



Gene expression profiling of CD8⁺ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis

James C. Lee,^{1,2,3} Paul A. Lyons,^{1,2} Eoin F. McKinney,^{1,2} John M. Sowerby,^{1,2} Edward J. Carr,^{1,2} Francesca Bredin,^{2,3} Hannah M. Rickman,^{1,2} Huzefa Ratlamwala,^{1,2} Alexander Hatton,^{1,2} Tim F. Rayner,^{1,2} Miles Parkes,^{2,3} and Kenneth G.C. Smith^{1,2}

¹Cambridge Institute for Medical Research, ²Department of Medicine, University of Cambridge School of Clinical Medicine, and ³NHS Department of Gastroenterology, Addenbrooke's Hospital, Cambridge, United Kingdom.

Crohn disease (CD) and ulcerative colitis (UC) are increasingly common, chronic forms of inflammatory bowel disease. The behavior of these diseases varies unpredictably among patients. Identification of reliable prognostic biomarkers would enable treatment to be personalized so that patients destined to experience aggressive disease could receive appropriately potent therapies from diagnosis, while those who will experience more indolent disease are not exposed to the risks and side effects of unnecessary immunosuppression. Using transcriptional profiling of circulating T cells isolated from patients with CD and UC, we identified analogous CD8⁺ T cell transcriptional signatures that divided patients into 2 otherwise indistinguishable subgroups. In both UC and CD, patients in these subgroups subsequently experienced very different disease courses. A substantially higher incidence of frequently relapsing disease was experienced by those patients in the subgroup defined by elevated expression of genes involved in antigen-dependent T cell responses, including signaling initiated by both IL-7 and TCR ligation — pathways previously associated with prognosis in unrelated autoimmune diseases. No equivalent correlation was observed with CD4⁺ T cell gene expression. This suggests that the course of otherwise distinct autoimmune and inflammatory conditions may be influenced by common pathways and identifies what we believe to be the first biomarker that can predict prognosis in both UC and CD from diagnosis, a major step toward personalized therapy.

Introduction

Crohn disease (CD) and ulcerative colitis (UC) are chronic forms of inflammatory bowel disease (IBD) that predominantly affect young adults and cause considerable morbidity. UC typically presents with bloody diarrhea, and CD typically presents with abdominal pain, weight loss, and altered bowel pattern. Although once considered to be “Western” diseases, principally affecting North America and Western Europe, it is now clear that the incidence and prevalence rates of these diseases are rapidly increasing in other parts of the world, with dramatic increases noted in India, Japan, China, and the Middle East (1). Accordingly, these diseases represent an increasing burden upon global health, which is likely to continue to increase in the future.

A major determinant of the impact of CD or UC at an individual level is the clinical course experienced by each patient, which is known to be highly variable (2, 3). Currently, clinicians cannot accurately predict at diagnosis how an individual's disease will behave, and consequently, a “step-up” management strategy is conventionally used; escalating immunosuppression with corticosteroids, thiopurines, and then biological therapies only if a treatment-refractory course evolves. Recently, an alternative “top-down” strategy has been advocated, following evidence that early use of anti-TNF- α therapies induced higher remission rates in CD (4). In support of this, subgroup analyses of anti-TNF- α trials demonstrate higher response rates when these treatments are introduced earlier (5).

While trial evidence supports early aggressive therapy, safety concerns exist regarding the indiscriminate implementation of this strategy. These include the unnecessary immunosuppression of patients who were destined to experience an indolent course, even without additional therapies (2, 3), and the rare but potentially life-threatening side effects of such drugs, including opportunistic infection (6), demyelination (7), and malignancy (8). Moreover, use of biological therapy in all patients would be prohibitively expensive.

Accordingly, the ability to reliably predict an individual's prognosis, such that treatment strategies could be appropriately personalized from diagnosis, would represent a major clinical advance. Previous attempts to identify prognostic markers have focused on clinical parameters, but these factors (including younger age at diagnosis, early steroid requirement, and perianal involvement in CD) lack specificity, and hence are not clinically useful (9, 10). Biomarkers have also been studied, including anti-*Saccharomyces cerevisiae* antibodies (ASCAs). ASCA seropositivity has been repeatedly associated with need for surgery in CD, but these reports are largely retrospective and confounded by the trend to increasing ASCA titers with time (11, 12). Likewise, genetic variants at *NOD2* (13) and *HLA-DRB1*0103* (14) are statistically associated with need for surgery in CD and UC, respectively, but their poor sensitivity and the low frequency of the **0103* variant precludes their clinical use.

We recently observed that a common CD8 T cell transcriptional signature could be detected in 2 unrelated, autoimmune diseases — SLE and ANCA-associated vasculitis (AAV) — and that this predicted disease prognosis in both (15). UC and CD are not classi-

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: *J Clin Invest.* 2011;121(10):4170–4179. doi:10.1172/JCI59255.

