Cutting Edge: Molecular Portrait of Human Autoimmune Disease

Kevin Maas, Sanny Chan, Joel Parker, Angela Slater, Jason Moore, Nancy Olsen and Thomas M. Aune

*J Immunol* 2002; 169:5-9; doi: 10.4049/jimmunol.169.1.5

http://www.jimmunol.org/content/169/1/5

---

**References**

This article cites 22 articles, 6 of which you can access for free at:

http://www.jimmunol.org/content/169/1/5.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:

http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:

http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:

http://jimmunol.org/alerts
Autoimmune diseases occur in 3–5% of the population (1, 2). The site of attack is organ- or tissue-specific: pancreatic β cells in type 1 diabetes (3), myelin basic protein in multiple sclerosis (MS)4 (4), thyroglobulin or thyroid peroxidase in thyroiditis (5), or more systemic: the synovial lining, lung, or heart in rheumatoid arthritis (RA) (6), or skin, kidney, or heart in systemic lupus erythematosus (SLE) (7). Despite these differences, clinically distinct autoimmune diseases exhibit common elements. First, most autoimmune diseases are thought to arise from alterations in the immune system. Second, coassociation of different autoimmune diseases is often found in families or individuals. Third, analysis of genome-wide linkage results demonstrates that multiple autoimmune diseases share common susceptibility loci. Taken together, these findings suggest that common groups of genes may contribute to development of clinically distinct forms of autoimmune disease (8–11).

cDNA microarray technology represents a powerful tool to investigate differences in gene expression profiles. We wanted to compare these profiles in a common cell population from individuals with clinically distinct forms of autoimmune disease. Given the ease of obtaining blood, we first tested whether PBMC represented such a source. To do so, we performed gene expression profiling of PBMC in control individuals following immunization with the influenza vaccine. The results show that the normal immune response is a highly dynamic process with coordinate changes in expression of genes that encode proteins with common functions. In contrast, genes with a distinct expression pattern in autoimmunity are conserved among the different autoimmune diseases but are not “immune response” genes.

Materials and Methods

Patient populations

Nine control subjects (27–58 years of age) were studied before and after influenza vaccination. Patients with RA (n = 20; 46–68 years of age), SLE (n = 24; 22–73 years), type I diabetes (n = 5; 20–46 years), and MS (n = 4; 37–54 years) were also enrolled in the study. A clinical diagnosis of each autoimmune disorder was the sole criterion for inclusion. Unaffected family members were also included in the study (n = 4, 33–54 years); three were parents of individuals with SLE and one was the child of an individual with RA. The ratio of females to males in the test groups was 3:1.

Sample preparation

PBMC were isolated from heparinized blood by centrifugation on a ficoll-hypaque (Sigma-Aldrich, St. Louis, MO) gradient. Leukocyte distribution in PBMC was determined by flow cytometry. Total RNA was isolated with Tri-Reagent (Molecular Research Center, Cincinnati, OH), processed, and 5 μg were hybridized to Research Genetics GF-211 membranes (Huntsville, AL). Filters were exposed to imaging screens for 24 h and screens were scanned by a phosphor imager (Molecular Dynamics, Piscataway, NJ). Data were normalized to yield an average intensity of 1.0 for each clone (4329) represented on the microarray. Reproducibility of the method was established by performing replicate hybridizations to separate microarrays. Linear regression analysis demonstrated that separate hybridizations yielded R2 values ranging from 0.87 to 0.96. Different exposure lengths of identical filters also produced high R2 values (0.99).

Data analysis

Eisen’s Cluster and Treeview software (Stanford University, Palo Alto, CA) (12) were used to compare similarities among individual samples. Data sets were analyzed using hierarchical, K-means, and self-organizing map algorithms (13). Research Genetics Pathways 3.0 program was used to identify differentially expressed genes in the immune and autoimmune disease classes. Expression levels of genes that did not change significantly (99% confidence, Chen test) over any of the conditions were removed from
FIGURE 1. Cluster analysis of the pre- and postimmune data set. A. An unsupervised self-organizing map was used to compare individuals before (control) or after immunization (IMM, days 6–9 postimmunization). Upper panel, profiles from analysis of all genes; lower panel, profiles after removing invariant genes. Individuals (designated 11–18) are connected by brackets. B, K-means analysis identified three distinct clusters based upon similarities in gene expression. Data are presented as the natural logarithm of the ratio of the experimental group indicated on the x-axis to the control group. Individual lines represent expression ratios of individual genes over the time course.

