



## Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients

T C T M van der Pouw Kraan, C A Wijbrandts, L G M van Baarsen, A E Voskuyl, F Rustenburg, J M Baggen, S M Ibrahim, M Fero, B A C Dijkmans, P P Tak and C L Verweij

*Ann Rheum Dis* 2007;66:1008-1014; originally published online 12 Jan 2007; doi:10.1136/ard.2006.063412

---

Updated information and services can be found at:  
<http://ard.bmj.com/cgi/content/full/66/8/1008>

---

*These include:*

**Data supplement**

*"web only appendix"*

<http://ard.bmj.com/cgi/content/full/ard.2006.063412/DC1>

**References**

This article cites 43 articles, 19 of which can be accessed free at:

<http://ard.bmj.com/cgi/content/full/66/8/1008#BIBL>

**Rapid responses**

You can respond to this article at:

<http://ard.bmj.com/cgi/eletter-submit/66/8/1008>

**Email alerting service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

---

**Notes**

---

To order reprints of this article go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Annals of the Rheumatic Diseases* go to:  
<http://journals.bmj.com/subscriptions/>

## EXTENDED REPORT

# Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients

T C T M van der Pouw Kraan, C A Wijbrandts, L G M van Baarsen, A E Voskuyl, F Rustenburg, J M Baggen, S M Ibrahim, M Fero, B A C Dijkmans, P P Tak, C L Verweij

*Ann Rheum Dis* 2007;**66**:1008–1014. doi: 10.1136/ard.2006.063412

See end of article for authors' affiliations

Correspondence to:  
Dr C L Verweij, VU  
University Medical Centre,  
Department of Molecular  
and Cellular Biology &  
Immunology, J295, PO Box  
7057, 1007 MB  
Amsterdam, The  
Netherlands;  
[c.verweij@vumc.nl](mailto:c.verweij@vumc.nl)

Accepted 16 December 2006  
Published Online First  
18 January 2007

**Background:** Rheumatoid arthritis (RA) is a heterogeneous disease with unknown cause.

**Aim:** To identify peripheral blood (PB) gene expression profiles that may distinguish RA subtypes.

**Methods:** Large-scale expression profiling by cDNA microarrays was performed on PB from 35 patients and 15 healthy individuals. Differential gene expression was analysed by significance analysis of microarrays (SAM), followed by gene ontology analysis of the significant genes. Gene set enrichment analysis was applied to identify pathways relevant to disease.

**Results:** A substantially raised expression of a spectrum of genes involved in immune defence was found in the PB of patients with RA compared with healthy individuals. SAM analysis revealed a highly significant elevated expression of interferon (IFN) type I regulated genes in patients with RA compared with healthy individuals, which was confirmed by gene ontology and pathway analysis, suggesting that this pathway was activated systemically in RA. A quantitative analysis revealed that increased expression of IFN-response genes was characteristic of approximately half of the patients (IFN<sup>high</sup> patients). Application of pathway analysis revealed that the IFN<sup>high</sup> group was largely different from the controls, with evidence for upregulated pathways involved in coagulation and complement cascades, and fatty acid metabolism, while the IFN<sup>low</sup> group was similar to the controls.

**Conclusion:** The IFN type I signature defines a subgroup of patients with RA, with a distinct biomolecular phenotype, characterised by increased activity of the innate defence system, coagulation and complement cascades, and fatty acid metabolism.

Rheumatoid arthritis (RA) is a systemic autoimmune disease characterised by chronic inflammation of the joints. There is growing evidence that patients with RA, as defined by the American College of Rheumatology Classification Criteria,<sup>1</sup> represent a highly heterogeneous group. However, a clinical approach to disease classification could erroneously suggest that the criteria be applied to classify one disease entity. The heterogeneity of RA is indicated by notable variability in clinical presentation, and the presence of distinct autoantibody specificities, such as rheumatoid factor and anticyclic citrullinated peptide antibodies in the serum.<sup>2–3</sup>

Disease heterogeneity is also apparent in histological features of the synovium, which shows different complexity levels of lymphoid organisation in subsets of patients.<sup>4–5</sup> Moreover, gene expression profiling of synovial tissue from patients with RA and those with osteoarthritis revealed marked variation in gene expression profiles that allowed us to identify molecularly distinct forms of RA synovium.<sup>6–7</sup>

The wide variation in responsiveness to virtually any treatment in RA is also consistent with the heterogeneous nature of the disease.<sup>8–10</sup> Together, these findings suggest that distinct disease mechanisms are at play in RA pathology. The relative contribution of the different mechanisms may vary between patients and, perhaps, in different stages of disease.

The heterogeneity most likely has its origin in the multifactorial nature of the disease, whereby specific combinations of environmental factors and a varying polygenic background are likely to influence not only susceptibility but also the severity and disease outcome. Findings from genetically identical twins, in whom the concordance rate is far less than

complete, are indicative of a major role of an environmental factor in the risk of developing RA.<sup>11–12</sup>

Given the heterogeneous nature of RA, and its systemic features, we investigated whether this heterogeneity is reflected in peripheral blood (PB) cells, because it can be anticipated that (aetio)pathogenic events in the host are reflected as phenotypic changes in the host cells. Large-scale gene expression profiling of PB cells from patients with RA could thus provide a molecular portrait that reflects the contributions of diverse cellular responses associated with RA, in general, and with disease subtypes, and could thus define the samples' unique biology.

