Modular Analysis of Peripheral Blood Gene Expression in Rheumatoid Arthritis Captures Reproducible Gene Expression Changes in Tumor Necrosis Factor Responders

M. Oswald  
*Northwell Health*

M. E. Curran  
S. L. Lamberth  
R. M. Townsend  
J. D. Hamilton

*See next page for additional authors*

Follow this and additional works at: [https://academicworks.medicine.hofstra.edu/articles](https://academicworks.medicine.hofstra.edu/articles)

Part of the [Medical Molecular Biology Commons](https://academicworks.medicine.hofstra.edu/articles)

**Recommended Citation**


This Article is brought to you for free and open access by Donald and Barbara Zucker School of Medicine Academic Works. It has been accepted for inclusion in Journal Articles by an authorized administrator of Donald and Barbara Zucker School of Medicine Academic Works. For more information, please contact academicworks@hofstra.edu.
Modular Analysis of Peripheral Blood Gene Expression in Rheumatoid Arthritis Captures Reproducible Gene Expression Changes in TNF Responders

Michaela Oswald¹, Mark Curran², Sarah Lamberth², Robert Townsend³, Jennifer D. Hamilton⁴, David N. Chernoff⁵, John Carulli⁶, Michael Townsend⁷, Michael Weinblatt⁸, Marlena Kern¹, Cassandra Pond¹, Annette Lee¹, and Peter K. Gregersen¹

¹Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, Manhasset, New York
²Systems Pharmacology and Biomarkers, Janssen Research and Development, Springhouse, Pennsylvania
³Clinical Biomarkers Group, Bristol-Myers Squibb, Princeton, New Jersey
⁴Regeneron Pharmaceuticals Inc, Tarrytown, New York
⁵Crescendo Bioscience, South San Francisco, California
⁶Biogen IDEC, Cambridge, Mass
⁷Genentech, South San Francisco, California
⁸Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Mass.

Abstract

Objective—To establish whether the analysis of whole blood gene expression can be useful in predicting or monitoring response to anti-TNF therapy in RA.

Methods—Whole blood RNA (PAXgene) was obtained at baseline and 14 weeks on three independent cohorts with a combined total of 250 patients with rheumatoid arthritis beginning anti-TNF therapy. We employed an approach to gene expression analysis that is based on gene expression “modules”.

Results—Good and Moderate Responders by EULAR criteria exhibited highly significant and consistent changes in multiple gene expression modules using a hyper geometric analysis after 14 weeks of therapy. Strikingly, non responders exhibited very little change in any modules, despite exposure to TNF blockade. These patterns of change were highly consistent across all three
cohorts, indicating that immunological changes after TNF treatment are specific to the combination of both drug exposure and responder status. In contrast, modular patterns of gene expression did not exhibit consistent differences between responders and non-responders at baseline in the three cohorts.

**Conclusions**—These data provide evidence that using gene expression modules related to inflammatory disease may provide a valuable method for objective monitoring of the response of RA patients who are treated with TNF inhibitors.

**INTRODUCTION**

The development of TNF inhibitors for the treatment of rheumatoid arthritis and other inflammatory disorders has been a seminal advance for the field of rheumatology(1). However, therapeutic approaches to RA remain a substantial clinical challenge, since approximately 30% of RA patients do not respond to TNF blockade, and many others obtain only partial reduction in joint inflammation(2). In addition, the determination of the clinical response to drug therapy is an inexact science, largely relying on both subjective perceptions of patients as well as physician assessments of disease activity which exhibit considerable variability. The integration of laboratory measures of inflammatory activity such as the CRP or the ESR is standard, but non-specific, and these measurements often do not follow the clinical parameters(3). The recent development of multi-parameter methods to determine disease activity may be helpful (4-5), but these biomarkers have not yet been widely used in formal studies of clinical drug response or demonstrated to benefit patient management.

Thus, there is a substantial unmet need for improvement in both prediction of drug response in order to individualize therapy, as well as in the development of quantitative laboratory methods to determine clinical activity of disease, which in turn will aid in the development of predictive biomarkers and enable more precise treatment concepts such as “treat to target” (6).

