policies on sharing data and materials, as detailed online in the journal guide for authors.

serum dependent repression of genes involved in cell adhesion, extracellular matrix remodelling and apoptosis is a critical discriminator between cultured OA and RA synovial fibroblasts.

Introduction

Determining the mechanisms underlying persistence of disease in arthritis remains a significant challenge. The fibroblast-like synoviocyte (FLS) has become prominent as a mediator of disease persistence, supported by the persistent in vitro phenotype of these cells exemplified by the severe combined immunodeficiency (SCID) mouse model of cartilage invasion [1]. Osteoarthritis (OA) synovial fibroblasts have been investigated less extensively. Although OA FLS show some degree of hyperplasia [2, 3], they do not possess significant invasive properties. Transcriptomic analyses in conditions of high serum have revealed unique functional groupings of activated genes [4]. Such apparently imprinted characteristics are not unique amongst fibroblasts. Chang et al demonstrated retained fibroblast regional memory after in vitro culture [5]. However, this group also demonstrated a canonical serum response programme (CSR) in fibroblasts exposed to high (10% fetal calf serum (FCS)) compared to very low (0.1% FCS) serum [6, 7]. The presence of the CSR in publicly available gene expression profiles of cancer tissues robustly predicted metastasis and poor survival [7]. Further parallels exist between invasive tumours, wound healing and FLS. These include the expression of tumour associated markers such as Fibroblast activation protein (FAP) [8], Galectin-3 [9] and S100A4 (FSP-1) [10], evidence of distant spread of pathologically involved cells [11], and activation of pathways implicated in wound healing such as TGF β mediated myofibroblast differentiation as shown by Kasperkovitz et al [12]. The findings of Chang and Kasperkovitz [12] illustrate the potential of microarray analysis to generate hypotheses about the mechanisms underlying tissue microenvironments, but also highlight the critical effects of both anatomical site and serum exposure upon transcriptomic profiles. No study has attempted to examine diseased synovial fibroblasts in the context of serum stimulation. In order to fully understand the impact of serum in disease, we analysed transcriptional responses to serum of fibroblasts derived from three matched anatomical locations in populations of RA and OA patients. We identified a strict hierarchy of transcriptional regulation, and an OA synovium-specific signature, which is repressed in response to serum stimulation and represents important regulatory pathways involving control of extracellular matrix remodelling and apoptosis.

Results

A canonical fibroblast serum response signature discriminates between normal, RA and OA synovial tissues

RA FALS display an agressive phenotype, which involves increased invasiveness and resistance to apoptosis [13]. Because of this, the expansive synovial tissue called 'pannus' at the cartilage-bone interface has been suggested to share many hallmarks of cancer. However, there is no evidence of increased cell proliferation in RA compared to normal and OA FLS. Invasiveness, apoptosis and proliferation in RA FLS derived from patients share the same properties than their *in vivo* counterparts [1]. We therefore formulated the hypothesis that a fibroblast serum-response signature, linked to cancer clinical outcome, may be able to discriminate RA and OA derived synovial fibroblasts. Consistent with the current understanding of the biology of



synovial FLS, we also hypothesize that migration and cell death but not proliferation may be the key difference between RA and OA derived FLS. We first tested this hypothesis by developing a statistical model predictive of disease type based on the original transcriptional signature developed by Chang et al [6]. We were able to map 95% of the genes originally described on a public domain dataset representing whole synovial tissue from normal, RA and OA individuals [14]. The hypothesis was indeed consistent with the prediction results as a K-nearest neighbour (KNN) based model was able to predict disease state with good accuracy. Normal samples were predicted with a sensitivity of 61% and a specificity of 95%. RA samples were predicted with a sensitivity of 88% and a specificity of 89% and OA samples were predicted with a sensitivity of 80% and a specificity of 81% (S1 Fig.). This observation therefore supported our hypothesis that fibroblast pathways involved in wound healing and neoplasia are relevant to the persistence of inflammation in RA and OA.

A distinct hierarchy of anatomical site and serum response distinguishes fibroblasts derived from arthritis patients

Since the 'core serum response' signature (CSR) developed by Chang et al was based on fibroblasts derived from normal skin, lung and gum we reasoned that it may not fully represent the responses of synovial fibroblast populations in diseased joints. We therefore set out to experimentally define a serum response signature based on fibroblasts derived from both RA and OA patients. In order to take account of the documented effect of site on fibroblast transcriptional profiles and to address the possibility of multiple site involvement in systemic disease, we compared the serum response of fibroblasts of synovial origin to those of bone marrow and skin from the same individual. We set out to identify differentially expressed genes in at least one of the fibroblast groups. By using a SAM multiclass approach we could identify a signature of 3633 genes (FDR < 3% followed by a two-fold cut-off). Microarray gene expression was correlated with qPCR data for validation purposes (S2 Table). Principal component analysis (PCA) was used to visualise the relationship between the genes (Fig. 1). This unsupervised analysis showed that fibroblasts cluster primarily on the basis of anatomical location. Furthermore all matched low and high serum samples were separated on the first principal component (PC1, x-axis). By contrast, minimal clustering on the basis of disease was seen. These results provide the first evidence of a molecular hierarchy of transcriptional control, where site of origin is the dominant factor, followed by response to serum stimulation and disease status.