## MATERIALS AND METHODS

### Patients and controls

PB was obtained in PAXgene RNA isolation tubes (PreAnalytix GmbH, Hombrechtikon, Switzerland) from 35 patients. Of these, 25 used methotrexate (MTX), and 10 patients were MTX and other DMARD naïve. All patients fulfilled the revised American College of Rheumatology 1987 criteria for RA,<sup>1</sup> except for three patients in the MTX naïve group, who were diagnosed as having probable RA with a disease duration of 3–12 months, as having mono-arthritis and as being positive for anticyclic citrullinated peptide antibodies. Two patients were diagnosed with RA after 6 and 12 months. Table 1 summarises the characteristics of these patients.

**Abbreviations:** DC, dendritic cell; FLS, fibroblast-like synoviocytes; GSEA, gene set enrichment analysis; IFN, interferon; MTX, methotrexate; PB, peripheral blood; RA, rheumatoid arthritis; SAM, significance analysis of microarrays; SLE, systemic lupus erythematosus; SS, Sjögren's syndrome

**Table 1** Patient characteristics

	Patients with RA receiving MTX treatment n = 25	Patients with RA not receiving MTX treatment n = 10
Age	49 (23–63)	49 (25–67)
Sex: female, male	17, 8	8, 2
Disease characteristics		
CRP	19 (3–76)	ND
ESR	27 (2–82)	41 (13–70)
Rheumatoid factor titre	114 (1–516)*	123 (12–832)
ACPA titre	935 (1–6264)*	1210 (1–4904)
Disease duration (months)	113 (8–417)	7 (1–12)
Erosions (n)	24	1
DAS28	5.5 (3.4–7.2)	5.1 (3.1–7.4)
Medication		
MTX dose (range)	21 (7.5–30)	0
Prednisone $\leq$ 10 mg/day (n)	6	0

ACPA, anticyclic citrullinated peptide antibodies; CRP, C-reactive protein; DAS28, Disease Activity Score 28; ESR, erythrocyte sedimentation rate; MTX, methotrexate; ND, not determined; RA, rheumatoid arthritis. Values are given as mean (range) unless otherwise specified. \*Available for 23 of 25 patients receiving MTX treatment.

The control group consisted of 15 healthy individuals (nine women, six men, mean age: 43 years, range 27–63). In all comparisons mentioned, the groups were age- and sex-matched. All patients and controls gave their informed consent, and the study protocol was approved by the medical ethics committees from the Academic Medical Center and VU Medical Centre.

### Sample preparation, labelling and hybridisation

This procedure was performed as described previously.<sup>13</sup> In brief, total RNA was isolated from PB using the PAXgene RNA isolation kit. Amplified RNA was labelled with aminoallyl-deoxyuridine-5-triphosphate during cDNA synthesis, followed by chemical coupling of the aminoallyl group to Cy3 or Cy5 for the experimental and reference samples, respectively. The labelled cDNA transcripts were hybridised together on human cDNA microarrays with 43 000 elements, representing about 26 000 genes, generated at Stanford University, as described.<sup>14</sup>

### Data filtering and analysis

Data were stored and preanalysed in the Stanford Microarray Database<sup>15</sup> at <http://genome-www.stanford.edu/microarray>. Data are expressed as  $\log_2$  ratios of fluorescence intensities of the experimental and the common reference sample. Intensity-dependent normalisation using local estimation (Loess) was performed separately on each sector of the array. Spots were included in the analysis when in at least 80% of the microarrays a reliable data point was obtained for that element (defined by a regression correlation coefficient  $R > 0.6$  for Cy3 and Cy5 pixel intensities, and a signal intensity of 2.5 times the local background for both Cy3 and Cy5). The use of a common reference allowed comparison of the expression levels across all samples.<sup>13</sup> Therefore, the expression levels (as  $\log_2$  ratios) were median centred—that is, each spotted element was expressed relative to the median expression level of that element across all samples. We corrected for array batch differences by applying “singular value decomposition”.<sup>16</sup> Genes represented more than once on the microarrays were averaged in the Stanford Microarray Database from sequences with the same Unigene identifier.

### Statistical analysis

Statistical analysis on microarray data was performed using significance analysis of microarrays (SAM).<sup>17</sup> Genes that were

expressed at significantly different levels between patients and controls, defined by a q-value of  $< 5\%$ , were analysed by supervised hierarchical clustering<sup>18</sup> to visualise the correlation of co-expressed genes in Treeview (available at <http://rana.lbl.gov/EisenSoftware.htm>).

For an interpretation of the biological processes that are represented by the genes that show a significantly different level of expression in patients with RA compared with the controls, we applied gene ontology analysis in the PANTHER database at <http://panthedb.org>.<sup>19</sup> PANTHER uses the binomial statistics tool to compare our gene list with a reference list (NCBI: Homo sapiens genes) to determine the statistically significant over-representation of functional groups of genes. A Bonferroni correction was applied to adjust for multiple comparisons. p Values  $< 0.05$  were considered significant.

For pathway analysis, we used gene set enrichment analysis (GSEA)<sup>20</sup> at <http://www.broad.mit.edu/gsea/>. Like SAM, it utilises data permutation to adjust for multiple testing, indicated by a false discovery rate. A total of 408 pathways from the Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg>) and Biocarta (<http://www.biocarta.com>) are applied in this analysis. The same gene may be present in more than one pathway or biological process. In addition, we incorporated several interferon (IFN)-response gene sets from published data.<sup>21, 22</sup> A minimal gene set size of 20 genes per pathway was applied, and pathways with a p value  $< 0.05$  and a false discovery rate of  $< 0.25$  were considered significant, according to the authors' suggestions.<sup>20</sup> For the comparison of mean gene expression levels in different gene sets, a Student's t test was used.