Expression microarrays have been utilized by a number of groups in order to develop biomarkers for drug response(7-15). However, the early applications of this technology have been compromised by a number of issues(16). First, small sample sizes for initial feature set definition lead to over-fitting of discriminative models and is further compromised by the semi-quantitative nature of disease activity phenotypes. Critically, there has been rather limited replication of specific predictive gene expression patterns between studies conducted to date. Although transcriptomics is a maturing technology, there are various approaches to expression analysis and sample collection methods are often not standardized. Many studies examine gene expression in purified peripheral blood mononuclear cells. This can lead to experimental noise due extracorporeal changes in gene expression, especially if RNA purification is not carried out immediately. More recently, the use of PaxGene or other technologies which stabilize whole blood RNA in the collection tube has helped to reduce this source of variability. Of course, this does not permit the analysis of expression in specific separated cell subsets of interest. Overall, none of these studies have provided convincing, replicated results using a standardized protocol for sample collection, microarray platform, and method of analysis(16).
In the current study we have attempted to address these issues by examining peripheral blood gene expression using a standard approach to blood sample collection (PaxGene) and the use of standard Illumina array platforms in three independent, prospectively collected cohorts of RA patients beginning therapy with TNF inhibitors. In addition, for our primary approach to analysis, we have taken advantage of an empirically derived set of “modules” of gene expression developed by Chaussabel and colleagues(17-18). The results reveal that patients who have a clinical response to TNF inhibition exhibit a common pattern of changes in these gene expression modules after 14 weeks of therapy, whereas clinical non-responders do not have major changes in modular gene expression. In contrast, we were not able to identify modular gene expression changes at baseline that identify patients who will, or will not, respond to TNF blockade. These results suggest to us that global gene expression patterns in peripheral blood can be a useful adjunct for determining whether clinical response has occurred. However, we expect that successful prediction of response to therapy is likely to require a focus on particular cell subsets along with the integration of genetic factors and additional biomarker data.

PATIENTS AND METHODS

Patient Cohorts

Three prospectively collected cohorts of patients with rheumatoid arthritis were studied, as summarized in Table 1. The ABCoN cohort was enrolled as part of the Autoimmune Biomarkers Collaborative Network, a NIAMS supported contract to develop new approaches to biomarkers for rheumatoid arthritis and lupus. Fifty (50) ABCoN patients had complete data available for study. Patients were treated with Enbrel (n=22), Remicade (n=19), Humira (n=9). At 14 weeks, the following distribution of EULAR responses were observed: Good Responder=14, Moderate Responder=21, Non-responder=15. A second patient cohort was taken from GO-FURTHER, a phase III clinical trial of Golimumab (19). All patients were started on 2 mg/kg IV and clinical assessment of response was determined at 14 weeks and 72 patients had gene expression data available for analysis. Of these, EULAR responses were as follows: Good Responder=29, Moderate Responder=37, Non-responder=6. A third patient cohort of 120 RA patients was taken from a multicenter cohort, BATTER-UP (BIOMARKERS OF ANTI- TNF-α THERAPY EFFICACY IN RHEUMATOID ARTHRITIS TO DEFINE UNRESPONSIVE PATIENTS). These patients were treated with a variety of anti-TNF agents including Enbrel (n=32), Remicade (n=23), Humira (n=41), CIMZIA (n=14), SIMPONI (n=9). EULAR response rates for the BATTER-UP cohort were as follows: Good Responder=42, Moderate Responder=39, Non-responder=39.

Microarray analysis

All subjects had PaxGene tubes drawn at baseline before starting therapy, and again at 14 weeks. RNA was isolated using Qiagen's QIAcube (Qiagen, Venlo, The Netherlands), following the manufacturer's protocol for Paxgene Blood RNA part 1, automated protocol. Extracted samples were eluted in 80ul of elution buffer (BR5), and subsequently run on Agilent's 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), for integrity, using the RNA 6000 Nano Chip. Samples with RNA Integrity Numbers (RINs) greater than 6.5 were
diluted to 30ng/ul in a total 11ul of RNAse free water. Samples were amplified using Life Technology’s Illumina RNA Total Prep Amplification Kit (Life Technologies, Carlsbad, CA). 750ng of cRNA was resuspended in 5ul of RNAse-free water for analysis the on the Human HT-12v4 chip and 1.2ug was resuspended in 10ul for analysis on the WG6v3 Bead Chip (Illumina, Inc., San Diego, CA). All samples were processed according to the manufacturer’s instructions. The Illumina WG6v3 was utilized for samples in the ABCoN cohort, while the HT12v4 was used for samples in the GO-FURTHER and BATTER-UP cohorts.