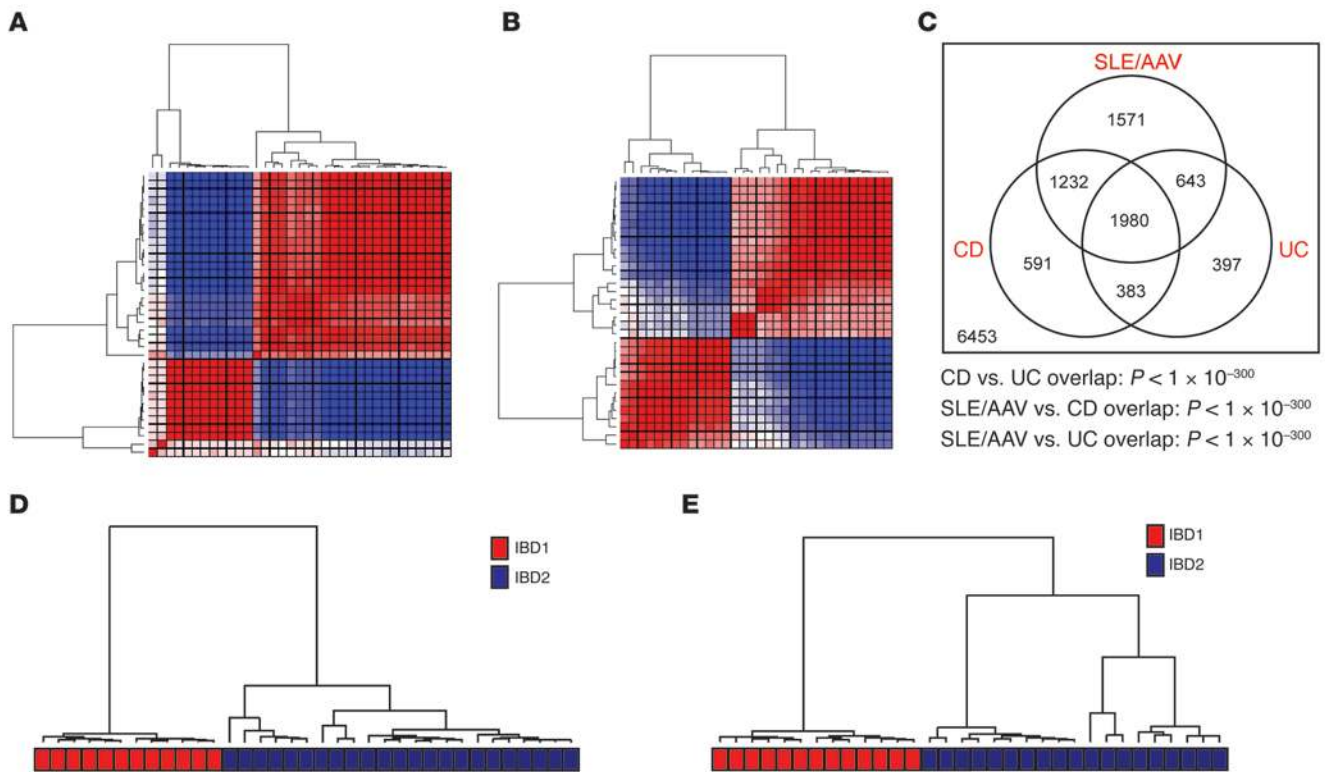


Figure 1

Overlapping CD8 T cell gene expression signatures divide patients with CD and UC into 2 distinct subgroups. Consensus clustering heat maps, demonstrating the merged output of 5,000 iterations of hierarchical and k-means clustering of (A) patients with CD and (B) patients with UC. Patient samples are arranged in the same order along the x and y axes. The colors of the intersecting squares represent the frequency with which samples cluster together, both within individual consensus clustering analyses and also between analyses using different methods of clustering. The color ranges from red (patients always cluster together) to blue (patients never cluster together). (C) Venn diagram illustrating the overlap between the gene signatures that distinguish the respective subgroups in CD, UC, and SLE/AAV (15). The statistical significance of each overlap was determined using a hypergeometric test. Numbers in the diagram refer to numbers of genes. Those inside the circles refer to the genes that are within each respective signature. The number outside the circles refers to the remaining genes expressed in CD8 T cells, which were not differentially expressed in any of the signatures. (D) The clusters of CD patients that were produced by k-means clustering for this cohort using the gene signature generated in the UC patients and (E) vice versa. The colored bar beneath each dendrogram corresponds to the original IBD1/2 subgroup membership.

cal autoimmune diseases, being thought to arise from an inappropriate immune response to gut microbiota, rather than to self antigens (16). However, irrespective of their classification, these diseases share a relapsing-remitting course driven by immunological responses to antigen(s); prompting us to hypothesize that an equivalent transcriptional signature might exist in CD and/or UC and, if present, would correlate with prognosis. We also investigated CD4 T cells for similar, clinically relevant transcriptional signatures, given that these are conventionally thought to be more important in IBD pathogenesis (17).

Results

We prospectively recruited 35 patients with active CD and 32 patients with active UC, prior to commencing treatment. Of these patients, 58% were newly diagnosed (23 out of 35 patients with CD and 16 out of 32 patients with UC). At enrollment, CD4 and CD8 T cells were positively selected from PBMCs for whole-genome transcriptional analyses. Patients were then managed conventionally using a step-up strategy by clinicians blinded to the microarray results (Supplemental Tables 1 and 2; supplemental material avail-

able online with this article; doi:10.1172/JCI59255DS1). We initially used an unsupervised approach to independently analyze the CD8 T cell gene expression data from each disease cohort, after filtering out genes that were not expressed. In both diseases, the distribution of the data was significantly different from that of the multivariate Gaussian distribution that would be expected if no substructure was present ($P < 1 \times 10^{-15}$). To investigate this further, we used consensus clustering (18), which iteratively resamples and clusters fractions of the data and provides a consensus output indicating whether stable and reproducible clusters/subgroups are present. This is superior to standard clustering algorithms, as it enables evaluation of which (and how many) clusters are genuinely present, as opposed to being artefacts of sampling variation. This demonstrated that the reason why we were unable to model either of the CD8 T cell transcriptional data sets with a Gaussian distribution was because 2 distinct patient subgroups were present within each disease cohort. Notably, these subgroups were detectable even if different clustering methods were used (k-means and hierarchical clustering) (Figure 1 and Supplemental Figure 1), although they could not be detected in unseparated PBMCs in either disease (Supplemental Figure 2).