## RESULTS

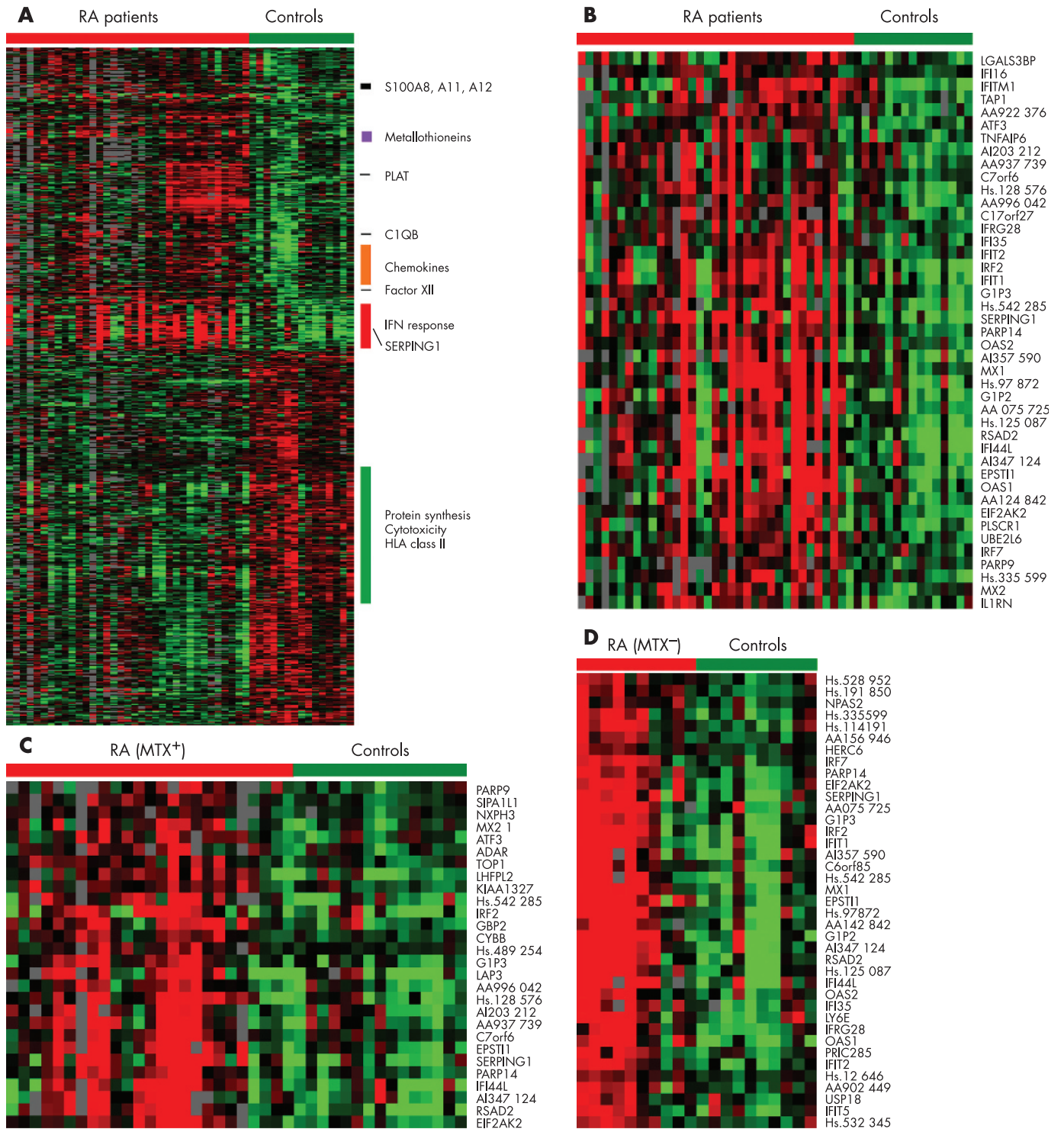
### Gene expression profiling in PB cells of patients with RA

Gene expression profiling of PB cells from 32 patients with RA, 3 patients with probable RA, and 15 age- and sex-matched healthy controls was performed on microarrays with a complexity of  $\sim 26\text{K}$  unique genes (43K elements). Data were analysed as, unpaired data belonging to two classes using SAM.<sup>17</sup> A total of 577 genes, of which 259 were upregulated and 318 genes downregulated, were selected, whose transcript levels were expressed at significantly different levels between the two groups. The significant gene expression differences between patients with RA and healthy controls were visualised in a heatmap (fig 1A).<sup>18</sup>

### Genes upregulated in RA

A global view of the significantly differential expressed genes revealed a prominent cluster of IFN-inducible genes that was upregulated in patients with RA. This cluster, highlighted in fig 1B, contains highly correlated genes such as IFRG28 (28 kDa interferon responsive protein), IFI35 (interferon-induced protein 35), IFI44L (interferon-induced protein 44-like), IFIT1 (interferon-induced protein with tetratricopeptide repeats 1), IFIT2, IRF2 (interferon regulatory factor 2), IRF7, GIP2 (interferon  $\alpha$ -inducible protein 2), GIP3, SERPING1 (serine proteinase inhibitor clade G member 1, C1 inhibitor), OAS1 (2'-5'-oligoadenylate synthetase 1), OAS2, MX1 (Myxovirus resistance 1), GIP2/ISG15 (interferon-induced protein 15) and RSAD2 (radical S-adenosyl methionine domain containing 2).