Results and Discussion
First, we wanted to test the hypothesis that the mononuclear cell population represented a suitable source to measure alterations in gene expression. To do this, we measured changes in gene expression in PBMC from healthy control subjects (n = 9) before and after immunization with influenza vaccine. In this case, it was most likely that we were viewing a secondary immune response because all adults had prior exposure to many influenza Ags. Samples were collected from subjects at three time points: 3, 6–9, and 19–21 days after immunization. A self-organizing map algorithm was used to compare the preimmune to the immune group. This method segregated individuals based upon identity rather than immune status, as demonstrated by the relative proximity of individual samples (Fig. 1A, upper panel). Thus, total gene expression patterns remained relatively unchanged after immunization. To focus on distinctions that arose from the most differentially expressed genes, we filtered out genes for which expression levels did not vary by $\geq 3$ SD from their respective means. After filtering, expression profiles were segregated primarily by pre- and postimmune status (Fig. 1A, lower panel). This suggested that uniform changes in expression levels of a smaller subset of genes distinguished pre- and postimmunization groups. To identify these genes, we used K-means clustering to group genes on the basis of similarity in expression patterns.

We found three distinct clusters associated with the normal immune response (Fig. 1B). The first cluster consisted of 304 genes that were overexpressed 3 days after immunization. This cluster mainly contained genes that encode proteins involved in key signal transduction pathways (e.g., protein kinase C, phospholipase C, 1,2-diacylglycerol kinase, mitogen-activated protein kinase, STATs, and STAT inhibitors, AP-1 transcription factors, interferon regulatory factors, and proteins required for proliferation). Genes in this cluster exhibited an increase in expression from 3- to 21-fold compared with the control group.

The second cluster of 88 late (19–21 days) response genes represented a shift away from signaling and proliferation pathways toward increased functional activity. Among the late immune response gene cluster, chemokines (SCYA3, SCYA13, SCYA14), complement components (C1S), IFN-inducible proteins (IFI35), and leukocyte homing/adhesion (ICAM2) genes were overexpressed. Receptors for serotonin, glutamate, estrogen, and retinoic acid were also overexpressed. Increases in expression levels of this group of genes varied from 2- to 11-fold.

The final immune response cluster contained 78 genes that exhibited reduced expression levels over the entire time course. Over 15% of these genes encode ribosomal proteins. This represents a decrease in the expression of one-third of all ribosomal protein-encoding genes present on the microarrays. Coordinate changes in ribosomal protein gene expression have been linked to differentiation in eukaryotic cells (15) and the observed changes may reflect differentiation of lymphocytes to an effector state in response to immunization. Taken together, these data illustrate dynamic, coordinate changes in mRNA expression that accompany the immune response, in vivo. First, genes were induced that are required for signal transduction and cell proliferation, two key elements of the early immune response. Later, we observed a shift away from these genes to other classes that are necessary to undertake the immune functions of lymphocytes.

Given the above results, we wanted to compare gene expression profiles of the normal immune response to those of autoimmune disease. We obtained samples from patients diagnosed with one of four common autoimmune disorders: RA, MS, type I diabetes, and SLE. The relatedness of global gene expression profiles associated with autoimmune disease was examined relative to the normal immune response using a hierarchical clustering algorithm (Fig. 2A). Other clustering algorithms yielded similar results (data not shown). Comparison between the RA/SLE class and the normal immune response class yielded four major branches from the clustering analysis. One major branch contained all normal immune samples and none of the autoimmune samples. The autoimmune samples segregated into the other three major branches. This analysis revealed that some of the RA samples (e.g., RA2 and RA5, or RA1, RA6, and RA4) and some of the SLE samples (e.g., SLE2, SLE3, and SLE4, or SLE6, SLE8, and SLE9) were highly related. However, unlike distinctions between the RA/SLE and the normal immune response samples, it was not possible to segregate the majority of RA samples from the majority of SLE samples. This argues that RA and SLE may represent a common autoimmune class that is distinct from the immune class. Similar results were obtained from clustering of normal immune response samples with MS/type I diabetes samples. Again, there was good segregation of the normal immune response group from the MS/type I diabetes group, but MS and type I diabetes profiles did not segregate from each other. This inability to segregate within autoimmune class was retained even when invariant genes were removed from the data set (data not shown). One explanation for the inability of the

* Complete data tables are available upon request.
clustering algorithm to distinguish among the phenotypically distinct autoimmune diseases is that the gene expression profiles are relatively similar.