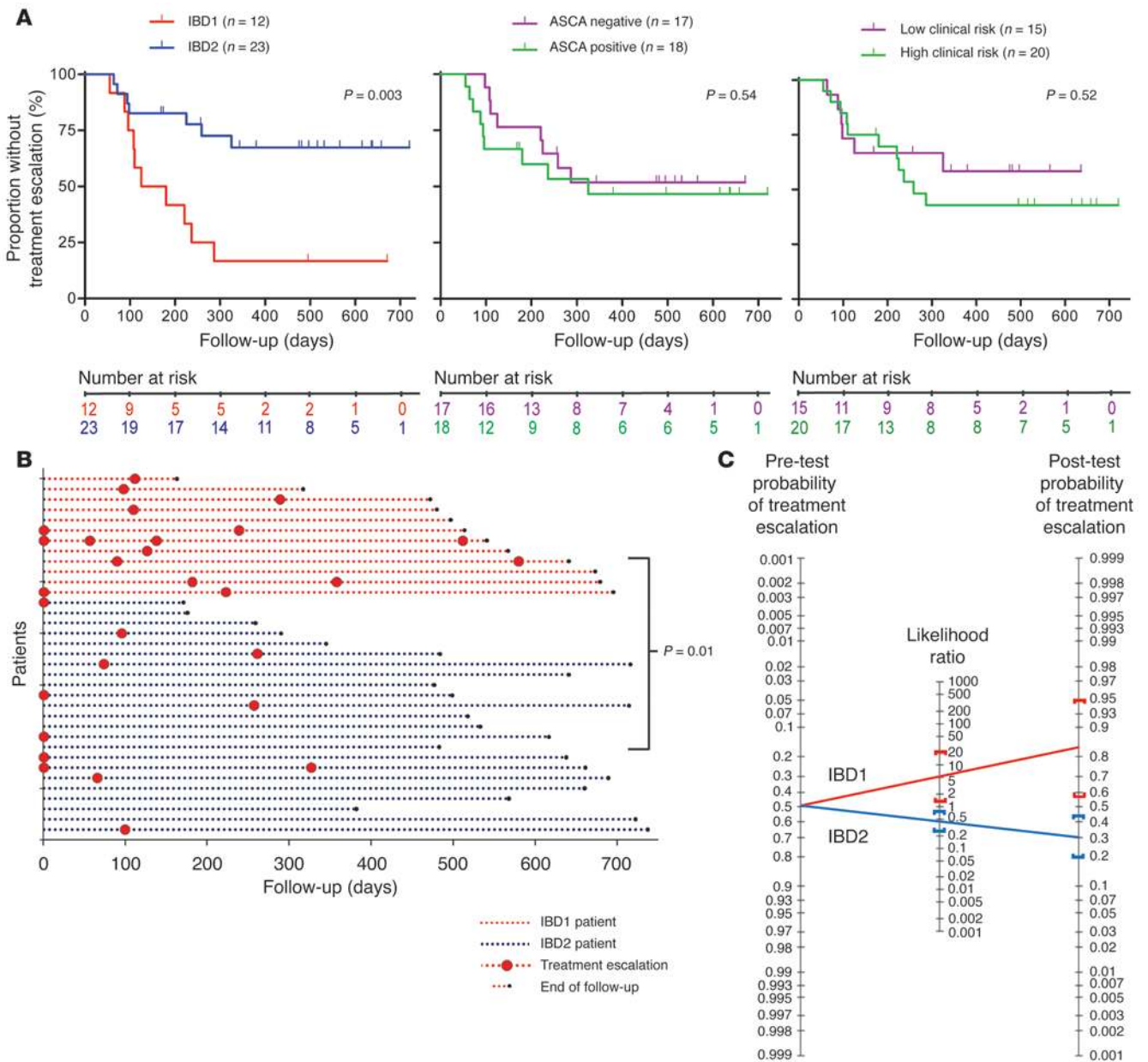


Figure 2

CD patients in subgroups IBD1 and IBD2 have significantly different disease courses. **(A)** Kaplan-Meier survival curves demonstrating the proportion of CD patients who did not require a subsequent treatment escalation (immunomodulator or surgery) after enrollment, as stratified by IBD1/2 subgroup (left) and ASCA serology (middle) and clinical parameters associated with complicated disease (right). Clinical parameters included age of less than 40 years at diagnosis, initial requirement for steroids, and perianal involvement. High risk of complicated disease was defined as 2 or more of these parameters; low risk was defined as fewer than 2 of these parameters. Statistical significance was determined using a log-rank test (df). Number at risk refers to the number of uncensored patients at each time point who remained at risk of requiring a treatment escalation. **(B)** Disease courses of all CD patients (y axis). The color of dotted lines reflects subgroup designation. In cases in which multiple treatment escalations are indicated, this universally reflects ongoing disease activity rather than intolerance to the initial treatment. Statistical significance was determined using a Fisher's exact test (2 df). **(C)** Bayes' nomogram demonstrating the effect that stratifying the CD cohorts by the IBD1/2 signature would have had upon the predicted requirement for treatment escalation: prior probability of treatment escalation, 48.6%; positive likelihood ratio, (sensitivity/[1-specificity]) 5.29 (95% CI, 1.35–21); negative likelihood ratio, ([1-sensitivity]/specificity) 0.46 (95% CI, 0.26–0.84); post-test probabilities, IBD1, 83% (95% CI, 56%–95%), and IBD2, 30% (95% CI, 20%–44%).

Next, we derived the lists of genes (signatures) that were differentially expressed between the 2 patient subgroups in each cohort. Of 13,250 genes that were considered to be expressed, 3,403 genes in the CD cohort and 4,186 genes in the UC cohort were significantly

differentially expressed between the patient subgroups after correction for multiple testing ($P < 0.05$) (data not shown; deposited in ArrayExpress, accession E-MTAB-331; <http://www.ebi.ac.uk/array-express/experiments/E-MTAB-331>). The expression of selected



genes was validated by qPCR (Supplemental Figure 3). Comparison of these signatures revealed highly significant overlap (hypergeometric $P < 1 \times 10^{-300}$; Figure 1). Indeed, this overlap was so considerable that both signatures could be used interchangeably, each being able to exactly reproduce — in the other disease cohort — the same subgroups that had been detected in the unsupervised analysis (Figure 1). Moreover, we were also able to generate smaller lists of genes, selected on their ability to distinguish the subgroups in half of the entire patient cohort, which could accurately designate patients with either CD or UC in the other independent half into their correct subgroups. The performance of these “classifiers” was similar, irrespective of the methods used to generate them or the number of genes they contained (between 4 and 100). An example is shown in Supplemental Figure 4 (positive predictive value 100%, negative predictive value 100%). We therefore termed the corresponding patient subgroups defined by these analogous signatures as “IBD1” (the subgroup characterized by elevated expression of the majority of differentially expressed genes) and “IBD2” (the subgroup characterized by lower expression of these genes).

Comparison with the SLE/AAV prognostic transcriptional signature. Given our original hypothesis, we then compared these transcriptional signatures with the prognostic signature detected in AAV and SLE (15). The genes within the analogous UC and CD signatures highly overlapped with those within the SLE/AAV signature (hypergeometric $P < 1 \times 10^{-300}$; Figure 1), with reciprocal enrichment of the upregulated and downregulated SLE/AAV signature genes in the corresponding parts of the IBD1/2 signature (Supplemental Figure 5). Accordingly, these data confirm that the transcriptional signatures that we have independently detected in CD and UC are analogous to a prognostic signature previously detected in unrelated autoimmune diseases.

The IBD1/2 signature predicts subsequent disease course in UC and CD. We then investigated whether subgroups IBD1 and IBD2 correlated with clinical course in UC and CD. In both diseases, an early indicator of an aggressive/progressive course is that an individual’s disease is persistently active, either in a continuous or a frequently relapsing fashion. The common end point for either of these patterns of disease behavior is the necessity to escalate treatment; initially with oral immunomodulators and subsequently — if disease activity persists — with alternative immunomodulators, biological therapies, or surgery. As patients with quiescent disease do not receive any such treatment escalations, we used the requirement to introduce such therapies after initial induction of disease remission to assess disease course (Supplemental Tables 1 and 2). In the CD cohort, patients in subgroup IBD1 (characterized by upregulation of the majority of differentially expressed genes) experienced relapsing or chronically active disease, such that a treatment escalation was necessary, significantly more frequently than those in subgroup IBD2 ($P = 0.003$; Figure 2). Moreover, of all the patients who required a treatment escalation at any stage, those in subgroup IBD1 were significantly more likely to continue to experience persistently active disease, despite the initial intervention, necessitating one or more further treatment escalations ($P = 0.01$; Figure 2 and Supplemental Figure 6). Of note, all additional treatment escalations were due to ongoing disease activity rather than intolerance to the initial treatment.

In the UC cohort, patients in subgroup IBD1 also experienced a more aggressive disease course than those in subgroup IBD2. This was similarly characterized by a higher incidence of recurrently

active disease requiring immunomodulators ($P = 0.0002$; Figure 3). The total number of escalations required throughout the follow-up period was also higher, as was observed in the CD IBD1 subgroup ($P = 0.002$; Figure 3).

To provide a direct estimate of the prognostic utility of this signature, we calculated the specificity (CD, 89%; UC, 84%) and sensitivity (CD, 59%; UC, 77%) of its ability to predict the future requirement for any treatment escalation. Using these values, we derived positive and negative likelihood ratios for each disease, which were then applied to the overall prevalence of requiring a treatment escalation in our cohorts (CD, 48.6%; UC, 40.6%). This was performed using a Bayes nomogram (19), which is a graphical tool for estimating how much a test can alter the probability of a particular outcome — in this case that a treatment escalation would be required in the future due to continuously active or frequently flaring disease. The results are shown in Figures 2 and 3 and demonstrate that prestratifying patients with CD or UC according to their IBD1/2 subgroup membership would have effectively distinguished those patients who were destined to experience more aggressive disease from those who went on to experience indolent disease.

We then investigated whether we could identify these patient subgroups by other means. We observed no significant association between the subgroups and any contemporaneous laboratory or clinical parameters recorded at enrollment (Table 1). There was also no association with prior duration of disease, the subgroups being detectable in both newly and previously diagnosed patients (Table 1). Accordingly, at enrollment, it was not possible to identify the subgroup to which a patient belonged by other means.