In addition, all patients showed increased expression of several inflammatory mediators, including the chemokines CXCL12, CXCL9, CCL15, CCL19, CCL7, CXCL12, CCL19, CCL7, CXCL3 and CCL8, as well as interleukin (IL)19 and the S100 family proteins S100 calcium-binding protein A8 (S100A8), S100A11 and S100A12. Other genes that were upregulated in



**Figure 1** (A) Cluster diagram of the expression of 577 significantly different by expressed genes in 35 patients and 15 healthy individuals. Genes are organised by hierarchical clustering, based on the overall similarity in expression patterns.<sup>18</sup> Red represents relative expression greater than the median expression level across all samples, and green represents an expression level lower than the median. Black indicates intermediate expression. Grey indicates missing data. Coloured bars to the right identify the locations of a category of clustered genes, with a correlated expression profile and related function. (B) Representation of the interferon (IFN)-response gene cluster with an enhanced expression in the group having patients with rheumatoid arthritis (RA). An expanded view of the genes in the IFN-response cluster of (A) is shown. Genes are either known genes with a unigene symbol characteristic for the defined gene cluster, or unknown, indicated by an accession number or unigene cluster ID. (C, D) The IFN-response programme is present in patients with RA irrespective of treatment. Representation of genes that are expressed at significantly different levels between patients with RA undergoing (C) or patients not undergoing (D) methotrexate (MTX) treatment and age- and sex-matched healthy controls. A selection of genes with a correlated expression profile that are indicative of an IFN-response programme is shown.

patients with RA were members of the antioxidant metallothionein family and the anti-inflammatory IL1 receptor antagonist.

#### Genes downregulated in RA

Genes that showed a lower expression in RA included CD3 $\zeta$ , TCR $\beta$  chain, TARP/TCR $\gamma$ V9, granzyme M, runx3 and KLRB1,

which are involved in cytotoxic functions, and many other genes with unknown function.

### Gene ontology analysis of genes with significant differential expression in RA

To systematically categorise the 577 genes with significant differential expression into functional groups, we used the PANTHER database consisting of a large collection of protein families that have been subdivided into functionally related subfamilies.<sup>19</sup> The differentially upregulated genes represented seven significant functional biological processes (table 2). There were no significant downregulated ontology groups. The “immunity and defense” ontology group represented a broad composite family that consists of more specified ontology subgroups. Within these subgroups, the most significant upregulated process that distinguished patients with RA from controls was “interferon-mediated immunity”.

### Pathway analysis

In addition, we performed GSEA<sup>20</sup> to identify pathways relevant to RA. In contrast to ontology analysis, this algorithm is based on the usage of all available gene expression data and derives its power from the analysis of sets of genes that are coordinately regulated in a defined biological process or pathway, while it uses data permutation to adjust for multiple testing. In addition to the intrinsic GSEA pathway gene sets, we included previously reported IFN-response sets in our analysis.<sup>21–22</sup> The results revealed that, besides the five GSEA intrinsic gene sets (table 3), the previously described type I IFN-induced genes by Baechler *et al*<sup>22</sup> (in their supplementary data) and the IFN $\alpha$ -induced genes were both significantly increased in patients with RA.

### IFN-induced genes in RA

We confirmed expression of key genes of the IFN pathway, RSAD2 and G1P2, in all samples by real-time PCR, which showed a high correlation with the microarray data ( $r = 0.78$  and  $0.87$ , respectively,  $p < 0.001$  in both cases, data not shown). To rule out an effect of MTX treatment on the IFN-induced genes, we compared patients undergoing ( $n = 25$ ) and not undergoing ( $n = 10$ ) MTX treatment with the appropriate age- and sex-matched controls. SAM revealed that both groups of patients showed a prominent IFN-induced gene expression signature (fig 1C, D). Thus, the IFN expression signature was present in patients with RA irrespective of MTX treatment.

The expression profiles of the three patients with probable RA, within the MTX naive group, did not differ from those of the other MTX-naive patients. Meanwhile, two of the patients with probable RA were diagnosed with “definite” RA at 6 and 12 months after Paxgene blood sampling, respectively, suggesting that the RA signature was present in the blood before diagnosis.

### Selective upregulation of type I IFN-response genes in RA

Type I IFNs are mainly produced directly after viral infection, whereas type II IFNs are secondarily produced by activated T- and natural killer cells. Type I and type II IFN response programmes share many of their genes. To disclose information on the inducing type of IFN, we obtained a specific type I IFN- and type II IFN-response gene set<sup>21</sup> (supplementary table 1). The type I IFN-response set consists of five genes that respond to both IFN $\alpha$  and IFN $\beta$ , but not to IFN $\gamma$ . The type II IFN-response set consists of 13 genes responding specifically to IFN $\gamma$ .<sup>21</sup> To investigate the relative contribution of either gene set to the RA gene expression profile, we calculated for each gene set the mean gene expression level ( $\log_2$  ratio) per patient and per healthy control, and compared the two groups with each other (fig 2). This analysis showed that the mean gene expression level of the type I IFN gene set was significantly higher in the group of patients with RA ( $p < 0.001$ ), whereas the mean gene expression level of the type II interferon (IFN) genes was similar between patients and controls. Hence, these findings provide evidence that type I IFNs rather than type II IFNs are responsible for the increased expression of IFN-induced genes.

