Crohn’s disease: Mucosal Immunology and Immune modulating therapy
Peters, Charlotte

Citation for published version (APA):
Gene expression profiles in peripheral blood as early biomarkers for response to induction therapy in Crohn’s disease

Abstract

Background
Monitoring of remission induction therapy for active Crohn’s disease (CD) is hampered by the lack of objective markers for therapy response. This also impedes development of new drugs since only subjective markers for therapeutic response can be used in clinical trials. Therefore, measurements of objective biomarkers are best applicable in peripheral blood. Biomarkers should differentiate therapy response from non-response, preferably early during therapy. This study aimed to identify biomarkers in peripheral blood distinguishing responders from non-responders at an early stage of remission induction therapy.

Methods
Whole blood RNA was obtained before therapy and after 3, 7, 14 and 56 days of remission induction from 10 responders and 10 non-responders. To identify a biomarker for therapy response in general and not for a specific treatment, both anti-TNF-α and prednisone remission induction strategies were included. Response was defined by ≥ 20% improvement in CDEIS at colonoscopy after 8 weeks of therapy. Differential gene expression profiles between responders and non-responders during induction therapy were determined using a linear model and a moderated F-test to compare baseline with all other time points ($P < 0.05$). Paired comparisons between baseline and any of the other time points were assessed using a moderated t test ($P < 0.01$).

Results
Before the start of therapy, 133 probesets were differentially expressed ($P < 0.001$) between responders and non-responders, including ATG16L2. Furthermore, after one week of induction therapy, 62 gene expression profiles could differentiate non-responders from responders ($P < 0.01$). The most differentially expressed genes comprise FLG2, CNOT4, ANK1, MICAL2, FHL2, IGF2BP2, S100A6 and MAPK8IP3. When taking the diagnostic accuracy into account, MICAL2, FLG2, IGF2BP2 and FHL2 remain interesting candidate biomarkers to predict response after one week of therapy.

Conclusions
Here we describe the first study on gene expression profiles from peripheral blood distinguishing non-responders from responders to CD remission induction therapy. These early gene expression profiles are candidates to be prospectively validated in an independent cohort as biomarkers for non-response to remission induction.
Introduction

Crohn’s disease (CD) is a chronic relapsing and remitting inflammation of the gastro-intestinal tract. When patients complain of diarrhoea, pain, fatigue and/or malaise, the treating physician should consider a flare of CD. However, symptoms may be due to other causes, such as irritable bowel syndrome (IBS) or enteric infections. Current guidelines underscore that clinicians can often poorly judge disease activity, therefore markers for objective disease are needed to confirm a flare of CD, before novel therapies are initiated. Objective markers include inflammatory markers in peripheral blood, such as C-reactive protein (CRP), faecal calprotectin and colonoscopy. Intensifying the current treatment or starting a novel therapy as remission induction is indicated when active disease is objectified. Nevertheless, monitoring of response to induction therapy is hampered by the lack of objective markers for therapy response. Normalization of CRP levels has been shown to correlate with long-term therapy response, however only 20%–25% of CD patients show increased levels of CRP, notwithstanding active intestinal inflammation. Subjective disease activity scores, such as the CD Activity Index (CDAI) have been considered to be “gold standard” to evaluate therapy efficacy in clinical trials. Nevertheless, in the SONIC trial steroid-free remission was achieved more often in patients with endoscopic inflammation before the start of infliximab and/or azathioprine, than for those who were treated based on an increase of the CDAI, indicating a considerable false positive rate of the CDAI when it comes to luminal disease activity. Furthermore, a recent post-hoc study showed that the CDAI, reflecting a subjective impression on patient’s clinical symptoms, did not accurately predict endoscopic remission or correlate with CRP levels. Recently, the EXTEND trial was the first to prospectively assess mucosal healing as primary end point. However, consecutive colonoscopies to visualize healing of the mucosa or amelioration of the CD Endoscopic Index of Severity (CDEIS) as an objective marker of therapy response may be applicable in clinical trials but not in daily clinical practice. Still, it is important to assess the outcome of an induction therapy as early as possible, to limit disease progression and prevent unnecessary side effects and costs of therapeutics. Therefore, we aimed to identify objective biomarkers in peripheral blood, reflecting early response to remission induction therapy. CD patients starting remission induction therapy for active disease were observed throughout the first 8 weeks of induction therapy. After 8 weeks of remission induction therapy the CDEIS was obtained during a second colonoscopy and patients were classified as responders or non-responders according to their individual changes in CDEIS. Consecutive blood samples were obtained before start and after 3, 7, 14 and 56 days. From these time series gene expression profiles with differential expression between responders and non-responders could be identified. In this study, we suggest novel peripheral blood biomarkers to identify non-responders early during induction therapy.
**Methods**

**Patients**
All CD patients scheduled for colonoscopy due to clinical symptoms suggesting active disease were asked for informed consent between 2009 and 2012. Patients were included when both active disease was observed during colonoscopy and/or imaging by MRI, and new remission induction therapy was initiated. Since we aimed to identify biomarkers for therapy response in general, and not only for one specific drug, patients were allowed to participate when starting anti-TNF-\(\alpha\) (infliximab or adalimumab) or prednisone.