We then considered other clinical and serological parameters, which have been reported to associate with disease course based on analysis of retrospective data. In CD, the development of more aggressive disease has been associated with 2 or more of the following factors: age of less than 40 years at diagnosis, early requirement for corticosteroids, and perianal disease (9). In UC, extensive disease and older age have been reported to be “negative prognostic predictors” (20). We therefore stratified both cohorts into high- and low-risk subgroups according to these clinical risk factors but did not detect any prognostic association with disease course in either cohort (Figures 2 and 3). We also assessed *S. cerevisiae* serology in the CD cohort. Fifty-one percent of patients were ASCA positive at enrollment, which correlated significantly with previous diagnosis ($P = 0.01$, Fisher’s exact test) but not with subsequent disease course ($P = 0.54$; Figure 2).

Pathway analysis. Due to the large size of these signatures, we investigated whether the differentially expressed genes were coordinately found within specific cellular pathways that could elucidate aspects of the underlying biology. This was performed using Gene-Set Enrichment Analysis (GSEA) (21), which is a computational method that determines whether the expression of whole lists of genes (predefined as being implicated in specific pathways) is significantly and concordantly different between 2 biological states (in this case IBD1 and IBD2). We did this instead of focusing on individual genes, as, while many are likely to have been of interest, they are — in isolation — less likely to have had substantial biological effects (21). We observed significant, reproducible enrichment for several pathways (Supplemental Table 3), including TCR signaling, CD28 costimulation, IL-7 signaling, and IL-2 signaling (Supplemental Table 4 and Supplemental Figure 7). These pathways, which are upregulated in IBD1 patients, are implicated in T

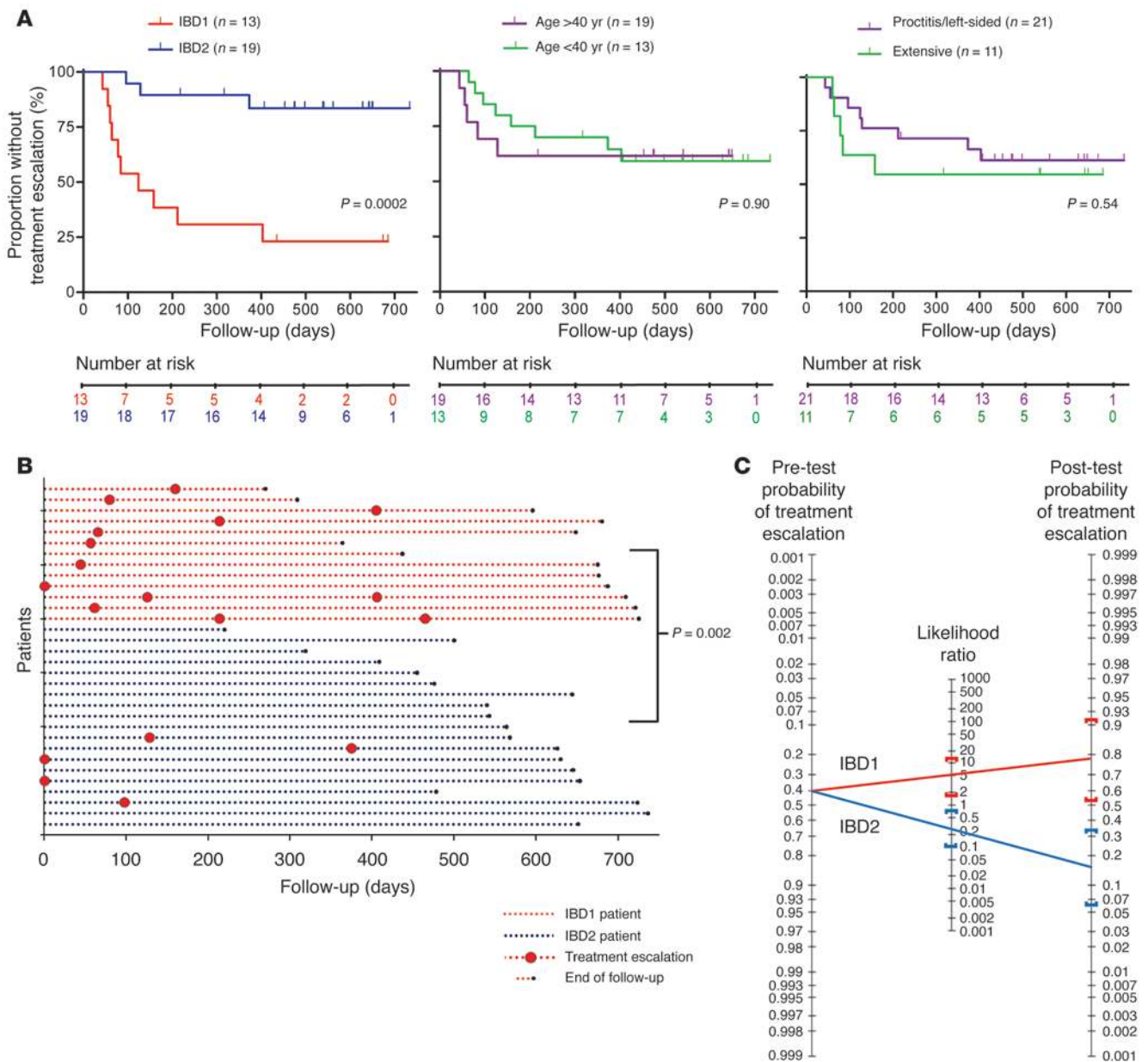


Figure 3 UC patients in subgroups IBD1 and IBD2 also have significantly different disease courses. **(A)** Kaplan-Meier survival curves demonstrating the proportion of UC patients who did not require a subsequent treatment escalation (immunomodulator or surgery) after enrollment as stratified by IBD1/2 subgroup (left), age at diagnosis (middle), and disease extent (right). Statistical significance was determined using a log-rank test (1 df). Number at risk refers to the number of uncensored patients at each time point who remained at risk of requiring a treatment escalation. **(B)** Disease courses of all UC patients (y axis). Format is identical to that used in Figure 2B. Statistical significance was determined using a Fisher’s exact test (2 df). **(C)** Bayes’ nomogram demonstrating the effect that stratifying the UC cohorts by the IBD1/2 signature would have had upon the predicted requirement for treatment escalation: prior probability of treatment escalation, 40.6%; positive likelihood ratio, 4.87 (95% CI, 1.65–14); negative likelihood ratio, 0.27 (95% CI, 0.10–0.75); post-test probabilities, IBD1, 77% (95% CI, 53%–91%), and IBD2, 16% (95% CI, 6%–34%).

cell activation and the subsequent development of antigen-specific T cell memory (22, 23), implying that there may be a difference in the activation status of CD8 T cells between the subgroups. To further investigate this, we then examined whether a signature of CD8 T cell activation, which was derived experimentally by stimulating primary human CD8 T cells (24), was enriched within the transcriptional signature that distinguished subgroups IBD1 and

IBD2. Experimentally derived signatures have one advantage over annotated signatures, as they are not reliant upon data curation but rather reflect the sum of the transcriptional changes – throughout the whole genome – that are induced upon activation. This demonstrated that there was a significant difference in CD8 T cell activation status between the IBD1 and IBD2 patients ($P = 0.048$; Supplemental Figure 8), with T cell activation genes being