Within anatomical locations, RA and OA fibroblasts are molecularly distinct and differ in their response to serum stimulation

While the PCA described above provides initial evidence for the existence of a molecular hierarchy of transcriptional control, it does not quantify the effect and it does not identify the specific genes linked to anatomical location, serum response and disease type. We therefore set out to directly quantify the differences in gene expression across all three levels of organisation and performed the following direct comparisons:

- 1. Anatomical location (Bone Marrow vs. Skin vs. Synovial tissue)
- 2. Serum Response (High vs. Low Serum)
- 3. Disease type (RA vs. OA)

Fig. 2A summarized the results of this analysis showing that as expected anatomical location has the highest number of differentially expressed genes followed by serum response and finally disease (3763, 435 and 180 differentially expressed genes, respectively at < 5% and a 2-fold cut-



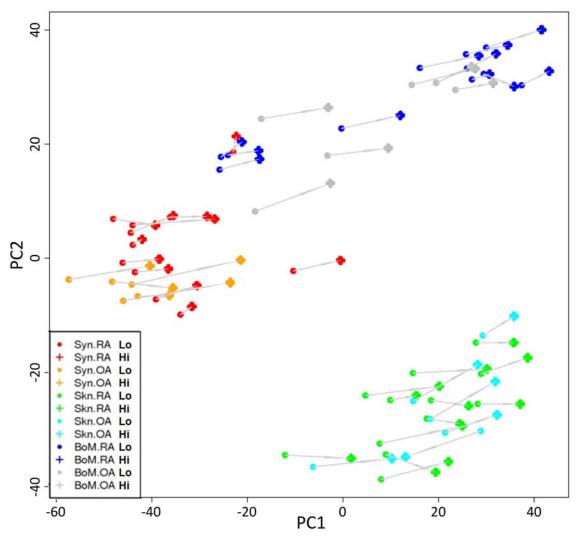


Fig 1. Identification of a hierarchy of molecular signatures in arthritis derived fibroblasts. The relative similarity of the different groups of fibroblast in a principal component (PC) plot is shown. Groups defined by disease state and anatomical location are indicated by coloured symbols whereas individual samples response to serum (Hi, High serum; Lo, low serum) is shown by a solid grey line. The PCA clearly separates synovium (Syn, red and orange) from bone marrow (BoM, blue and grey) and skin (Skn, green and light blue) on the first PC. Skin and bone marrow samples are only separated on the second PC. Additionally it can be observed that all samples are separated by their serum response mainly on the first PC.

off). We then focused within each anatomical location and performed a more detailed analysis. More precisely, within each tissue and within each disease group we identified genes differentially expressed between low and high serum (Fig. 2B). We found that in both RA and OA the synovial tissue has the highest number of differentially expressed genes (738 and 1419 genes respectively). Bone marrow tissues both expressed similar number of genes (RA: 469, OA: 558) while skin showed a much higher number of genes in RA (633) than OA (276) (Fig. 2B). We then identified within each anatomical location and within each serum level the number of differentially expressed genes between RA and OA samples (Fig. 2C). Here interestingly a large difference in the low serum but not the high serum condition could be observed in the synovial tissue (302 vs. 0 differentially expressed genes respectively). Skin as well as bone marrow both expressed similar numbers of differentially expressed genes (96±2) (Fig. 2C). The fold changes of differentially expressed genes were broadly similar between groups indicating that numbers

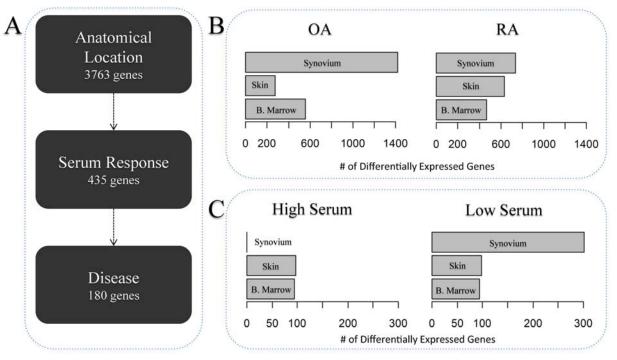


Fig 2. Differential gene expression analyses at tissue, serum and disease level. The figure shows the differential gene expression analysis performed within each level of the study. For each level of organisation we quantified the number of differentially expressed genes at an FDR < 5% followed by a 2-fold filter. This allowed us to devise a hierarchy of organisation that follows Anatomical Location then serum response and finally disease status (A). A more detailed tissue-by-tissue analysis to identify serum response genes in RA and OA fibroblasts (B) and to identify disease genes within high and low serum (C) was then performed. The arrows in panel A represent the direction of the hierarchy defined by the number of genes differentially expressed. BoM, Bone marrow.

of genes are a useful approximation of the degree of difference observed between transcriptional profiles. For each anatomical location we also performed a multi-class SAM analysis to identify genes that change in at least one of the different serum and disease types. We then visualised the resulting gene-lists using a PCA (Fig. 3). This clearly demonstrated that bone marrow and skin RA and OA fibroblasts responded to serum following parallel trajectories along the first principal component (PC) and were separated by disease predominantly according to the second PC (Figs. 3A and 3B), while RA and OA synovium fibroblasts separated along the second PC but converged in response to serum (Fig. 3C). In order to investigate the serum response overlap between bone marrow and skin fibroblasts, we extracted the most contributing genes within the first principle component, and found that 120 and 285 genes contributed to the first principle component of bone marrow and skin PCA analyses respectively, sharing 36 genes. This exclusive skin/bone marrow overlap suggested a specific synovial serum response.

Serum response signatures defined by anatomical location and disease type define distinct functional profiles

Having defined the serum response signatures in each anatomical location and within each disease group we asked whether the serum response programs overlap significantly Samples derived from all three different anatomical locations of RA patients exhibited a greater overlap in serum response profile than OA samples, which exhibited more divergent, site-specific serum responses (11% vs 3% in common, supplementary S2 Fig.). Next, in order to address the biological significance of our findings we used the downstream effect analysis pipeline within the



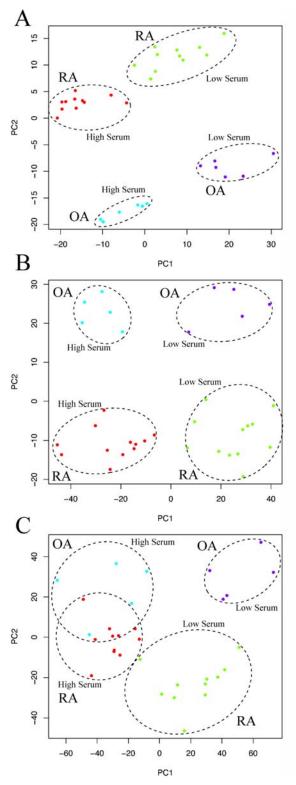


Fig 3. Serum response in arthritis derived fibroblasts within anatomical sites. The figure describes the relationships between RA and OA derived fibroblasts in low and high serum conditions using PCA. Bone marrow (A), skin (B) and synovium (C) derived fibroblasts are represented separately.