We further analyzed the data set to identify genes that were most differentially expressed in autoimmune diseases relative to the normal immune response. Nonautoimmune groups were segregated into control (no treatment) and immune (6–9 days after immunization). Individual samples from the autoimmune groups were segregated based upon disease type and compared with the immune response gene profiles. Gene expression differences among different groups were plotted as the natural logarithm of the ratio between experimental condition and control group.

Two clusters of differentially expressed genes distinguished between 1) patients with autoimmune disease and 2) control and immune individuals (Fig. 2B). The first major cluster consisted of 95 genes that were overexpressed in all four autoimmune diseases (type I diabetes, MS, RA, and SLE). The genes in this overexpressed autoimmune cluster were relatively heterogeneous, representing several distinct functional categories: receptors (CSF3R, HLA-DMB, HLALS, TGFBR2, and BMPR2), inflammatory mediators (MSTP9, BDNF, CES1, ELA3, and CYR61), signaling/second messenger molecules (FASTK, DGKA, and DGKD), and autoantigens (GARS and GAD2). The second major cluster contained 117 genes that were strongly underexpressed in all autoimmune groups. Levels of expression of these genes did not change in the immune response group. Many of the down-regulated genes play key roles in apoptosis (TRADD, TRAP1, TRIP, TRAF2, CASP6, CASP8, TP53, and SIVA) and ubiquitin/proteasome function (UBE2M, UBE2G2, and POH1). Inhibitors of various cellular functions were also widely represented in this cluster. These include direct inhibitors of cell cycle progression (CDKN1B, CDKN2A, and BRCA1), as well as inducers of cell differentiation (LIF and CD24). Certain enzyme inhibitors (APOC3 and KAL1) were also found in this class.

A striking result of the K-means clustering was that we were unable to identify clusters of genes that overlapped between the immune and autoimmune classes. This argues that the expression patterns that characterize the normal immune response bear little resemblance to those found in autoimmune disease. In addition, we were unable to identify clusters of genes that distinguished among the distinct autoimmune diseases. Rather, the results point to pervasive changes in gene expression that are relatively uniform among four distinct autoimmune diseases.

To examine this hypothesis in greater detail, we compared expression levels of single genes between preimmune controls and individuals with each of four autoimmune diseases. We selected 10 genes that exhibited the greatest level of over- and underexpression (Fig. 3) at the population level and were highly consistent in each individual with autoimmune disease. Overexpressed genes in the autoimmune population showed greater individual variation in
expression (Fig. 3A). In this group, no individual gene was overexpressed in all autoimmune individuals compared with all control individuals. However, each of these genes was significantly overexpressed in the autoimmune population (p < 0.05). In contrast, expression levels of underexpressed genes (Fig. 3B) were lower in all autoimmune individuals than in all control individuals.

Differences in gene expression between the control and the autoimmune populations may be attributed to alterations in distribution or activation status of cells that make up the PBMC. We performed two analyses to test this possibility. First, we analyzed PBMC preparations for frequency of CD3 (T cells), CD14 (monocytes), CD19 (B cells), and leukocyte alkaline phosphatase (neutrophils) by flow cytometry. All PBMC preparations from both subject groups contained 75–80% T cells, ~10% monocytes, ~5% B cells, and <1% neutrophils. Second, we determined whether expression levels of genes that are either restricted to a given subpopulation or reflect activation status were differentially expressed in the control compared with the autoimmune population (Table I). Expression levels of these genes varied by less than 2-fold between the control and autoimmune groups and this difference did not achieve statistical significance. Taken together, these data argue that alterations in the composition or activation status of PBMC did not account for the observed differences in gene expression between the control and autoimmune populations.