Raw data were exported from GenomeStudio and further analyzed with the R programming language. All 3 data sets were background corrected using the R/Bioconductor package ‘lumi’ (20). Data were further transformed using a vst (variance stabilization transformation) and quantile normalized. The moderated t-statistic from the R/Bioconductor package ‘limma’ (21) was used for differential expression analysis. We analyzed modules of co-expressed genes that have previously demonstrated coordinated regulation of expression across a set of diseases (whole blood) as described by Chaussabel et al (17). A total of 28 modules were analyzed, each containing 22-325 gene probes. Because modules were originally defined using Affymetrix probesets, these probes were mapped to Entrez Gene ID which in turn were mapped to Illumina probes as shown in Figure 1. Group comparisons (e.g. modular expression at baseline vs 14 weeks; responders vs. non responders) were carried out using a moderated t statistic comparing groups and categorizing all individual probes as up or down regulated using a significance cutoff p<0.05. For each module, we compiled the list of all significantly changed probes (up or down) at the cutoff significance P<0.05. We then determined what fraction of the probes are significantly up or down regulated within a given module. When comparing groups (e.g. responders vs. non responders) the significance of the observed difference in distribution of up/down regulated probes in a module was calculated using a hypergeometric test (equivalent to a 2×2 table analysis by Fischer’s exact test).

RESULTS

Consistent changes in modular gene expression are observed after 14 weeks of anti-TNF therapy in clinical responders

As originally reported by Chaussabel, 28 gene expression modules can be observed in whole blood across a variety of inflammatory and infectious diseases (17). Where possible, these modules were labeled after the fact according to the identity of the gene transcripts present in these modules (e.g. B cells, myeloid lineage, T cells, Interferon Inducible etc.). However, many modules could not be easily categorized, and were designated as “Undetermined”. We have utilized these designations as well as the original module numbers (M1.1-1.8; M2.1-2.10; M3.1-3.9) in our discussion here.

We initiated analysis by investigating whether changes in modular gene expression occurred after TNF blockade, and whether these changes differ among responders and non responders at the 14 week time point. The results for the three datasets are summarized separately for EULAR Responders (Good and Moderate Responders combined) and EULAR Non-responders in Figure 2. For each module, the direction of change in module expression is
shown for each dataset, with statistical significance indicated by the shading of the bars. The fraction of probes that are significantly changed (up or down) within each module is indicated by the height of the bars. Note that for the Responder groups, there is a consistent and statistically significant increase in modules containing transcripts related to plasma cells (M 1.1), B cells M1.3, MHC and ribosomal proteins (M 1.7), and T cells (M2.8), as well as a variety of “Undetermined” modules (M1.8, M3.4, M3.7, M3.8, M3.9). In addition, there is a consistent down modulation of transcripts in several myeloid lineage modules (M1.5, M2.6) and platelets (M2.1) as well as Inflammation 1 (M3.2) and Inflammation 2 (M3.3). These results stand in marked contrast to the lack of any consistent modular changes across the three datasets for nonresponders. Indeed, in general, non responders show a very small fraction of probes significantly changing when comparing 14 weeks to baseline within any of the 28 modules.

In order to provide a more detailed representation of these modular changes and their magnitude, we have plotted the fold changes for all probes achieving a significant difference (p<0.05, moderated t-test) within each module, for each dataset and responder group. This is shown in figure 3A, 3B and 3C for the ABCoN, GO-FURTHER and BATTERUP cohorts, respectively. Within the responders from the three datasets, there is striking consistency in the pattern of changes in modules. Significant changes (p<0.005, hypergeometric test) in probe distribution, up or down, are indicated in red and blue respectively. Indeed, some of these changes achieve remarkably compelling levels of statistical significance. For example, the p values associated with reduction in myeloid modules 1.5 and 2.6 after TNF blockade range from 10^{-20} to 10^{-87}, and reductions in the “Undetermined” modules 3.2 and 3.3 range from 10^{-12} to 10^{-85} across the various datasets. These dramatic changes are only seen in the Responder group, with minimal change in patients who are non responders.