Ingenuity Pathway Analysis (IPA) knowledge management system. This algorithm assesses the probability that the observed transcriptional response is consistent with a phenotypic change. The rationale behind this approach is as follows: The IPA database contains experimentally observed causal relationships that provide a direction of effect between genes and functional terms. Given that generally a number of genes may be linked to any given function, a statistical approach to quantify this effect was developed by the IPA Team. Functional terms that have significantly more "increased" or "decreased" predictions are found by comparing the submitted gene lists with a random selection with equal probability. Fig. 4 represents a summary of the resulting functional terms linked to the serum response of anatomical locations and disease types. Most of the genes submitted by each comparison can be classified into functions relating to migration and invasion of tumor cell lines, proliferation of cells and cell viability (See S3 Table for more details). However there are a number functions specific to each anatomical site and disease type. Particularly bone marrow and synovial tissue show a number of overlapping functional terms such as invasion, synthesis of DNA and apoptosis. Moreover, bone marrow RA specific functions are one of the largest set of functions associated to a number of genes (423 genes associated to 30 increased and 104 genes to decreased functions) including differentiation, angiogenesis, vasculogenesis, activation of tumor cell lines and apoptosis of tumor cell lines (decreased). Within the synovial tissue, RA samples are characterized by proliferation of fibroblasts, adhesion of tumor cell lines and necrosis and cell death (decreased 145 genes in 2 functions). Similarly to RA, OA also regulate genes associated with apoptosis but with a much larger number of genes (246 genes associated to 2 functions). Interestingly, OA synovial samples are the only sample type with functions relating to sensitization of cells which may have an impact on their ability to respond to serum.

Analysis of the synovium fibroblast response to serum identifies OA-specific repression of a tissue remodelling transcriptional signature

A key question in arhritis research is whether RA and OA fibroblasts are different in their ability to respond to grow-factor signals. As shown in Fig. 3C, the transcriptional state of RA and OA derived fibroblasts converged in high serum culture conditions. We therefore set out to further characterize the difference in response between RA and OA in low serum conditions within the synovium and to assess whether these changes in expression follow serum stimulation. In the low serum condition, 302 genes were differentially expressed between synovial OA and RA fibroblasts. Genes significantly down-regulated in OA synovial fibroblasts during low to high serum transition and not included in classical GO classifiers included tumour suppressor and pro-apoptotic genes PTEN, SLIT3, TRERF1 (TReP-132), SHOX, SH3D19, ubr5 and EDD; genes with roles in cytoskeletal rearrangement supervillin, paxillin, and genes involved in ubiquitination and signalling included PIAS1, crebbp and the IGFBP family of genes. Genes regulated at higher levels in RA than OA fibroblasts included Secretogranin-II and tensin 3. Of the 302 genes, 98% were not differentially expressed in bone marrow or skin fibroblasts, confirming a synovium-specific response (Fig. 5A). Fig. 5B shows the relative expression in low and high serum of differentially expressed genes in synovial fibroblasts. The majority of genes (88%) were down-regulated in response to high serum in OA synovial fibroblasts. This relationship was also seen to a lesser extent in skin and bone marrow OA fibroblasts, while RA fibroblasts showed no convincing regulation of this gene list in response to serum. To define the biological relevance of the serum mediated synovium-specific disease signature, we performed a network analysis using Ingenuity Pathway Analysis software. This identified three interlinked gene networks representing the control of actin remodelling through insulin-like growth factor associated proteins (IGFBP5, IGFBP6, IGFBP7 and GnRH1) (Fig. 6A), control of tissue



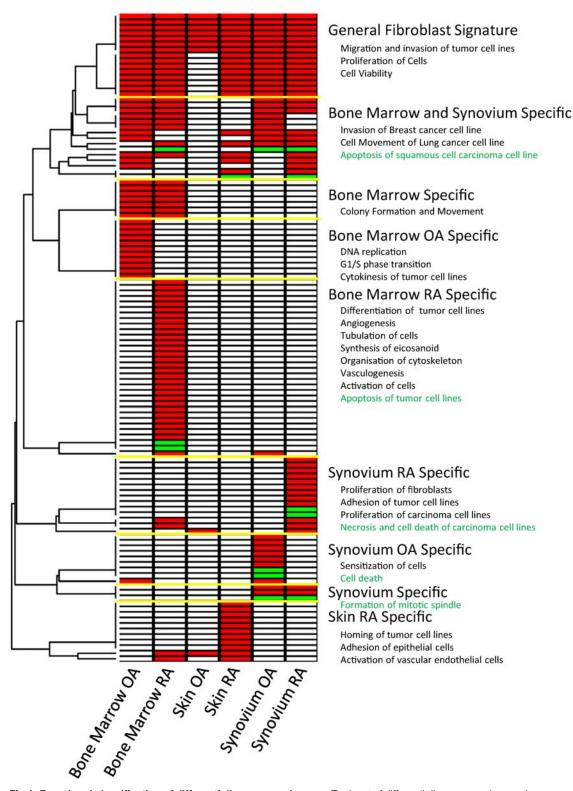


Fig 4. Functional classification of differentially expressed genes. Each set of differentially expressed genes in response to serum was submitted to Ingenuity and searched for functional terms with predicted increased (red) or decreased (green) activity.