Because genes underexpressed in autoimmune disease exhibited a more homogeneous pattern of expression than overexpressed genes (Fig. 3), we wanted to determine whether we could use this pattern to classify individuals with autoimmune disease and predict whether new samples were derived from autoimmune or control individuals. First, we selected 35 genes that were most underexpressed in the SLE population compared with the control population with the greatest statistical significance. Second, to give each gene equal weight in the determination, we summed the average of the control population and the SLE population for each gene and divided by 2 ((control + SLE)/2). Third, we inspected expression levels of each of the 35 genes in each sample and assigned them a value of 0 if they were less than the average and 1

### Table I. Expression levels of genes encoding proteins that distinguish among lymphocyte subsets or activation state

<table>
<thead>
<tr>
<th>Population Expression Level</th>
<th>Control</th>
<th>SLE</th>
<th>RA</th>
<th>Insulin-Dependent Diabetes Mellitus</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>T cell Ags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD38</td>
<td>0.7 ± 0.2</td>
<td>0.6 ± 0.4</td>
<td>0.5 ± 0.2</td>
<td>0.5 ± 0.2</td>
<td>0.4 ± 0.2</td>
</tr>
<tr>
<td>CD3y</td>
<td>0.5 ± 0.1</td>
<td>0.6 ± 0.9</td>
<td>0.4 ± 0.1</td>
<td>0.3 ± 0.1</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>CD88 (Tc)</td>
<td>0.8 ± 0.3</td>
<td>0.8 ± 0.2</td>
<td>0.6 ± 0.2</td>
<td>0.5 ± 0.2</td>
<td>0.5 ± 0.2</td>
</tr>
<tr>
<td>CD44 (memory)</td>
<td>0.5 ± 0.1</td>
<td>0.8 ± 0.5</td>
<td>0.7 ± 0.4</td>
<td>0.8 ± 0.5</td>
<td>0.7 ± 0.4</td>
</tr>
<tr>
<td>CD69 (activation)</td>
<td>0.5 ± 0.2</td>
<td>0.7 ± 0.3</td>
<td>0.6 ± 0.2</td>
<td>0.8 ± 0.3</td>
<td>0.7 ± 0.4</td>
</tr>
<tr>
<td>CD62 (L-selectin)</td>
<td>1.3 ± 0.6</td>
<td>1.4 ± 0.9</td>
<td>1.8 ± 1.1</td>
<td>1.7 ± 1.1</td>
<td>1.9 ± 1.1</td>
</tr>
<tr>
<td>CD122 (IL-2Rβ)</td>
<td>0.4 ± 0.1</td>
<td>0.4 ± 0.2</td>
<td>0.5 ± 0.2</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>B cell Ags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD79a</td>
<td>0.6 ± 0.3</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
</tr>
<tr>
<td>CD79b</td>
<td>0.5 ± 0.2</td>
<td>0.6 ± 0.3</td>
<td>0.8 ± 0.7</td>
<td>0.8 ± 0.4</td>
<td>0.7 ± 0.3</td>
</tr>
<tr>
<td>CD72</td>
<td>0.4 ± 0.1</td>
<td>0.4 ± 0.3</td>
<td>0.3 ± 0.2</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>CD22</td>
<td>0.3 ± 0.1</td>
<td>0.4 ± 0.3</td>
<td>0.4 ± 0.4</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>Monocyte Ags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD14</td>
<td>0.5 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.2</td>
<td>0.3 ± 0.2</td>
</tr>
<tr>
<td>CD163</td>
<td>0.3 ± 0.1</td>
<td>0.4 ± 0.2</td>
<td>0.4 ± 0.2</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.2</td>
</tr>
<tr>
<td>CD32 (B/mθ)</td>
<td>0.3 ± 0.1</td>
<td>0.5 ± 0.4</td>
<td>0.5 ± 0.3</td>
<td>0.3 ± 0.1</td>
<td>0.4 ± 0.2</td>
</tr>
<tr>
<td>Activation-induced Ags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD54 (ICAM-1)</td>
<td>4.4 ± 1.8</td>
<td>3.1 ± 2.1</td>
<td>4.3 ± 0.7</td>
<td>4.3 ± 2.2</td>
<td>3.9 ± 1.0</td>
</tr>
<tr>
<td>CD38</td>
<td>0.4 ± 0.3</td>
<td>0.3 ± 0.2</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>CD71</td>
<td>0.2 ± 0.1</td>
<td>0.2 ± 0.1</td>
<td>0.2 ± 0.1</td>
<td>0.2 ± 0.1</td>
<td>0.2 ± 0.1</td>
</tr>
</tbody>
</table>

*Average expression level of the indicated genes within the different populations.