### The IFN signature defines a subgroup of patients with RA

Consistent with the heterogeneous nature of RA, we observed that the IFN response showed a large variation between patients with RA (figs 2 and 3). To obtain more insight into the differential expression of IFN-induced genes in individual patients, we calculated for each individual the average expression of the IFN-cluster genes that were upregulated in patients with RA, as described in fig 1B. Next, we defined which patients show an altered IFN response, by calculating the 95% limits of the controls (normal values, defined as the mean (SD) expression of the 43 IFN genes,  $\pm 1.96$  SD). We identified 20 patients with an average expression level above normal

**Table 2** Ontology analysis of the genes that were expressed at significantly different levels between all 35 patients, or the subgroups consisting of 20 IFN<sup>high</sup> and 15 IFN<sup>low</sup> patients, and healthy controls

Biological process	All patients p Value	IFN <sup>high</sup> p Value	IFN <sup>low</sup> p Value
Immunity and defence	1.61E-11	3.41E-15	NS
Interferon-mediated immunity	1.09E-06	3.38E-08	NS
Macrophage-mediated immunity	6.83E-05	1.27E-04	NS
Lipid and fatty acid transport	1.54E-04	NS	NS
Transport	4.82E-04	3.33E-02	NS
Cytokine/chemokine-mediated immunity	1.63E-03	2.61E-05	NS
Ligand-mediated signalling	3.52E-03	1.92E-02	NS
Cell motility	NS	4.24E-03	NS
Blood clotting	NS	5.30E-03	NS
Oncogenesis	NS	1.04E-02	NS
Cell structure and motility	NS	1.55E-02	NS
Inhibition of apoptosis	NS	1.66E-02	NS
Signal transduction	NS	3.21E-02	NS
Proteolysis	NS	3.97E-02	NS

NS, nonsignificant.

Processes with a p value  $< 0.05$  were considered significant (with Bonferroni correction).

**Table 3** Pathways which are overexpressed in all patients, and in the subgroups of IFN<sup>high</sup> and IFN<sup>low</sup> patients, all compared with healthy controls, analysed by gene set enrichment analysis

Pathway	No. of genes	All patients p Value	IFN <sup>high</sup> p Value	IFN <sup>low</sup> p Value
Type I IFN-induced genes <sup>22</sup>	83	0.004	<0.001	NS
IFN $\alpha$ -induced <sup>21</sup>	36	0.022	<0.001	NS
IFN $\beta$ -induced <sup>21</sup>	51	NS	<0.001	NS
Cytokine–cytokine receptor interaction	129	0.006	0.005	NS
Neuroactive ligand–receptor interaction	90	0.009	NS	<0.001
Jak-STAT signalling pathway	86	0.016	0.008	NS
Complement and coagulation cascades	44	0.025	0.009	NS
Fatty acid metabolism	41	0.027	0.015	NS
IFN $\gamma$ -induced <sup>21</sup>	34	NS	0.019	NS

IFN, interferon; NS, nonsignificant.

Pathways are shown with a p value <0.05, and a false discovery rate value <0.25.

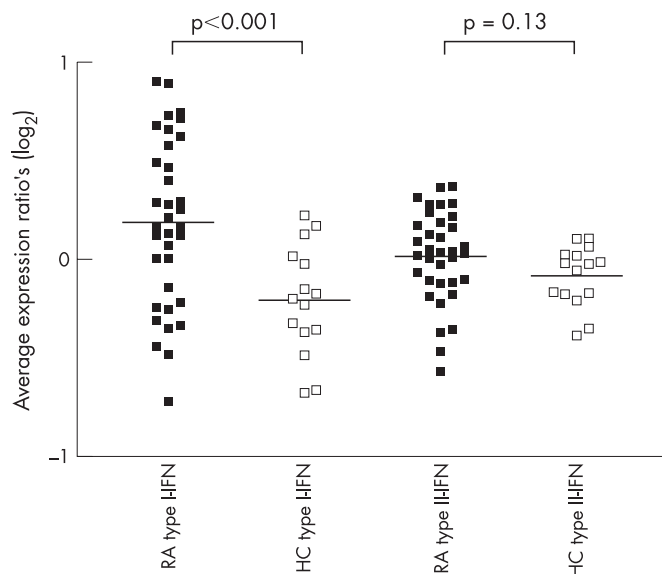
values, further defined as the IFN<sup>high</sup> group, while the remainder of the patients, with an expression level equal to that of controls, was defined as the group IFN<sup>low</sup> (fig 3).

### Distinct characteristics of the IFN<sup>high</sup> group

To further characterise the IFN<sup>high</sup> group, we performed SAM analysis, which revealed that 484 genes were upregulated in IFN<sup>high</sup> patients, compared with the healthy controls, while 229 genes were downregulated. The same analysis for the IFN<sup>low</sup> patients revealed only 57 upregulated genes and 93 downregulated genes. These data indicate that, within the patients with RA, the patients with an IFN signature represent the most distinct group compared with normal controls.

When we applied gene ontology analysis, we found that nearly all of the processes that were identified as upregulated in the whole RA group were also upregulated in the IFN<sup>high</sup> group. Moreover, an additional group of 10 biological processes were selectively upregulated in the IFN<sup>high</sup> group (table 2). No downregulated processes were identified.

Gene ontology analysis of the IFN<sup>low</sup> group revealed no significant downregulated or upregulated processes.