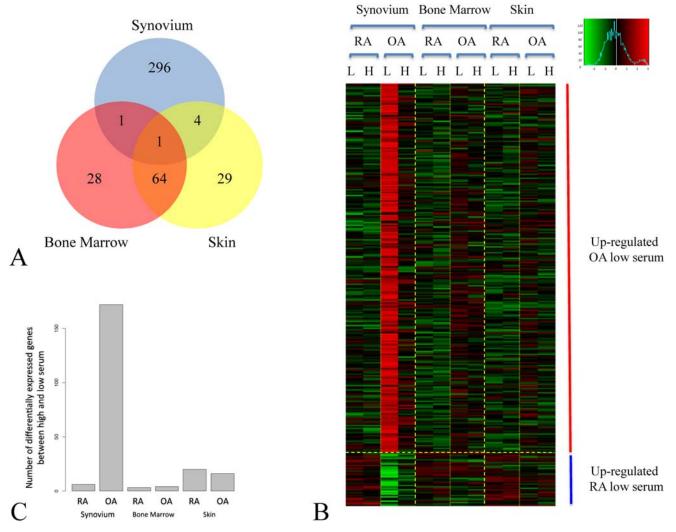


Fig 5. A synovium specific gene signature in low serum. Panel A shows a Venn diagram showing the overlap between genes differentially expressed in synovium, bone marrow and skin derived fibroblasts between RA and OA groups in the low serum state. The high degree of overlap between skin and bone marrow and the very different signature in the synovium are evident. Panel B shows the result of a two factor cluster analysis of the 296 synovium specific genes identified in Panel A. Panel C represents the number of statistically differentially expressed genes between low and high serum states within each tissue and disease. As expected OA synovium is associated with the largest number of differentially expressed genes.

remodelling by β 3 integrin (Fig. 6B), and regulation of apoptosis through CD44 (Fig. 6C), each representing facets of the fibroblast phenotype in OA. This last observation shows that molecular signatures characterizing RA fibroblast serum response not only discriminate between RA and OA FLS but that the functional profile is related primarily to adhesion and migration pathways. This is consistent with our initial hypothesis.

Discussion

We hypothesised that the CSR of Chang et al would differentiate persistent inflammatory RA from OA and normal synovium, and showed this to be the case at the tissue level. This observation places inflammatory arthritis at the centre of the long-standing debate regarding the relationship between wound healing, cancer and persistent inflammation [15]. Comparison of serum responses across site and disease led to the discovery of a hierarchy of transcriptomic



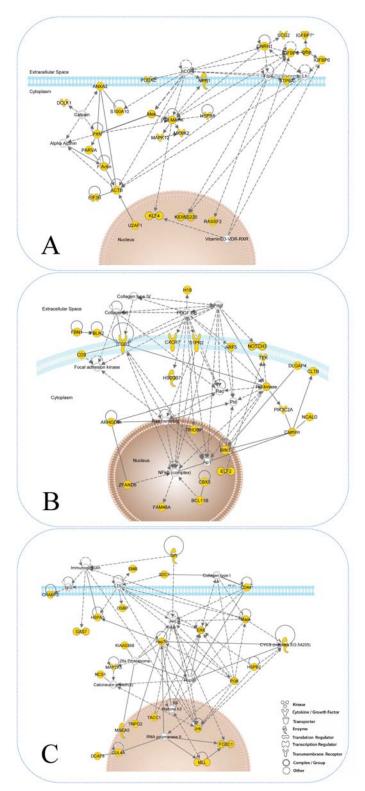


Fig 6. Ingenuity Network analyses. The three major networks identified were: (A) Actin cytoskeleton remodelling by extracellular insulin-like growth factor binding proteins through gonadotropic hormones. (B) ITGB3 signalling connected to differentiation and angiogenesis. (C) Regulation of apoptosis through CD44. Yellow nodes represent genes resulting from input whereas empty nodes represent added genes by Ingenuity. Dashed and solid lines represent indirect and direct relationships respectively.



signatures. The primacy of site as the strongest determinant of transcriptional fibroblast response fits with both Chang et al's findings and clinical observations of organ specific disease [16]. Within anatomical sites, skin and bone marrow fibroblasts exhibited parallel serum responses regardless of disease. However the serum responses of RA fibroblasts exhibited greater overlap than OA fibroblasts and were broadly functionally aligned. This alignment of serum responses in RA could result from exposure to a systemic cytokine milieu, or from a field effect in mesenchymal precursors residing in bone marrow or local niches. Intriguingly, cultured skin and bone marrow fibroblasts also retained a robust transcriptional memory of disease following in vitro culture analogous to the disease memory exhibited by RA fibroblasts in the SCID mouse invasiveness model [1]. This is highly suggestive of underlying epigenetic mechanisms, just as site specific gene expression results from noncoding RNA (HOTAIR—a 2.2kb ncRNA fragment in the HOXC locus) driven epigenetic regulation of developmentally expressed genes [17]. We have previously shown that transcriptional profiles in lymphoid and peripheral fibroblasts may be modified under the influence of cytokine treatment [16]. The phenotype of all fibroblasts in RA may therefore result from in vivo modification of the serum response programme under the influence of persistent cytokine exposure. This finding has wider implications, not least for mesenchymal stem cell therapy, where allogeneic rather than autologous therapy may be more effective. Small numbers of genes linking fibroblast behaviour in RA and cancer have been previously identified [18]; our novel approach identified links to additional genes that are relatively silenced in RA compared to OA fibroblasts. An example is the tumour suppressor PTEN, the low expression of which at sites of synovial invasion and growth has been described by Pap et al. [19] SLIT3 specifically, along with other SLIT/ROBO family members, is inactivated by promoter hypermethylation in multiple human cancers [20], and antagonizes invasiveness of RA synovial fibroblasts [21]. SH3D19 codes for the tumour suppressor EBP that inhibits ras driven transformation in acute myeloid leukaemia [22]. ubr5 (EDD), a ubiquitin ligase, similarly acts as a DNA damage checkpoint with under-expression leading to cell cycle activation [23]. Responsiveness of these pathways in OA may help to explain the different clinical and radiological phenotypes seen in OA and invasive RA. Ingenuity pathway analysis identified three inter-linked gene networks: The first network represents the insulin-like growth factor signalling network. Transcription of the IGFBP family of genes which mediate profibrotic and migratory changes [24] has been described in a "low inflammation" group of RA fibroblasts by Kasperkovitz et al [12] and by Seki et al using differential subtraction techniques [25]. The second network represents the interaction between β 3 integrin (ITGB3) and matrix components with signalling to differentiation and angiogenesis pathways. ITGB3 binds to multiple extracellular matrix components and growth factors including platelet-derived growth factor (PDGF), Collagens and Fibulin 2 [26-29]. Treatment of cells with PDGF-BB has been shown to increase the expression of ITGB3 [30] as well as increase recruitment of ITGB3 to the cell surface [28]. ITGB3:PTPN11 complexes also inhibit dephosphorylation of insulin-like growth factor-I receptor (IGF-IR), leading to inhibition of apoptosis via Akt and BclXL; hence decreased ITGB3 expression promotes apoptosis [31]. Angiogenesis related genes were consistently identified in RA fibroblast serum responses. However, although identified as a potential target in RA, an $\alpha v\beta$ 3 monoclonal antibody (Vitaxin) was withdrawn after trials failed to show therapeutic benefit [32]. Furthermore, β 3 and α v β 3 integrin knockout mice exhibit enhanced pathological angiogenesis of implanted tumours while ITGB3 over-expressing models demonstrate small new vessels deficient in pericytes [33, 34]. The role of $\alpha v \beta 3$ integrins in angiogenesis therefore remains controversial. The CD44 pathway covers a broad spectrum of genes, including the MAPKinases and PI3Kinase that are all implicated in control of apoptosis and survival [35, 36]. Furthermore FOXC1 and CYCS are regulators of apoptosis, highlighted by gene ontology analysis as the most enriched term in this pathway. Our