**Modules do not exhibit consistent differences between responders and non-responders at baseline**

While substantial changes in modular gene expression are seen in responders after TNF blockade, the baseline pattern of module expression does not show consistent differences among the three cohorts (Figure 4). While some of the datasets suggest presence of meaningful differences between responders and non responders at baseline in one cohort (for example several “undetermined” modules are relatively less prominent in responders in the ABCoN cohort), these changes are not replicated in the other two datasets. This emphasizes the importance of having replication data sets to validate initial findings, and also demonstrates that an improved understanding of RA patient heterogeneity will be required for the development of an accurate predictor of drug response.

**DISCUSSION**

In this report we have demonstrated a consistent change in whole blood gene expression in clinically responding RA patients after exposure to TNF blocking agents for 14 weeks. These changes are not observed in subjects who do not exhibit a clinical response to treatment, using standard EULAR criteria. The data are based on a consideration of sets of genes that are expressed together in a modular fashion across different inflammatory
conditions, as first demonstrated by Chaussabel(17). These data provide compelling evidence that these gene expression changes are not simply a result of exposure to TNF blockade, but rather reflect a change in both the immunological and disease state in subjects who are responding to TNF therapy. Interestingly, we have recently had an opportunity to examine a small number of subjects who did not receive study drug (placebo group: 7 responders, 16 non-responders). In this case, there was no difference between responders and non-responders (data not shown), suggesting that these changes are specific to physiological effects of drug therapy in the setting of response. However, a larger study will be required to confirm these findings.

In contrast to the gene expression changes observed in responders over time, these modules do not predict response to future response to TNF blockade when these modular expression patterns are interrogated at baseline before starting therapy, as shown in Figure 4. The development of useful biomarkers for the management of rheumatoid arthritis has been extremely challenging. At present, there is no clinical parameter or laboratory test that can distinguish likely responders to anti-TNF blockade prior to initiation of therapy. Given the fact that at least 25% of patients will not respond to anti-TNF therapy, while at most one third will have a robust clinical response, there is clearly a need to develop such markers. Several small studies have proposed biomarkers to predict such response, but none have been validated for any biologic therapy(16). It is important to note that published studies suffer from reporting primary results and have not had multiple independent cohorts available for validation and replication of results. A primary component of our work here is replication of findings in multiple independent cohorts. In this context, our use of modular approaches to whole blood microarray analysis identified suggestive feature sets for prediction of response in individual cohorts but these marker sets failed to replicate or yield a convincing gene expression biomarker for future drug response. This emphasizes the biological complexity of the RA response phenotype. If the scientific community is to succeed in the goal of response prediction, it is likely that multiple parameters will need to be combined, including genetic and genomic data along with cellular and biochemical data that more precisely reflect the underlying TNF dependence of the current inflammatory state and inflammation drivers in a given patient.

Our current results are likely to have utility in the context of determining the presence of reduced disease activity in response to drug. This is an important issue, since the current measures, such as the ACR and EULAR criteria, are highly dependent on primarily subjective patient and physician based data points. The current results suggest that analysis of modular gene expression can enhance the accuracy of a responder vs. non-responder designation. For example, while the responder and non-responder groups show consistent differences in the pattern of change in these modules, individuals may have a more or less robust change. In the future we will generate an individually applicable “modular change score”. A diagnostic biomarker of this type may be useful in assigning likelihood for membership in the responder or non-responder group, allow for decisions on change from current therapy and to track the timing and durability of the response. As part of such an effort we plan to incorporate other current measures of disease activity, such as the proteomic panel described by the Vectra DA test(4, 22). Development of an improved RA disease status scoring method will require careful exploration of combinations of modules.
and other laboratory measures in order to optimize and validate such a method. Fortunately, an additional validation dataset of over 100 additional subjects in the BATTERUP study will shortly be available to permit a robust replication of the current findings as well as new combinatorial approaches to individual categorization of response at a molecular level. Our preliminary findings suggest that some patients who have a clinical response may not show strong evidence of a biological response, possibly due to placebo effect. Likewise, subjects with a biological response may exist within the “non-responder” group, again underlying heterogeneity in the mechanisms of disease. Methods to better quantitatively classify patients with regard to biological response will facilitate clinical treatment decisions and enable the practice of “individualized medicine”(23) including such concepts as treat to target(6). Finally, the outcome data and molecular interpretation presented in this manuscript are restricted to anti-TNF treatment. In future studies it will be interesting to examine the effect of alternative mechanisms of action on transcript modules.