**Figure 2** Type I interferon (IFN)-induced genes are overexpressed in rheumatoid arthritis (RA). Each square indicates the mean expression levels of genes known to be specifically induced by either type I (13 genes) or type II IFN (5 genes) per individual patient or per healthy control (HC). These genes are extracted from the gene sets used for pathway analysis in table 3.<sup>21</sup>

In accordance with gene ontology analysis, pathway analysis by GSEA revealed that the IFN<sup>high</sup> patients were responsible for the upregulated pathways in the overall RA group (table 3). This was particularly clear for the IFN-type I induced gene sets, complement and coagulation cascades. On the other hand, the IFN<sup>low</sup> group was associated with increased expression of the “neuroactive ligand–receptor interaction” pathway. We did not identify any downregulated pathways in either group of patients with RA. Overall, these analyses indicate that, within the whole group of patients, the IFN<sup>high</sup> group is more distinct from controls than the IFN<sup>low</sup> group.

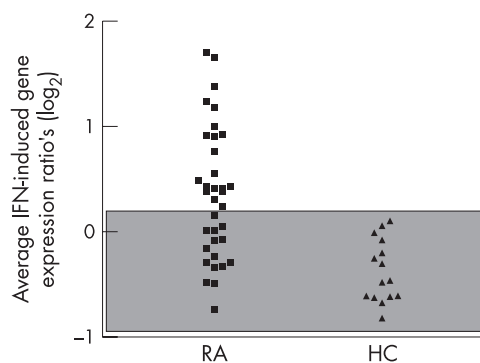
The molecular stratification of RA was not associated with clinical parameters described in table 1.

### DISCUSSION

Since RA is a systemic disease, several investigators addressed the question whether disease characteristics are reflected by changes in gene expression levels in PB cells. Whereas these studies provided insight into the genes that were differentially expressed between patients with RA and healthy controls, the issue of transcript-based disease heterogeneity has not been addressed so far, except for a comparison between recent-onset arthritis and longstanding disease.<sup>23</sup>

Using large-scale gene expression profiling, we identified a large number of genes, including genes involved in the immune/inflammatory response, such as the previously described calcium-binding proteins S100A8, S100A12 and IL1RA.<sup>24–25</sup> Pathway level analysis was used to classify gene expression data in biological processes and pathways. The clear induction of IFN-response genes in patients with RA prompted us to incorporate several IFN-response gene sets from published data<sup>21–22</sup> in the analysis. This analysis revealed that the type I IFN-mediated immunity was the most significantly upregulated pathway in patients with RA, independent of MTX treatment. Albeit that inclusion of the type I IFN gene set is a biased decision, this approach provides a method to demonstrate the significance of the type I IFN response programme in RA.

Most interestingly, our analysis revealed a striking heterogeneity between patients with RA on the basis of the differential expression of genes involved in the innate defence system—in particular, the type I IFN system. These findings suggest that different pathogenic mechanisms may contribute to the disease. The IFN<sup>high</sup> group was further characterised by gene sets reflecting increased activity of complement and coagulation cascades.<sup>26</sup> Next to complement activation, the other pathways associated with the IFN type I signature, such as “fatty acid metabolism” and “coagulation” may all contribute to the increased risk for cardiovascular disease in a subgroup of patients with RA.<sup>27</sup>



**Figure 3** A subgroup of patients with rheumatoid arthritis (RA) show increased expression of interferon (IFN)-response genes (IFN<sup>high</sup>). Each square represents a single individual with the average expression ratio of all 43 IFN-response genes, which are shown as a distinct cluster in fig 1A and B. The shaded box indicates the normal range within the 95% confidence limits. Patients with RA outside the shaded box are defined as the IFN<sup>high</sup> group. HC, healthy control.

The most significant genes from the complement and coagulation pathway are indicated in fig 1, including complement subcomponent C1q chain B (C1QB), coagulation factor XII (F12), tissue plasminogen activator (PLAT) and serpin peptidase inhibitor, clade G (C1 inhibitor), member 1 (SERPING1). These genes are involved in activation as well as inhibitory components of the pathways.

Upregulation of IFN-induced genes has also been observed in PB cells of (a subset of) patients with other autoimmune diseases like systemic lupus erythematosus (SLE),<sup>22–28</sup> systemic sclerosis,<sup>29</sup> multiple sclerosis,<sup>13</sup> and in tissues from patients with Sjögren's syndrome (SS),<sup>30</sup> type I diabetes mellitus<sup>31</sup> and dermatomyositis.<sup>32</sup> These findings suggest that an activated IFN gene expression programme is a common hallmark in chronic autoimmune diseases.

Type I IFNs, which are the early mediators of the innate immune response that influence the adaptive immune response through direct and indirect actions on dendritic cells (DCs), T- and B cells, and natural killer cells, could affect the initiation or amplification of autoimmunity and tissue damage through their diverse and broad actions on almost every cell type and promotion of T helper 1 responses.<sup>33</sup> This appears to be the case for SLE, but for RA both clinical and pathophysiological data have suggested that tumour necrosis factor alpha (TNF $\alpha$ ) rather than type I IFN is essential for persistence of the disease. Hence, it is believed that mutually exclusive cytokine expression patterns are characteristic for distinct autoimmune diseases. However, since we observed an IFN type I signature in the PB of a subgroup of patients with RA this could mean that cytokine profiles are a patient-specific rather than a disease-specific phenomenon.