investigations shed light on the findings of Del Rey et al [4] who compared transcriptional profiles of synovial fibroblasts in high serum. They were surprised at the small number of differentially expressed genes seen, an explanation for which is provided by the serum conditions in which their experiments were conducted. Intriguingly, both studies identified unique aspects of the osteoarthritis synovial fibroblast phenotype, including cell adhesion and motility factors, and a conspicuous absence of proinflammatory cytokine pathways. We have demonstrated that serum response programmes segregate both tissues and fibroblasts from different sites and diseases. This may provide a molecular explanation for the ability of $TGF\beta$ linked pathways, that form a part of this response, to sub-classify disease. These data also demonstrate at a transcriptome level that low serum conditions will increase the likelihood of identifying disease-specific signatures.

Materials and Methods

Analysis of synovial tissue datasets from public repository

We retrieved a dataset from the NCBI GEO database which contained synovial tissue biopsy samples from RA, OA and normal patients (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12021) [14]. The dataset was normalized using robust multi-array average (RMA) methodology; genes with expression close to the detection level were removed. We mapped cDNA common serum response (CSR) genes to Affymetrix Ids using the ID converter facility in DAVID [37, 38]. Expression of successfully mapped genes was then determined in the GSE12021 dataset. To determine the ability of the CSR to classify synovial tissue from different diseases, tissue datasets were classified by CSR genes using a k-nearest neighbour (KNN) classification algorithm, with validation by a leave one out cross-validation (LOOCV) procedure.

Patients and fibroblast culture

Fibroblasts were isolated as previously described [39] from synovium, bone marrow and skin tissue samples taken at the time of knee or hip replacement surgery from 12 rheumatoid arthritis patients meeting the 1987 ACR criteria [40] and 6 osteoarthritis patients diagnosed on the basis of characteristic x-ray findings and the absence of features suggestive of inflammatory arthritis (S1 Table). Only one hip sample was present in either disease group. Fibroblasts were maintained in fibroblast medium (consisting of 81.3% RPMI 1640, 10% FCS, 0.81x MEM nonessential amino acids, 0.81 mM sodium orthopyruvate, 1.62 mM glutamine, 810 U/ml penicillin and 81 μ g/mL streptomycin) at 37°C in a humidified 5% CO2 atmosphere. Ethical approval for the use of this material was given by the 'NRES Committee West Midlands—The Black Country' and all patients gave written informed consent (LREC reference 5735). RA disease activity was quantified using the DAS28 disease activity score [41].

RNA extraction and microarray hybridization

Cells were seeded in duplicate at the required density using 10% FCS. Once attached, the medium was removed and the cells washed in serum free medium before replacing with fibroblast medium, which was identical but contained either 10% or 0.1% FCS [6]. mRNA was extracted from all fibroblasts at passage 5, under identical conditions. In order to remove confluence of cells as a confounding factor, proliferation rates were measured, and appropriate numbers of cells seeded to achieve 75% confluent cultures at harvest. Fibroblasts were retrieved from culture by trypsin digestion and washed in fibroblast medium then PBS before resuspending in 200 µl of RLT RNA protection buffer (Quiagen). 102 fibroblast samples in all were prepared according to Affymetrix manufacturer instructions using 50 ng total RNA, which was amplified



using Nugen™ Ovation™ Biotin RNA Amplification prior to hybridization with U133 plus 2.0 Affymetrix® Genechips. All chips passed QC criteria and visual inspection of Expressionist thumbnail images. The data is available under GEO accession number GSE56409.

Array validation using quantitative PCR

Synovial fibroblasts from 3 OA and 3 RA patients were cultured as above. Cells were trypsinised and RNA extraction performed using an RNeasy kit (Qiagen). RNA from 1×10^5 fibroblasts was reverse transcribed using Superscript Vilo (Life Technologies). 1 μ l of the resulting cDNA was added to the real-time PCR master mix (Applied Biosystems) and TaqMan primers/probes were used to measure the relative quantities of SLIT3, PIAS1 and TNS3 (VIC labelled) with GAPDH (FAM labelled) as control using an ABI 7900HT machine (Applied Biosystems).