**Table 1**
Baseline characteristics

| | CD | | | UC | | |
|-------------------------------------|------------------|------------------|------|------------------|------------------|------|
| | IBD1 (n = 12) | IBD2 (n = 23) | P | IBD1 (n = 13) | IBD2 (n = 19) | P |
| Age (yr) ^A | 27.5 (25.9–32.6) | 34.0 (27.7–44.1) | 0.19 | 48.8 (36.3–50.8) | 40.8 (29.3–49.5) | 0.38 |
| Hemoglobin (g/dl) ^A | 13.2 (12.2–14.4) | 12.5 (11.6–13.8) | 0.51 | 14.0 (13.5–14.4) | 13.1 (12.4–14.8) | 0.51 |
| Albumin ^A | 35.5 (32–38) | 36 (34–39) | 0.45 | 40 (38–41) | 40 (39–41) | 0.60 |
| CRP (mg/l) ^A | 24 (14–35) | 23 (10–51) | 0.93 | 7 (6–33) | 10 (6–19) | 0.80 |
| ESR (mm/h) ^A | 25 (16–35) | 25 (18–41) | 0.66 | 16 (11–27) | 18 (11–26) | 0.87 |
| HBSI ^A | 10.5 (8–12.5) | 10 (9–12) | 0.77 | – | – | – |
| SCCAI ^A | – | – | – | 10 (8–11) | 8 (7–9) | 0.27 |
| Follow-up (d) ^A | 526 (476–647) | 531 (428–659) | 0.82 | 575 (435–685) | 562 (475–643) | 0.27 |
| Previously diagnosed ^B | 5 (42%) | 7 (30%) | 0.71 | 8 (62%) | 8 (42%) | 0.45 |
| Sex (M) ^B | 5 (42%) | 9 (39%) | 1.00 | 6 (46%) | 6 (32%) | 0.47 |
| Smoker ^B | 5 (42%) | 11 (48%) | 1.00 | 1 (8%) | 0 (0%) | 0.41 |
| ASCA seropositive ^B | 5 (42%) | 13 (57%) | 0.49 | – | – | – |
| Disease distribution ^{B,C} | | | | | | |
| CD – L1 (ileal) | 2 (17%) | 5 (22%) | 1.00 | – | – | – |
| CD – L2 (colonic) | 4 (33%) | 5 (22%) | 0.69 | – | – | – |
| CD – L3 (ileocolonic) | 6 (50%) | 13 (57%) | 0.74 | – | – | – |
| CD – L4 (upper GI) | 0 (0%) | 2 (9%) | 0.54 | – | – | – |
| Perianal | 2 (17%) | 2 (9%) | 0.59 | – | – | – |
| UC – E1 (proctitis) | – | – | – | 3 (23%) | 3 (16%) | 0.67 |
| UC – E2 (left-sided) | – | – | – | 4 (31%) | 9 (47%) | 0.47 |
| UC – E3 (extensive) | – | – | – | 6 (46%) | 7 (37%) | 0.72 |

Data shown in parentheses indicate median (interquartile range) values or percentages unless specified. ^AStatistical significance was calculated using a Mann-Whitney test (2 tailed) for continuous variables and ^Ba Fisher's exact test (2 tailed) for dichotomous variables. ^CDisease distribution was classified according to the Montreal Classification (29). GI, gastrointestinal; HBSI, Harvey-Bradshaw severity index; M, male; SCCAI, simple clinical colitis activity index.

expressed at a relatively higher level in the IBD1 patients – consistent with the implications of the pathway analysis results. Finally, we examined whether we could detect any differences in the relative size or surface phenotype of the cellular subsets of CD8 T cell memory (including central and effector memory populations; ref. 25) by contemporaneous flow cytometry. Notably, when examining these memory subpopulations – which will have contained cells specific to a range of antigens (not just those specific to IBD-related epitopes) – we were not able to detect any significant differences between the subgroups (Supplemental Figure 9).

Transcriptional differences in CD4 T cells do not predict disease course. There is a large body of evidence implicating roles for various CD4 T cell subsets in the pathogenesis of both CD and UC. We therefore used the same methodology to analyze the expression data from the 13,709 genes deemed to be expressed in CD4 T cells. Although we detected significant substructure in each disease cohort that was best described by the presence of 2 subgroups, these were less distinct than those observed in the CD8 T cell gene expression data (Supplemental Figure 10), and different clusters of patients were produced in both cohorts depending on the clustering algorithm used. We compared the lists of genes that were differentially expressed between the CD4 subgroups in each disease but failed to demonstrate significant overlap ($P = 0.12$). Furthermore, these subgroups did not correlate with subsequent disease course (Supplemental Figure 10) or any other clinical parameter.

Discussion

A major barrier to personalized medicine in IBD is the lack of suitable biomarkers to guide treatment early in the disease course. In oncology, gene expression profiling has been used to

identify transcriptional signatures that predict several aspects of disease behavior, including risk of metastasis and response to chemotherapy (26–28). Where such techniques have previously been applied to autoimmune and inflammatory conditions, they have generally not detected signatures with equivalent prognostic utility. This may have been because the tissues commonly examined, such as PBMCs or mucosal biopsies, are heterogeneous, and hence any transcriptional variation detected will predominantly reflect differences in the cellular composition between samples. Indeed, such analyses have been shown to be insensitive to the differences that may have been detectable in separated cell subsets (29). Recently, using separated CD8 T cells, we have identified what we believe to be a novel transcriptional signature that predicted outcome in SLE and AAV and that was enriched for genes within the IL-7 and TCR signaling pathways. We therefore set out to investigate whether this transcriptional signature might exist in patients with CD and UC, which, despite being very different conditions, share a relapsing-remitting course driven by immunological responses to antigen(s). In both UC and CD, we identified, using an unsupervised approach, a common transcriptional signature in separated CD8 T cells that was analogous to the prognostic signature previously described in SLE and AAV. In both IBD cohorts, patients whose CD8 T cells were enriched for this signature (subgroup IBD1) had a substantially higher incidence of experiencing treatment-refractory, relapsing, or chronically active disease – concordant with the clinical phenotype observed in SLE and AAV. These observations therefore validate our a priori hypothesis that the signature may be present and that, if present, it would associate with altered prognosis.



The patterns of relapsing or chronically active intestinal inflammation that were commonly experienced by patients in the IBD1 subgroup are associated with considerable morbidity, including the development of medically irreversible complications, and represent the courses of disease most likely to benefit from early top-down therapy (5). Accordingly, these data suggest that gene expression profiling may represent the first method by which treatment strategies could be appropriately personalized at diagnosis. Such preemptive stratification would be expected to improve the therapeutic outcome in those patients with IBD destined to run a refractory course and avoid unnecessary immunosuppression in those with indolent disease. However, as a top-down strategy was not used in our cohorts, these expected results require confirmation in a prospective trial.