**FIGURE 4.** Classification and prediction of autoimmune disease. The score (y-axis) is shown for each individual sample analyzed from the different populations (x-axis). Values of p are depicted in the legend. The 35 genes used to derive this score were (gene symbols or accession no.): TGM2, SSP29, TAF21, LLGQ2, TNFAIP2, SPP1, BPH1, TP53, DIPA, ASL, DFR2,58601922, MAN1A1, R09503, LOC51643, BMPS, ORC1L, W04674, R04175, CDH1, SUSD, EPB72, CDKN1B, CASP6, TXK, MYO1B, LIF, HSJ2, BRCA1, GUCY1B3, AP352, N68565, SC55, UBE2G2, SLC16A4, and MMP17.
if they were greater than the average. In this analysis, the maximum possible score was 35 and the minimum possible score was 0 (Fig. 4). The range of scores for control individuals was 18–35. Most achieved a score of 35 (8 of 11). The range of scores for the normal immune samples was 26–35. In contrast, the range of scores in individuals with autoimmune disease was 0–5 (SLE), 0–6 (RA), 0–1 (type I diabetes), and 0 (MS) \((p < 0.000001)\). Next, we tested a new set of SLE and RA patients that were not included in the initial data analysis to determine whether these individuals would obtain a similar score. The range of scores for this second autoimmune group was 0–5 (SLE) and 0–6 (RA).

One question raised by these results is whether differences in gene expression in individuals with autoimmune disease reflect chronic illness, autoimmune disease, or mode of therapy. Alternatively, differences in gene expression may result from a genetic trait or other unrecognized factors. To examine this question in greater detail, we determined gene expression profiles of four unaffected first-degree relatives of autoimmune patients. Each of these individuals obtained a score of 0 in our test and exhibited a gene expression profile that was very similar to the gene expression profile of an individual with autoimmune disease. Although this is a relatively small sample, these data support the possibility that the gene expression profile outlined in this study may reflect a genetic trait rather than the disease process.

In designing this study, we set out to test the hypothesis that individuals with autoimmune disease exhibited uniform changes in gene expression in PBMC. For this reason, we studied patients with variable disease activity, duration of disease, modes of therapy, and type of autoimmune disease. In our population, the most common therapy for RA is methotrexate, for SLE it is prednisone, and for type of autoimmune disease. In our population, the most common therapy for RA is methotrexate, for SLE it is prednisone, and for MS it is IFN-β. None of these patients were receiving cyclophosphamide.

The fact that the observed changes in gene expression are independent of these parameters and are also observed in first-degree unaffected family members argues that these changes are genetic in origin rather than a reflection of the disease process or therapy. Analysis of patients on prednisone indicates that this treatment partially corrects the expression levels of certain genes (data not shown), but these genes are not included in the set that comprises the common gene expression profile found in autoimmune disease. This suggests that other clusters of genes, as yet unidentified, are involved in the disease process.

Individual genes represented within the autoimmune expression profile illustrate the potential impact of these variations upon the immune response. First, a large number of genes that encode proteins involved in multiple apoptosis pathways are underexpressed. Defects in apoptosis have been clearly demonstrated in animal models of autoimmune disease and in human autoimmune disease (16–19). However, the pervasive nature of defects in the expression of a large number of genes involved in apoptosis pathways was unexpected. Second, inhibitors of a variety of functions, ranging from cell cycle control, proteasome function, enzyme activities, and cell differentiation are also underexpressed. Third, genes required for induction of tolerance are also underexpressed.

Defects in expression of these genes may increase the likelihood that lymphocytes avoid the normal processes used by the immune system to eliminate unwanted lymphocytes or to down-regulate an immune response (20–22). Normally, autoreactive lymphocytes that survive the process of thymic selection and circulate in the periphery can respond to self-Ag if presented by the correct MHC. This process is held in check by the combined actions of apoptosis and anergy and results in a state of peripheral tolerance. However, if patients carry this autoimmune gene expression signature, signaling pathways essential for the maintenance of tolerance may not function properly. This may permit lymphocytes to escape tolerance and adopt a prosurvival agenda that increases the likelihood of autoimmune diseases.

Acknowledgments

We thank S. Sriman, A. Powers, H. Fuchs, and J. Lipp for providing valuable patient samples.

References