Dataset and data processing

Out of the 108 samples retrieved six samples did not meet quality control and hence did not undergo further analysis. These included 1x Bone Marrow RA low serum, 1x Bone Marrow RA high serum, 1x Skin OA low serum, 1x Synovial OA low serum, 1x Synovial OA high serum and 1x Skin OA high serum sample. Gene expression data from the Affymetrix arrays were processed and normalized using RMA methodology. Data were processed to remove genes with expression close to the detection level.

Identification of differentially expressed genes

Statistical analysis for microarrays (SAM) [42] as implemented by the R package (samr) was used to identify differentially expressed genes between groups. A cut-off of FDR < 3% and an additional cut-off of two fold change was applied on all differentially expressed gene sets. Initially we identified differentially expressed genes for each of the three levels of organisation (anatomical location, serum response and disease type). Here all samples belong to the given level were compared in either a "Multiclass" comparison (anatomical location) or a "Two class unpaired" comparison (serum response and disease type) using the SAM algorithm. The number of differentially expressed genes were then filtered to the previously mentioned thresholds. To identify genes differentially expressed in each tissue and within serum and disease types, we first subset the data into the three tissues. Then for each serum level we identify differentially expressed genes between RA and OA. Similarly for each disease type we identify differentially expressed genes between high and low serum. Apparent discrepancies in total gene numbers in tables compared to the text result from the common phenomenon of genes with multiple Affymetrix. IDs (See S1 Dataset for more details).

K-nearest neighbour classification

To test whether the CSR signature is able to distinguish between normal, RA and OA samples retrieved from the public domain we applied a K-nearest neighbour algorithm as implemented by the knn.cv function in the 'class' library in the statistical environment R. This methodology uses a leave-one-out-crossvalidation. For each sample in the training set the k nearest neighbours are identified and the classification decided by majority vote. We have tested 2 < k < 10 and observed that the accuracy was independent of the value of k (standard deviation of the prediction accuracy over the values of k was < 0.05% in all comparisons).



Data reduction techniques

Genes differentially expressed between all sample groups were visualised using Principal component analysis (PCA), a technique designed to represent highly multi-dimensional objects in a lower dimensionality space (Fig. 2). The first principal component (PC) on the x-axis represents the majority of the variance information, whilst the second PC on the y-axis, is orthogonal to the first PC and holds less variance. All data were mean centred prior to application of PCA using the prcomp function in the statistical environment R [43].

Ingenuity Pathway Analysis (IPA) of differentially expressed gene sets

Gene sets derived from differential gene expression and Venn diagram analyses were analysed using the IPA application (Palo Alto, http://www.ingenuity.com) that uses the Ingenuity Pathways Knowledge Base (IPKB) to explore biological interaction networks. Each gene identifier is mapped to a global molecular network within IPKB. An algorithm then generates networks based on the connectivity of focus genes, limited to a maximum of 35 genes for simpler visualisation. For the downstream effect analysis gene lists and calculated fold changes were uploaded to the IPKB. For each biological function within the IPKB an activation z-score is calculated by the software. Here the molecular network within IPKB is interrogated and experimentally observed causal relationships between genes and function extracted. These connections provide a literature-derived direction of effect which is used in combination with the foldchange to calculate the activation state of the given functional terms. As multiple genes may be connected to a function a statistical approach is necessary to calculate its significancy as compared to a randomly derived variable. This variable is obtained by randomly choosing 1 or -1 with equal probability representing an "increased" or "decreased" state respectively. For each number of genes affecting the functional term the number of "increased" (N_+) and "decreased" (N_-) predictions is summed up for the random. This however assumes that there is only 1 possible downstream effect to a given function. To correct for this the IPAKB uses a weighted z-score statistics in which:

$$w_{i} = \frac{|M_{activating} - M_{inhibiting}|}{M_{activating} + M_{inhibiting} + 1}$$
(1)

where $M_{activating}$ and $M_{inhibiting}$ represent the number of findings underlying an edges "activation" or "inhibition" respectively. This is then used to correct for the direction of change within the z score algorithm. The z-score can therefore be calculated as:

$$z = \frac{x}{\sigma_x} = \frac{\sum_i w_i x_i}{\sqrt{\sum_i w_i^2}} \tag{2}$$

Supporting Information

S1 Fig. Classification of normal and arthritis synovium biospies using a canonical serum response signature. The figure shows the probability of a synovium biopsy sample to be correctly classified as normal, RA or OA on the basis of a canonical serum response signature. The x axis represents each individual sample biopsy in the 3 groups and the y axis the probability of correct classification estimated using a leave one out cross-validation (LOOCV) procedure. The percentage of correct classification using a majority rule is indicated below the sample labels.

(TIFF)



S2 Fig. Venn diagram representing disease specific serum response. A) Shows the Venn diagram for the response to serum in all tissues in OA. B) represents the serum response of all tissues in RA patients. (TIFF)

S1 Table. Patient demographic and disease data. RhF, rheumatoid factor; DMARD, Disease-modifying antirheumatic drug; MTX, methotrexate, weekly dose; D-PEN, D-penicillamine; Gold, intramuscular gold, monthly dose; AZA, azathioprine; SSA, sulphasalazine. (DOCX)

S2 Table. Fold changes in gene expression in OA versus RA synovial fibroblasts in low serum. Data were obtained from 3 synovial fibroblast lines. qPCR, quantitative PCR; PIAS1, protein inhibitor of activated STAT-1 (signal transducer and activator of transcription-1). (DOCX)

S3 Table. Functional associations in serum response. Overview of the Functional associations in the serum response between RA and OA fibroblasts within each anatomical site. The numbers represent "number of genes increased"/"number of genes decreased" within each term.

(DOCX)

S1 Dataset. Results from the differential gene expression analysis. An excel file containing all results from the SAM analysis for all comparisons performed, including up and down regulated genes with their associated Gene ID, Fold Change and q-value (%). (XLSX)

Acknowledgments

Supported by research grants from the Medical Research Council and Arthritis Research UK (Grant 18547). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 305549. AF, CB and KR have received research grant funding from Wyeth. Partially funded by the BBSRC/EraSysBio Linknet BB/I004556/1 grant to FF.