In patients with SLE, the IFN signature is related to disease severity.<sup>22</sup> It is at present unclear what the role of type I IFNs in RA pathogenesis could be. In analogy to systemic sclerosis<sup>29</sup> and multiple sclerosis,<sup>13</sup> no clinical associations were found for RA so far. We have previously suggested that IFN/STAT-1 activation in RA synovium could be a reactive attempt to limit inflammation.<sup>34</sup> This suggestion was recently supported by a study showing that IFN $\beta$  deficiency could prolong experimental arthritis and resulted in increased activation of fibroblast-like synoviocytes (FLS) *in vitro*.<sup>35</sup> In addition, IFN $\beta$ -competent fibroblasts were able to ameliorate arthritis in IFN $\beta$ -deficient recipients. It should be noted, however, that systemic administration of IFN $\beta$  was unsuccessful in the treatment of RA, which may be due to pharmacokinetic issues.<sup>36</sup>

Concerning the origin of type I IFNs, infectious and endogenous agents, such as viruses, bacteria, unmethylated CpG DNA, single- or dsRNA, heat shock protein 60, or fibrinogen fragments, could all be proximal mediators of type I IFN production and thus lead to the more downstream activation programme. In sera from patients with SLE, IFN $\alpha$  levels correlate with IFN-response gene expression levels of PB cells.<sup>37</sup> For SLE and SS, it has been demonstrated that immune complexes of autoantibodies and DNA- or RNA-containing autoantigens can induce type I IFN production by PB plasmacytoid DCs.<sup>28–30, 38–39</sup> This response depends on interaction with Fc $\gamma$ RIIa and Toll-like receptors.<sup>30–39</sup> Further studies are needed to determine whether the increased type I IFN-response genes in RA are the result of endogenous or infectious factors.

Besides a role for PB cells as producers of type I IFNs, cells at the site of inflammation may also be responsible for production. Cells with morphological and phenotypic characteristics of plasmacytoid DCs were shown to infiltrate skin lesions in SLE and actively produce type I IFN locally.<sup>40</sup> In patients with SS, numerous IFN $\alpha$ -producing cells were detected in the affected salivary gland biopsies.<sup>30</sup> In RA, IFN $\beta$  protein has been detected in the synovium.<sup>41</sup> Moreover, FLS were responsible for increased levels of IFN $\beta$  in the RA synovium.<sup>42</sup> The endogenous TLR3 ligand, dsRNA, derived from necrotic synovial fluid cells, has been shown to stimulate the production of IFN $\beta$  in RA FLS.<sup>43</sup>

In conclusion, we demonstrated that genomic profiling powers disease subclassification and has led to the identification of subgroups of patients, on the basis of differential expression of genes involved in non-specific immunity.



Supplementary table is given at <http://ard.bmj.com/supplemental>

#### Authors' affiliations

**T C T M van der Pouw Kraan\***, **L G M van Baarsen**, **F Rustenburg**, **J M Baggen**, **C L Verweij**, VU Medical Center, Department of Molecular Cell Biology & Immunology, Amsterdam, The Netherlands  
**C A Wijbrandts\***, **P P Tak**, Academic Medical Center, Division of Clinical Immunology and Rheumatology, University of Amsterdam, Amsterdam, The Netherlands  
**A E Voskuyl**, **B A C Dijkmans**, VU Medical Center, Department of Rheumatology, Amsterdam, The Netherlands  
**S M Ibrahim**, Section of Immunogenetics, University of Rostock, Rostock, Germany  
**M Ferro**, Stanford Functional Genomics Facility, Stanford University, Stanford, California, USA

\*These authors contributed equally to this work.

This work was supported in part by the EURO-RA Marie Curie Trainings network, the European Community's FP6 integrated program funding (AUTOCURE), the Innovation Oriented research Program (IOP) on Genomics and the Centre for Medical Systems Biology (a centre of excellence approved by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research). This publication reflects only the author's views. The European Community is not liable for any use that may be made of the information herein. These sponsors had no involvement in the study design, analysis or interpretation of the data and publications.

Competing interests: None.

#### REFERENCES

- 1 **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 Rheum* 1988;**31**:315–24.
- 2 **Zendman AJ**, van Venrooij WJ, Puijnt GJ. Use and significance of anti-CCP autoantibodies in rheumatoid arthritis. *Rheumatology (Oxford)* 2006;**45**:20–5.