It is not possible to fully elucidate the biological differences that account for such transcriptional variation with this sort of microarray-based study. However, it is noteworthy that pathway analysis identified several pathways – upregulated in IBD1 patients – that are associated with the activation of CD8 T cells in response to MHC class I-bound antigen and their subsequent proliferation and differentiation into effector cells (TCR signaling, CD28 costimulation, and IL-2 signaling). This would imply that concordant differences in the activation status of CD8 T cells may contribute to the differences between these subgroups. Consistent with this, we observed that CD8 T cells from IBD1 patients were relatively enriched for a CD8 T cell activation signature in comparison with those from IBD2 patients. Studies of how memory T cells are generated after antigen exposure have shown that the more activated T cells become in response to an antigen (reflected in the “clonal burst” size), the more memory T cells are subsequently formed (30). Notably, the other key determinant of memory T cell generation is IL-7 signaling, which facilitates the survival and differentiation of effector cells into long-lived antigen-specific memory cells (23) and which was also upregulated in IBD1 patients. Accordingly, the combination of greater T cell activation in conjunction with greater IL-7 signaling would be expected to result in more antigen-specific memory being generated in IBD1 patients, which would, in turn, facilitate more rapid and potent CD8 effector responses upon future reencounters with that antigen (31). Such antigen-specific, cytotoxic CD8 T cells are known to be detectable in increased numbers in the intestinal mucosa in active CD and UC (32, 33), and therefore it would seem plausible that differences in antigen-specific T cell memory between the subgroups could manifest clinically as differences in the future behavior of these relapsing-remitting immune-driven diseases. However, while these data are consistent with a role for CD8 T cells in determining the natural history of both UC and CD, they do not prove that this is the case. Indeed, confirmation of whether correlation exists between the quality or quantity of antigen-specific memory formed during a flare and subsequent disease course will probably require knowledge of the specific antigens involved, as no differences were detectable in the overall CD8 T cell memory compartments (as assigned by expression of surface markers). Importantly, this result was not unexpected, and it does not disprove that differences in memory T cells may exist between the subgroups. This is because the memory compartments examined by flow cytometry will have contained cells specific to a variety of previously encountered antigens – not just those associated with IBD. This heterogeneity would obviously limit our ability to

detect differences that only involve a subpopulation of memory cells (e.g., those specific to IBD-associated antigens). However, these differences could still be detectable at a gene expression level, due to the increased transcriptional activity of this subpopulation of cells during an antigen-specific immune response (i.e., a disease flare). A further potentially confounding issue is that the assignment of “memory” function on the basis of the expression of a few surface markers is known to be imperfect (34). Accordingly, any discrepancy between immunophenotype and function would also reduce our ability to detect potential differences between the subgroups.

Another question, which these data present, but which also cannot be directly addressed in such a microarray-based study, is the mechanism(s) that underlies such transcriptional variation, as this could reveal novel therapeutic targets. One possibility would be that the differences arise as a primary phenomenon (e.g., due to differences in genetic, ref. 35, or epigenetic variation, ref. 36–38), such that CD8 T cells from the 2 subgroups were predestined to respond differently to antigen. This could still feasibly be consistent with differences in IL-7 signaling between the IBD1 and IBD2 subgroups, as this pathway is also critical in T cell development through supporting the differentiation of common lymphoid precursors (39). However, if this were true, one might have expected the resulting differences in the number of effector and memory cells to be detectable by flow cytometry, given that all antigen responses would be expected to be similarly affected. Alternatively, these differences could arise as a secondary effect and hence be contingent upon the context in which CD8 T cells encounter antigen. This would implicate other cells in driving the transcriptional differences, such as antigen-presenting cells and/or helper CD4 T cells (22).

The lack of clinically relevant transcriptional variation in CD4 T cells is noteworthy in the context of the CD8 T cell prognostic signature, particularly as CD4 T cells are traditionally thought to be more important in IBD pathogenesis. While this result may therefore appear surprising, there are several reasons why these data do not conflict with current models of disease pathogenesis. First, although a prognostic signature was not present in CD4 T cells, this does not imply that they are not involved in affecting disease prognosis, only that their involvement is of a nature that cannot be wholly reflected in simple transcriptional changes. Indeed, it is even possible that the CD8 T cell transcriptional signature is a secondary phenomenon arising from other immunological events associated with disease pathogenesis. These could include CD4 T cell responses, which are known to be required for the development of effective CD8 T cell memory (22). Second, although there is no reason to suppose that a prognostic CD4 T cell signature must exist, the relative heterogeneity of circulating CD4 T cells (which comprise varying proportions of Th subsets [including Th1, Th2, and Th17] and regulatory T cells) will have reduced our sensitivity to detect such a signature (29). It could therefore be speculated that an equivalent prognostic signature might exist in one of the CD4 Th subsets, although in order to test this hypothesis, it would have been necessary to contemporaneously isolate and array these individual subpopulations. Third, while the full complexity of the adaptive immune response in IBD is not well understood, it is clear that CD8 T cells – as well as CD4 T cells – do have an important role. Indeed, in addition to activated effector CD8 T cells being detectable in the mucosa in IBD, several animal models (40–43)



have suggested it is the destruction of intestinal epithelial cells by CD8 T cells that represents the primary event leading to the loss of barrier function and enhanced exposure to microbial antigens that ultimately drives disease activity. Accordingly, it has been proposed that CD8 T cells may play an earlier role in IBD development or relapse than the CD4 T cells that are more traditionally associated with disease pathogenesis (44), which may be relevant in the context of our observations.

It is interesting that our data confirm both the existence of this signature and its association with more aggressive disease behavior in 2 diseases that, although being immunologically driven, are not traditionally classified as being autoimmune. Genetic studies have previously shown that certain polymorphisms can contribute to the development of several different diseases (45). Our data suggest that as well as sharing common genomic variation, distinct disease states can also be influenced by common variation at the level of the transcriptome and that this can have direct relevance for disease prognosis. The common underlying biology accounting for this could include, for example, features of the CD8 T cell response to MHC class I-restricted antigen, which would be consistent with the pathway analysis results and the differences observed in subsequent disease course. Interestingly, this would mean that such transcriptional differences can occur – and can be detected – irrespective of the nature of the causative antigen(s) (self vs. foreign) or the genetic and environmental susceptibility factors that combine with the immune response to determine the resulting disease phenotype. While this is only one possible explanation, it is not necessary to understand the precise mechanism(s) to appreciate that there are wide-ranging implications of the observation that generic, rather than disease-specific, factors may significantly affect the natural history of distinct immune-driven diseases. Indeed, we would ultimately hope that better characterization of the biology that underlies this transcriptional variation might reveal novel therapeutic targets that could be relevant for several diseases in which effector T cells play a role in pathogenesis.

In conclusion, we have shown that a gene expression signature, detectable at diagnosis in patients with UC and CD, is associated with a significantly more aggressive disease course in both conditions. This represents, to our knowledge, the first biomarker that has been prospectively shown to predict the course of both UC and CD from diagnosis. This could therefore enable patients with either condition to be stratified to receive personalized therapy according to their disease prognosis and, accordingly, represents a major step toward individualized management in the treatment of these common and disabling conditions.

Methods

Patient recruitment. Patients with active CD and UC were recruited from a specialist IBD clinic at a tertiary referral hospital, prior to commencing treatment. Diagnosis was made using standard endoscopic, histologic, and radiological criteria (46). Patients receiving immunomodulators or corticosteroids were excluded due to potential effects on gene expression. Enrolled patients were managed conventionally using a step-up strategy (Supplemental Tables 1 and 2). Assessment of disease activity was in accordance with national and international guidelines and included consideration of symptoms, clinical signs, and objective measures, including blood tests (C-reactive protein [CRP], erythrocyte sedimentation rate [ESR], hemoglobin concentration, and serum albumin), stool markers (calprotectin), and mucosal assessment (by sigmoidoscopy or colonoscopy) where appropriate. Validated scoring

tools were used as another means of assessing disease activity (Harvey-Bradshaw severity index, ref. 47, or simple clinical colitis activity index, ref. 48, for CD and UC, respectively), although these were not used to guide treatment decisions. All clinicians were blinded to the microarray results.