Author Contributions

Conceived and designed the experiments: AF PA GNP HML MO DS KR CDB FF. Performed the experiments: AF PA GNP MJP AMT. Analyzed the data: AF PA FF. Contributed reagents/materials/analysis tools: AF HML MO MJP AMT CDB FF. Wrote the paper: AF PA GNP HML MO MJP AMT DS KR CDB FF.

References

- Müller-Ladner U, Kriegsmann J, Franklin B, Matsumoto S, Geiler T, Gay R, et al. Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice. The American journal of pathology. 1996; 149(5):1607. PMID: 8909250
- Izquierdo E, Cañete JD, Celis R, Del Rey MJ, Usategui A, Marsal S, et al. Synovial fibroblast hyperplasia in rheumatoid arthritis: Clinicopathologic correlations and partial reversal by anti–tumor necrosis factor therapy. Arthritis & Rheumatism. 2011; 63(9):2575–2583. doi: 10.1002/art.30433
- Scaife S, Brown R, Kellie S, Filer A, Martin S, Thomas A, et al. Detection of differentially expressed genes in synovial fibroblasts by restriction fragment differential display. Rheumatology. 2004; 43 (11):1346–1352. doi: 10.1093/rheumatology/keh347 PMID: 15292528
- Del Rey MJ, Usategui A, Izquierdo E, Cañete JD, Blanco FJ, Criado G, et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Annals of the rheumatic diseases. 2012; 71 (2):275–280. doi: 10.1136/annrheumdis-2011-200281 PMID: 22021863



- Chang HY, Chi JT, Dudoit S, Bondre C, Van De Rijn M, Botstein D, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99(20):12877. doi: 10.1073/pnas.162488599 PMID: 12297622
- Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS biology. 2004; 2(2):E7. doi: 10.1371/journal.pbio.0020007 PMID: 14737219
- Chang HY, Nuyten DSA, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proceedings of the National Academy of Sciences. 2005; 102(10):3738. doi: 10.1073/pnas. 0409462102
- Bauer S, Jendro M, Wadle A, Kleber S, Stenner F, Dinser R, et al. Fibroblast activation protein is expressed by rheumatoid myofibroblast-like synoviocytes. Arthritis Research & Therapy. 2006; 8(6): R171. doi: 10.1186/ar2080
- Ohshima S, Kuchen S, Seemayer CA, Kyburz D, Hirt A, Klinzing S, et al. Galectin 3 and its binding protein in rheumatoid arthritis. Arthritis & Rheumatism. 2003; 48(10):2788–2795. doi: 10.1002/art.11287
- Ošlejšková L, Grigorian M, Gay S, Neidhart M, Šenolt L. The metastasis associated protein S100A4: a
 potential novel link to inflammation and consequent aggressive behaviour of rheumatoid arthritis synovial fibroblasts. Annals of the rheumatic diseases. 2008; 67(11):1499. doi: 10.1136/ard.2007.079905
 PMID: 18056757
- Lefèvre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R, et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nature medicine. 2009; 15(12):1414–1420. doi: 10.1038/nm. 2050 PMID: 19898488
- 12. Kasperkovitz, P, Verbeet, N, Smeets, T, Tak, P, Huizinga, T, Baltus, B, et al. Fibroblast-like synoviocytes derived from patients with rheumatoid arthritis show the imprint of synovial tissue heterogeneity: evidence for the existence of distinctive pathways relevant to disease. 2004;.
- Bartok B, Firestein GS. Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis. Immunological reviews. 2010; 233(1):233–255. doi: 10.1111/j.0105-2896.2009.00859.x PMID: 20193003
- 14. 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 Research and Therapy. 2008; 10(4):R98. doi: 10.1186/ar2485 PMID: 18721452
- Dvorak H. Tumors: wounds that do not heal: similarities between tumor stroma generation and wound healing. The New England journal of medicine. 1986; 315(26):1650–1659. doi: 10.1056/ NEJM198612253152606 PMID: 3537791
- Parsonage G, Falciani F, Burman A, Filer A, Ross E, Bofill M, et al. Global gene expression profiles in fibroblasts from synovial, skin and lymphoid tissue reveals distinct cytokine and chemokine expression patterns. Thromb Haemost. 2003; 90(4):688–697. PMID: 14515190
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. Cell. 2007; 129(7):1311–1323. doi: 10.1016/j.cell.2007.05.022 PMID: 17604720
- Watanabe N, Ando K, Yoshida S, Inuzuka S, Kobayashi M, Matsui N, et al. Gene expression profile analysis of rheumatoid synovial fibroblast cultures revealing the overexpression of genes responsible for tumor-like growth of rheumatoid synovium. Biochemical and biophysical research communications. 2002; 294(5):1121–1129. doi: 10.1016/S0006-291X(02)00608-3 PMID: 12074593
- Pap T, Franz JK, Hummel KM, Jeisy E, Gay R, Gay S. Activation of synovial fibroblasts in rheumatoid arthritis: lack of expression of the tumour suppressor PTEN at sites of invasive growth and destruction. Arthritis research. 2000; 2(1):59–64. doi: 10.1186/ar69 PMID: 11219390
- Dickinson R, Dallol A, Bieche I, Krex D, Morton D, Maher E, et al. Epigenetic inactivation of SLIT3 and SLIT1 genes in human cancers. British journal of cancer. 2004; 91(12):2071–2078. doi: 10.1038/sj.bjc. 6602222 PMID: 15534609
- Denk A, Kaufmann S, Stark K, Schedel J, Lowin T, Schubert T, et al. Slit3 inhibits Robo3-induced invasion of synovial fibroblasts in rheumatoid arthritis. Arthritis Research & Therapy. 2010; 12(2):R45. doi: 10.1186/ar2955
- 22. Nguyen TDT, Ma LN, Slovak ML, Bangs CD, Cherry AM, Arber DA. Identification of novel Runx1 (AML1) translocation partner genes SH3D19, YTHDf2, and ZNF687 in acute myeloid leukemia. Genes, Chromosomes and Cancer. 2006; 45(10):918–932. doi: 10.1002/gcc.20355 PMID: 16858696
- Munoz MA, Saunders DN, Henderson MJ, Clancy JL, Russell AJ, Lehrbach G, et al. The E3 Ubiquitin Ligase EDD Regulates S-Phase and G2/M DNA Damage Checkpoints. Cell cycle. 2007; 6(24):3070– 3077. doi: 10.4161/cc.6.24.5021 PMID: 18073532