In addition to these considerations, the need for physician based data to categorize patient response is a major barrier to carrying out large studies of biomarkers for drug response. Without physician based data, studies of drug response, even when prospective, are justifiably viewed with great skepticism. This means that any prospective study to investigate potential new predictive biomarkers is logistically difficult and extremely expensive. The current results suggest that it may be possible to combine changes in modular gene expression, with solely patient based data to reach a convincing categorization of drug response. If this can be achieved, it will dramatically enhance opportunities to recruit large cohorts of patients starting new drugs without the traditional expenses associated with physician led enrollment sites.

Acknowledgments

This work has been supported by the National Institutes of Health (NO1-AR1-2256), the BATTER-UP consortium (Biogen/IDEC, Bristol Myers-Squibb, Crescendo, Genentech, Janssen Research and Development LLC, Regeneron), and Janssen Research and Development LLC. We are also grateful to the Boas Family and Eileen Ludwig Greenland for general research support of Dr. Gregersen's laboratory over many years. All of the authors thank the patients who have participated in these studies.

REFERENCES


Arthritis Rheumatol. Author manuscript; available in PMC 2016 February 01.


Figure 1.
Summary of experimental approach and data analysis.
Figure 2.
Summary of change in modules at 14 weeks compared with baseline in responders (R) and non responders (NR) to anti-TNF therapy. The three cohorts - ABCoN (A), GO-FURTHER (G) and BATTER-UP (B) - are indicated by the three sets of bars for each module. The height of the bars shows the fraction of probes significantly changed (up or down) within each module. Statistical significance of change is indicated by shading: black bars P<0.0005, dark grey bars p<0.005, light grey bars p<0.05, open bars NS. Modules are designated according to the original nomenclature in ref (17), with description accompanying selected modules.
Figure 3.
The distribution of significantly changed (p<0.05, see methods) probes within modules between 14 weeks and baseline. The responders and non-responders are shown for each dataset: panel A – ABCoN; panel B – GO_FURTHER; panel C – BATTERUP. The colors indicate probes in modules with significant changes up (red) or down (blue), or no significant modular changes (hypergeometric test, p<0.005). As noted in the text, many of these modules exhibit changes with extremely high significance. For example p values for modules module 2.5 range are $10^{-37}$, $10^{-82}$ and $10^{-68}$ in panels A, B and C respectively.
Figure 4.
Comparison of module differences at baseline between responders and non-responders in all three cohorts. Bar height indicates fraction of probes that are different between responders and non-responders, bar shading indicates significance of changes (see legend Figure 2).
### Table 1
Clinical characteristics of patients in studied in the ABCoN, GO-FURTHER and BATTER-UP cohorts

<table>
<thead>
<tr>
<th>Column1</th>
<th>ABCoN (n=50)</th>
<th>GO_FURTHER (n=72)</th>
<th>BATTER-UP (n=118)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>55.62 sd 11.7</td>
<td>50.36 sd 10.91</td>
<td>53.3 sd 13.7</td>
</tr>
<tr>
<td>SEX F/M</td>
<td>39/11</td>
<td>53/19</td>
<td>96/22</td>
</tr>
<tr>
<td>CCP+and/orRF+</td>
<td>43 (86%)</td>
<td>72 (100%)</td>
<td>74 (62.7%)</td>
</tr>
<tr>
<td>DAS</td>
<td>5.24 sd .79</td>
<td>6.01 sd .84</td>
<td>5.12 sd 1.1</td>
</tr>
<tr>
<td>Good Responder</td>
<td>14</td>
<td>29</td>
<td>42</td>
</tr>
<tr>
<td>Moderate Responder</td>
<td>21</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Non Responder</td>
<td>15</td>
<td>6</td>
<td>39</td>
</tr>
<tr>
<td>Infliximab</td>
<td>20</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>Enbrel</td>
<td>21</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Golimumab</td>
<td></td>
<td>72</td>
<td>9</td>
</tr>
<tr>
<td>Humira</td>
<td>9</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Certolizumab Pegol</td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>White</td>
<td>44</td>
<td>72</td>
<td>100</td>
</tr>
<tr>
<td>Black</td>
<td>2</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>range</td>
<td>91-160</td>
<td>95-110</td>
<td>70-182</td>
</tr>
<tr>
<td>mean</td>
<td>103.4</td>
<td>99.3</td>
<td>105.1</td>
</tr>
<tr>
<td>SD</td>
<td>11.9</td>
<td>1.94</td>
<td>19.49</td>
</tr>
</tbody>
</table>