Cell separation. A 110-ml blood sample was taken from eligible patients. A 10-ml sample was used for flow cytometric immunophenotyping and to obtain serum. PBMCs were isolated by density centrifugation, and CD4 and CD8 T cells were positively selected as previously described (49) (median purity, 93.8% and 91.0%, respectively).

RNA extraction and microarray analysis. RNA was extracted from PBMC and CD4 and CD8 T cell lysates using RNEasy Mini Kits (Qiagen) according to the manufacturer's instructions. RNA quantity and quality were determined using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) and an Agilent 2100 Bioanalyzer (Agilent Technologies). 200 ng RNA was processed for hybridization onto Affymetrix Human Gene ST 1.0/1.1 microarrays, according to the manufacturer's instructions, prior to scanning.

Data analysis. Raw data were preprocessed (normalization [vsnrma], refs. 50, 51; quality-control evaluation, ref. 52; and batch effect correction, ref. 53) using BioConductor (<http://www.bioconductor.org/>) in R (<http://www.r-project.org/>). A gene filter was used to exclude genes that were not expressed but which could affect any results through noise. This excluded probes whose signal intensity value was below a background level (set using a reference control data set) in more than 10% of samples. To test our a priori hypothesis that a clinically relevant transcriptional signature would exist, we first used unsupervised consensus clustering (5,000 iterations, 80% subsampling) to investigate whether any subgroups were present in either the CD4 or CD8 T cell gene expression data from both diseases (18). This was performed using k-means and hierarchical clustering to enable comparison of the results between clustering algorithms (54). The significance of any apparent substructure was determined by comparing the distribution of our data with that which would be expected if no substructure was present (55).

Differentially expressed genes were identified using linear modeling and an empirical Bayes method (56). We used a family-wise error rate method to correct for multiple testing. This particularly stringent correction was chosen over more liberal methods, because avoidance of type-1 errors was more important than optimization of statistical power (at the expense of including false positives). GSEA (21) was used to assess whether the differentially expressed genes were enriched within specific biological pathways. We initially used the database of pathways that have been curated from BioCarta (<http://www.biocarta.com>) by the Molecular Signatures Database (<http://www.broadinstitute.org/gsea/msigdb>) (21) and subsequently sought to confirm the implications of this analysis by using an experimentally derived signature of CD8 T cell activation (24). The initial exploratory analysis was performed using independent subsets of the overall data set to confirm that any enrichment was reproducible. Pathways were considered significantly enriched if the enrichment ($P < 0.05$) was reproduced in the second independent cohort ($P < 0.05$; false discovery rate $q < 0.25$). We also used these cohorts to generate and test simpler methods of ascribing subgroup designation (using smaller numbers of genes than are assayed on a whole microarray) to ascertain whether a more practical test could be developed. Several methods of generating such classifiers were investigated, including weighted voting and random forests. Further details of the bioinformatic analyses are provided in the Supplemental Methods. Raw data, transformed data, and the transcriptional signatures are deposited in ArrayExpress (accession E-MTAB-331; <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-331>).

Quantitative PCR. *CD44*, *IL10RB*, *ILF2*, and *IL2RG* mRNA levels were determined using TaqMan Expression Assays (Applied Biosystems) on an ABI Prism 7900HT instrument according to the manufacturer's instructions. Transcript abundance was calculated by comparison with a standard curve.



Immunophenotyping. Flow cytometric immunophenotyping of T cell memory compartments was performed on a CyAn ADP flow cytometer (Beckman Coulter) using the following antibodies: CD3-PE/Cy5, CD8-Pacific Blue, CD45RA-PE/Cy7, CD62L-FITC, CD127-PE, CD25-PE, and IgG1-PE isotype control (BD Biosciences). Data were analyzed using FlowJo software (TreeStar).

ELISAs. Sera were tested for the presence of *S. cerevisiae* IgA and IgG antibodies using a commercially available sandwich ELISA (ASCA Screen ELISA, IBL Hamburg) according to the manufacturer's instructions.

Statistics. The statistical/bioinformatic analyses of microarray data are described in the *Data Analysis* section, with additional details in the Supplemental Methods. Other statistical analyses were as follows. The statistical significance of the overlap among any gene signatures was determined with a hypergeometric test (1 degree of freedom [1 df]). Comparison of the clinical and laboratory parameters among patients in different subgroups was performed using Fisher's exact test (1 df) for dichotomous variables and Mann-Whitney test (2 tailed) for continuous variables. Comparison of the initial requirement for a treatment escalation was performed using a log-rank test (1 df) and graphically represented with a Kaplan-Meier plot. Comparison of the total number of treatment escalations required was performed using a Fisher's exact test (2 df). The α value for these analyses was 0.05 and was corrected for multiple testing (as detailed) where appropriate.

Study approval. Ethical approval was for this work was obtained from the Cambridgeshire Regional Ethics Committee (REC08/H0306/21). All participants provided written informed consent.

Acknowledgments

We are grateful to the patients who have provided samples for this study and to the IBD service at Addenbrooke's Hospital, Cambridge, for helping identify suitable patients. We thank David Clayton, Jeffrey Barrett, and Johan Rung for discussions and guidance regarding statistical analysis and Arthur Kaser for critical reading of the manuscript. This work was supported by the UK National Institute of Health Research Cambridge Biomedical Research Centre, the Wellcome Trust (programme grant number 083650/Z/07/Z), and Crohn's and Colitis UK (NACC). J.C. Lee holds a Wellcome Trust Clinical PhD Programme Fellowship, E.F. McKinney holds a Wellcome Trust Clinical Training Fellowship, and E.J. Carr holds a MRC Doctoral Training Account studentship. K.G.C. Smith is a Lister Prize Fellow. The Cambridge Institute for Medical Research is in receipt of Wellcome Trust Strategic Award 079895.

Received for publication May 31, 2011, and accepted in revised form August 10, 2011.

Address correspondence to: Kenneth G.C. Smith, Cambridge Institute for Medical Research — Box 139, Addenbrooke's Hospital, Cambridge, CB2 0XY, United Kingdom. Phone: 44.1223.336848; Fax: 44.1223.336846; E-mail: kgs2@cam.ac.uk.