- Feghali-Bostwick CA, Gadgil AS, Otterbein LE, Pilewski JM, Stoner MW, Csizmadia E, et al. Autoantibodies in patients with chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine. 2008; 177(2):156–163. doi: 10.1164/rccm.200701-014OC PMID: 17975205
- Seki T, Selby J, H"aupl T, Winchester R. Use of a differential subtraction method to identify genes that characterize the phenotype of cultured rheumatoid arthritis synoviocytes. Arthritis & Rheumatism. 1998; 41(8):1356–1364. doi: 10.1002/1529-0131(199808)41:8%3C1356::AID-ART4%3E3.0.CO;2-X
- Kobayashi N, Kostka G, Garbe JHO, Keene DR, B"achinger HP, Hanisch FG, et al. A Comparative Analysis of the Fibulin Protein Family. Journal of Biological Chemistry. 2007; 282(16):11805. doi: 1074/jbc.M611029200 PMID: 17324935
- 27. Brooks PC, Strömblad S, Sanders LC, von Schalscha TL, Aimes RT, Stetler-Stevenson WG, et al. Localization of matrix metalloproteinase MMP-2 to the surface of invasive cells by interaction with integrin [alpha] v [beta] 3. Cell. 1996; 85(5):683–693. doi: 10.1016/S0092-8674(00)81235-0 PMID: 8646777
- Roberts MS, Woods AJ, Shaw PE, Norman JC. ERK1 associates with V 3 integrin and regulates cell spreading on vitronectin. Journal of Biological Chemistry. 2003; 278(3):1975. doi: 10.1074/jbc. M208607200 PMID: 12393886
- Ding Q, Stewart J, Olman MA, Klobe MR, Gladson CL. The pattern of enhancement of Src kinase activity on platelet-derived growth factor stimulation of glioblastoma cells is affected by the integrin engaged. Journal of Biological Chemistry. 2003; 278(41):39882. doi: 10.1074/jbc.M304685200 PMID: 12881526
- Graf K, Neuss M, Stawowy P, Hsueh WA, Fleck E, Law RE. Angiotensin II and {alpha} v {beta} 3 Integrin Expression in Rat Neonatal Cardiac Fibroblasts. Hypertension. 2000; 35(4):978. doi: 10.1161/01. HYP.35.4.978 PMID: 10775572
- Edderkaoui M, Hong P, Lee JK, Pandol SJ, Gukovskaya AS. Insulin-like growth factor-I receptor mediates the prosurvival effect of fibronectin. Journal of Biological Chemistry. 2007; 282(37):26646. doi: 1074/jbc.M702836200 PMID: 17627944
- 32. Lainer-Carr D, Brahn E. Angiogenesis inhibition as a therapeutic approach for inflammatory synovitis. Nature Clinical Practice Rheumatology. 2007; 3(8):434–442. doi: 10.1038/ncprheum0559 PMID: 17664950
- Hynes RO, et al. A reevaluation of integrins as regulators of angiogenesis. Nature medicine. 2002; 8 (9):918–921. doi: 10.1038/nm0902-918 PMID: 12205444
- 34. Kanamori M, Berg SRV, Bergers G, Berger MS, Pieper RO. Integrin {beta} 3 Overexpression Suppresses Tumor Growth in a Human Model of Gliomagenesis: Implications for the Role of {beta} 3 Overexpression in Glioblastoma Multiforme. Cancer research. 2004; 64(8):2751. doi: 10.1158/0008-5472. CAN-03-3354 PMID: 15087390
- 35. Oka N, Nakahara S, Takenaka Y, Fukumori T, Hogan V, Kanayama Ho, et al. Galectin-3 Inhibits Tumor Necrosis Factor–Related Apoptosis-Inducing Ligand–Induced Apoptosis by Activating Akt in Human Bladder Carcinoma Cells. Cancer Research. 2005; 65(17):7546–7553. PMID: 16140916
- Thalhamer T, McGrath M, Harnett M. MAPKs and their relevance to arthritis and inflammation. Rheumatology. 2008; 47(4):409–414. doi: 10.1093/rheumatology/kem297 PMID: 18.187523
- Da Wei Huang BTS, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols. 2008; 4(1):44–57. doi: 10.1038/nprot.2008.211
- **38.** Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol. 2003; 4(5):P3. doi: 10.1186/gb-2003-4-5-p3
- Filer A, Parsonage G, Smith E, Osborne C, Thomas AMC, Curnow SJ, et al. Differential Survival of Leukocyte Subsets Mediated by Synovial, Bone Marrow, and Skin Fibroblasts. Arthritis & Rheumatism. 2006; 54(7):2096–2108. doi: 10.1002/art.21930
- Arnett FC, Edworthy SM, Bloch DA, Mcshane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis & Rheumatism. 2005; 31(3):315–324. doi: 10.1002/art.1780310302
- 41. Van Gestel A, Prevoo M, Van't Hof M, Van Rijswijk M, Van de Putte L, Van Riel P. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis: comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism criteria. Arthritis & Rheumatism. 1996; 39(1):34–40.
- **42.** Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to transcriptional responses to ionizing radiation. Proc Natl Acad Sci USA. 2001; 98:5116–5121. doi: 10.1073/pnas.091062498 PMID: 11309499
- **43.** R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2009. ISBN 3-900051-07-0 Available from: http://www.R-project.org.