- 3 Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, Fries JF, et al. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. *Arthritis Rheum* 2005;**52**:2645–55.
- 4 Tak PP, Smeets TJ, Doha MR, Kluin PM, Meijers KA, Brand R, et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. *Arthritis Rheum* 1997;**40**:217–25.
- 5 Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM, et al. Lymphoid neogenesis in rheumatoid synovitis. *J Immunol* 2001;**167**:1072–80.
- 6 van der Pouw Kraan TC, Van Gaalen FA, Huizinga TW, Pieterman E, Breedveld FC, Verweij CL. Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair. *Genes Immun* 2003;**4**:187–96.
- 7 van der Pouw Kraan TC, Van Gaalen FA, Kasperkovitz PV, Verbeet NL, Smeets TJ, Kraan MC, et al. Rheumatoid arthritis is a heterogeneous disease: evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. *Arthritis Rheum* 2003;**48**:2132–45.
- 8 Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR, et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000;**343**:1594–602.
- 9 Kremer JM, Westhovens R, Leon M, Di Giorgio E, Altan R, Steinfeld S, et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N Engl J Med* 2003;**349**:1907–15.
- 10 Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med* 2004;**350**:2572–81.
- 11 Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol* 1993;**32**:903–7.
- 12 Deighton CM, Wentzel J, Cavanagh G, Roberts DF, Walker DJ. Contribution of inherited factors to rheumatoid arthritis. *Ann Rheum Dis* 1992;**51**:182–5.
- 13 van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T, et al. A subtype of multiple sclerosis defined by an activated immune defense program. *Genes Immun* 2006;**7**:522–31.
- 14 Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
- 15 Ball CA, Awad IA, Demeter J, Gollub J, Hebert JM, Hernandez-Boussard T, et al. The Stanford Microarray Database accommodates additional microarray platforms and data formats. *Nucleic Acids Res* 2005;**33**:D580–2.
- 16 Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, O'Connell JX, et al. Molecular characterisation of soft tissue tumours: a gene expression study. *Lancet* 2002;**359**:1301–7.
- 17 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;**98**:5116–21.
- 18 Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;**95**:14863–8.
- 19 Mi H, Lazareva-Ulitsky B, Loo R, Kejarivala A, Vandergriff J, Rabkin S, et al. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* 2005;**33**:D284–8.
- 20 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;**102**:15545–50.
- 21 Der SD, Zhou A, Williams BR, Silverman RH. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc Natl Acad Sci USA* 1998;**95**:15623–8.
- 22 Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci USA* 2003;**100**:2610–5.
- 23 Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J, et al. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. *Ann Rheum Dis* 2004;**63**:1387–92.
- 24 Batliwalla FM, Baechler EC, Xiao X, Li W, Balasubramanian S, Khalili H, et al. Peripheral blood gene expression profiling in rheumatoid arthritis. *Genes Immun* 2005;**6**:388–97.
- 25 Bovin LF, Rieneck K, Workman C, Nielsen H, Sorensen SF, Skjodt H, et al. Blood cell gene expression profiling in rheumatoid arthritis. Discriminative genes and effect of rheumatoid factor. *Immunol Lett* 2004;**93**:217–26.
- 26 Wouters D, Wiessenberg HD, Hart M, Bruins P, Voskuyl A, Doha MR, et al. Complexes between C1q and C3 or C4: novel and specific markers for classical complement pathway activation. *J Immunol Methods* 2005;**298**:35–45.
- 27 Solomon DH, Goodson NJ, Katz JN, Weinblatt ME, Avorn J, Setoguchi S, et al. Patterns of cardiovascular risk in rheumatoid arthritis. *Ann Rheum Dis* 2006;**65**:1608–12.
- 28 Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. *Arthritis Rheum* 2005;**52**:1491–503.
- 29 Tan FK, Zhou X, Mayes MD, Gourh P, Guo X, Marcum C, et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. *Rheumatology (Oxford)* 2006;**45**:694–702.
- 30 Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML, et al. Activation of the type I interferon system in primary Sjogren's syndrome: a possible etiopathogenic mechanism. *Arthritis Rheum* 2005;**52**:1185–95.
- 31 Huang X, Yuang J, Goddard A, Foulis A, James RF, Lernmark A, et al. Interferon expression in the pancreases of patients with type 1 diabetes. *Diabetes* 1995;**44**:658–64.
- 32 Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. *Ann Neurol* 2005;**57**:664–78.
- 33 Biron CA. Interferons alpha and beta as immune regulators—a new look. *Immunity* 2001;**14**:661–4.
- 34 Tak PP. IFN-beta in rheumatoid arthritis. *Front Biosci* 2004;**9**:3242–7.
- 35 Treschow AP, Teige I, Nandakumar KS, Holmdahl R, Issazadeh-Navikas S. Stromal cells and osteoclasts are responsible for exacerbated collagen-induced arthritis in interferon-beta-deficient mice. *Arthritis Rheum* 2005;**52**:3739–48.
- 36 van Holten J, Pavelka K, Vencovsky J, Stahl H, Rozman B, Genovese M, et al. A multicentre, randomised, double blind, placebo controlled phase II study of subcutaneous interferon beta-1a in the treatment of patients with active rheumatoid arthritis. *Ann Rheum Dis* 2005;**64**:64–9.
- 37 Hua J, Kirou K, Lee C, Crow MK. Functional assay of type I interferon in systemic lupus erythematosus plasma and association with anti-RNA binding protein autoantibodies. *Arthritis Rheum* 2006;**54**:1906–16.
- 38 Lovgren T, Eloranta ML, Bave U, Alm GV, Ronnblom L. Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG. *Arthritis Rheum* 2004;**50**:1861–72.
- 39 Lovgren T, Eloranta ML, Kastner B, Wahren-Herlenius M, Alm GV, Ronnblom L. Induction of interferon-alpha by immune complexes or liposomes containing systemic lupus erythematosus autoantigen- and Sjogren's syndrome autoantigen-associated RNA. *Arthritis Rheum* 2006;**54**:1917–27.
- 40 Farkas L, Beiske K, Lund-Johansen F, Brandtzaeg P, Jahnsen FL. Plasmacytoid dendritic cells (natural interferon-alpha/beta-producing cells) accumulate in cutaneous lupus erythematosus lesions. *Am J Pathol* 2001;**159**:237–43.
- 41 van Holten J, Smeets TJ, Blankert P, Tak PP. Expression of interferon beta in synovial tissue from patients with rheumatoid arthritis: comparison with patients with osteoarthritis and reactive arthritis. *Ann Rheum Dis* 2005;**64**:1780–2.
- 42 Pilling D, Akbar AN, Girdlestone J, Orteu CH, Borthwick NJ, Amft N, et al. Interferon-beta mediates stromal cell rescue of T cells from apoptosis. *Eur J Immunol* 1999;**29**:1041–50.
- 43 Brentano F, Schorr O, Gay RE, Gay S, Kyburz D. RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via toll-like receptor 3. *Arthritis Rheum* 2005;**52**:2656–65.