1. Lakatos PL. Recent trends in the epidemiology of inflammatory bowel diseases: up or down? *World J Gastroenterol.* 2006;12(38):6102–6108.
2. Jess T, et al. Changes in clinical characteristics, course, and prognosis of inflammatory bowel disease during the last 5 decades: a population-based study from Copenhagen, Denmark. *Inflamm Bowel Dis.* 2007;13(4):481–489.
3. Solberg IC, et al. Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study. *Clin Gastroenterol Hepatol.* 2007;5(12):1430–1438.
4. D'Haens G, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet.* 2008;371(9613):660–667.
5. D'Haens G. Top-down therapy for IBD: rationale and requisite evidence. *Nat Rev Gastroenterol Hepatol.* 2010;7(2):86–92.
6. Keane J, et al. Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N Engl J Med.* 2001;345(15):1098–1104.
7. Fromont A, De Seze J, Fleury MC, Mailliefert JF, Moreau T. Inflammatory demyelinating events following treatment with anti-tumor necrosis factor. *Cytokine.* 2009;45(2):55–57.
8. Siegel CA, Marden SM, Persing SM, Larson RJ, Sands BE. Risk of lymphoma associated with combination anti-tumor necrosis factor and immunomodulator therapy for the treatment of Crohn's disease: a meta-analysis. *Clin Gastroenterol Hepatol.* 2009;7(8):874–881.
9. Beaugerie L, Seksik P, Nion-Larmurier I, Gendre JP, Cosnes J. Predictors of Crohn's disease. *Gastroenterology.* 2006;130(3):650–656.
10. Loly C, Belaiche J, Louis E. Predictors of severe Crohn's disease. *Scand J Gastroenterol.* 2008; 43(8):948–954.
11. Seow CH, et al. Novel anti-glycan antibodies related to inflammatory bowel disease diagnosis and phenotype. *Am J Gastroenterol.* 2009;104(6):1426–1434.
12. Ferrante M, et al. New serological markers in inflammatory bowel disease are associated with complicated disease behaviour. *Gut.* 2007; 56(10):1394–1403.
13. Buning C, et al. Mutations in the NOD2/CARD15 gene in Crohn's disease are associated with ileocecal resection and are a risk factor for reoperation. *Aliment Pharmacol Ther.* 2004;19(10):1073–1078.
14. Ahmad T, et al. The contribution of human leucocyte antigen complex genes to disease phenotype in ulcerative colitis. *Tissue Antigens.* 2003;62(6):527–535.
15. McKinney EF, et al. A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. *Nat Med.* 2010;16(5):586–591.
16. Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature.* 2007;448(7152):427–434.
17. Powrie F. T cells in inflammatory bowel disease: protective and pathogenic roles. *Immunity.* 1995; 3(2):171–174.
18. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning.* 2003;52:91–118.
19. Fagan TJ. Letter: Nomogram for Bayes theorem. *N Engl J Med.* 1975;293(1):257.
20. Romberg-Camps MJ, et al. Influence of phenotype at diagnosis and of other potential prognostic factors on the course of inflammatory bowel disease. *Am J Gastroenterol.* 2009;104(2):371–383.
21. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550.
22. Bevan MJ. Helping the CD8(+) T-cell response. *Nat Rev Immunol.* 2004;4(8):595–602.
23. Kaech SM, Tan JT, Wherry EJ, Konieczny BT, Surh CD, Ahmed R. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nat Immunol.* 2003;4(12):1191–1198.
24. Wang M, Windgassen D, Papoutsakis ET. Comparative analysis of transcriptional profiling of CD3+, CD4+ and CD8+ T cells identifies novel immune response players in T-cell activation. *BMC Genomics.* 2008;9:225.
25. Sallusto F, Lenig D, Forster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature.* 1999; 391(6754):708–712.
26. Chang JC, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet.* 2003;362(9381):362–369.
27. Watanabe T, Kobunai T, Tanaka T, Ishihara S, Matsuda K, Nagawa H. Gene expression signature and the prediction of lymph node metastasis in colorectal cancer by DNA microarray. *Dis Colon Rectum.* 2009;52(12):1941–1948.
28. Watanabe T, et al. Gene expression signature for recurrence in stage III colorectal cancers. *Cancer.* 2009; 115(2):283–292.
29. Lyons PA, et al. Novel expression signatures identified by transcriptional analysis of separated leucocyte subsets in systemic lupus erythematosus and vasculitis. *Ann Rheum Dis.* 2010;69(6):1208–1213.
30. Hou S, Hyland L, Ryan KW, Portner A, Doherty PC. Virus-specific CD8+ T-cell memory determined by clonal burst size. *Nature.* 1994;369(6482):652–654.
31. Kaech SM, Wherry EJ, Ahmed R. Effector and memory T-cell differentiation: implications for vaccine development. *Nat Rev Immunol.* 2002;2(4):251–262.
32. Muller S, et al. Activated CD4+ and CD8+ cytotoxic cells are present in increased numbers in the intestinal mucosa from patients with active inflammatory bowel disease. *Am J Pathol.* 1998;152(1):261–268.
33. Probert CS, Chott A, Saubermann LJ, Stevens AC, Balk SP, Blumberg RS. Prevalence of an ulcerative colitis-associated CD8+ T cell receptor beta-chain CDR3-region motif and its association with disease activity. *J Clin Immunol.* 2001;21(2):126–134.
34. Badovinac VP, Harty JT. Programming, demarcating, and manipulating CD8+ T-cell memory. *Immunol Rev.* 2006;211:67–80.
35. Morley M, et al. Genetic analysis of genome-wide variation in human gene expression. *Nature.* 2004; 430(7001):743–747.
36. Juelich T, et al. Interplay between chromatin remodeling and epigenetic changes during lineage-specific commitment to granzyme B expression. *J Immunol.* 2009;183(11):7063–7072.
37. Araki Y, et al. Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8+ T cells. *Immunity.* 2009;30(6):912–925.
38. Yamada T, Park CS, Mamontkin M, Lacorazza HD. Transcription factor ELF4 controls the proliferation



- tion and homing of CD8(+) T cells via the Kruppel-like factors KLF4 and KLF2. *Nat Immunol.* 2009; 10(6):618–626.
39. Akashi K, Reya T, Dalma-Weiszhausz D, Weissman IL. Lymphoid precursors. *Curr Opin Immunol.* 2000;12(2):144–150.
40. Kontoyiannis D, et al. Genetic dissection of the cellular pathways and signaling mechanisms in modeled tumor necrosis factor-induced Crohn's-like inflammatory bowel disease. *J Exp Med.* 2002; 196(12):1563–1574.
41. Nancey S, et al. CD8+ cytotoxic T cells induce relapsing colitis in normal mice. *Gastroenterology.* 2006;131(2):485–496.
42. Westendorf AM, et al. Autoimmune-mediated intestinal inflammation-impact and regulation of antigen-specific CD8+ T cells. *Gastroenterology.* 2006;131(2):510–524.
43. Steinhoff U, et al. Autoimmune intestinal pathology induced by hsp60-specific CD8 T cells. *Immunity.* 1999;11(3):349–358.
44. Cheroutre H. In IBD eight can come before four. *Gastroenterology.* 2006;131(2):667–670.
45. Lees CW, Barrett JC, Parkes M, Satsangi J. New IBD genetics: common pathways with other diseases [published online ahead of print February 7, 2011]. *Gut.* doi:gut.2009.199679.
46. Silverberg MS, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol.* 2005;19 suppl A:5–36.
47. Harvey RF, Bradshaw MJ. Measuring Crohn's disease activity. *Lancet.* 1980;1(8178):1134–1135.
48. Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. *Gut.* 1998;43(1):29–32.
49. Lyons PA, et al. Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. *BMC Genomics.* 2007;8:64.
50. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics.* 2002;18 suppl 1:S96–S104.
51. Irizarry RA, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249–264.
52. Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics.* 2009;25(3):415–416.
53. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118–127.
54. Simpson TI, Armstrong JD, Jarman AP. Merged consensus clustering to assess and improve class discovery with microarray data. *BMC Bioinformatics.* 2010;11:590.
55. Liu Y, Hayes DN, Nobel A, Marron JS. Statistical significance of clustering for high-dimension, low-sample data. *J Am Stat Assoc.* 2008;103(483):1281–1293.
56. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol.* 2004;3:Article3.