**Tissue collection**
Before colonoscopy, blood for RNA isolation, plasma, CRP, leucocytes and albumin, the CDAI and IBDQ were obtained from every patient. During colonoscopy the CDEIS was obtained and biopsies were collected from inflamed and non-inflamed regions of the intestine. After the start of induction therapy at 3, 7, and 14 days blood for RNA isolation, plasma, CRP, leucocytes and albumin was obtained. After 8 weeks patients again completed the CDAI and IBDQ, and donated blood. Furthermore a second colonoscopy was performed after 8 weeks of induction therapy including scoring of the CDEIS and collection of biopsies of inflamed and non-inflamed regions, when applicable.

**Identification of top 10 responders and non-responders**
Response to the induction therapy was determined primarily by the changes in CDEIS and secondly by CRP after 8 weeks of therapy. Furthermore, CDAI and IBDQ scores were obtained before and after 8 weeks. The CDEIS was considered to be “gold standard”, however patients classified as “responders” should show at least 20% improvement on all activity scores. Patients were classified to be “non-responders” when CD activity scores remained equal or deteriorated after 8 weeks of therapy. Individual fold changes were calculated by dividing patient’s activity score after 8 weeks (e.g. CDEIS) by the corresponding score at baseline. Thereby fold changes above 1 reflect deterioration of disease when focussing on CDEIS, CRP or CDAI and fold changes below 1 reflect amelioration of disease. Contrariwise, the IBDQ score increases when quality of life improves, hence fold changes above 1 reflect improvement of IBDQ and fold changes below 1 reflect declined quality of life.

**RNA extraction**
In PAXgene tubes containing RNA stabilizing solution (Qiagen, Venlo, the Netherlands) 2,5 ml of peripheral blood was collected. PAXgene blood tubes were incubated at room temperature for 24 hours to ensure lysis of blood cells. Total RNA should be stable for up to 3 days at 18–25°C, which makes this an applicable tool for clinical practice. After 24 hours at room temperature, PAXgene tubes were stored in -20°C for 24 hours to ensure the tubes would not burst when stored directly in the -80°C, followed by storage in a -80°C freezer up to the isolation of total RNA. Total RNA was isolated with PAXgene Blood miRNA Kits (Qiagen, Venlo, the Netherlands) using the automated method, according to the manufacturer’s protocol. In brief, PAXgene tubes were centrifuged after which the nucleic acid pellet was washed and resuspended into...
optimized buffers together with proteinase K for protein digestion. Next, additional centrifugation through a column was performed to homogenize the cell lysate and remove residual debris. For binding conditions, ethanol was added and during a brief centrifugation in a spin column, only RNA bound the silica membrane as contaminants are lost. To exclude remnant contaminants, the column was treated with DNase I and washed several times, after which RNA was eluted and heat-denatured. Quality of the RNA was ascertained by analyses of all samples with the 2100 Bioanalyzer Instrument (Agilent Technologies, Santa Clara, California, United States). The RNA Integrity Number (RIN) and RNA concentration was determined by using Agilent RNA 6000 Nano Kits (Agilent Technologies, Santa Clara, California, United States). Samples which had a RIN of at least 8.0 were used for microarray analysis.

Microarray experiment and pre-processing
Microarray labelling and hybridization were performed at the Microarray Department, University of Amsterdam according to the manufactures instructions. To avoid confounding with the time point of blood collection, samples were randomly allocated to the Affymetrix Human Genome U133 Plus 2.0 Arrays (Affymetrix), which comprised 54,675 probe sets covering the whole genome. Analyses were carried out with Bioconductor packages in the statistical software package R (version 3.0.0). Raw data were extracted from the CEL files using the package affy. Data was normalized and summarized at the probeset level using robust multiarray averaging (RMA) with default settings (function rma, package affy). In total 21,215 probesets without presence call (function mas5calls, package affy) on at least one array were excluded. The package ArrayQualityMetrics was used to assess the quality of the microarray data both before and after normalization. Based on the quality control, one array at the 8-week time point was left out and remaining arrays were renormalized. Probesets were annotated using the package hgu133plus2.db

Statistical analyses
Genes that respond differently to medication in responders versus non-responders were determined using a linear model with a fixed main effect for group (responder/non-responder), and nested interactions of group with patient and time points. Significant differences in response were determined using the appropriate contrast and a moderated F-test to compare time point 0 versus all other time points ($P < 0.05$, package limma). Significant differences in response for the pairwise comparisons between time point 0 and any of the other time points were assessed using a moderated t test ($P < 0.01$, package limma). Genes that are differentially expressed between responders and non-responders at time point 0 were determined by extracting the appropriate contrast from a linear model with a fixed main effect for each combination of group and time point ($P < 0.001$, moderated t-test). Resulting p-values were corrected for multiple testing using the Benjamini-Hochberg False Discovery Rate adjustment, such that a corrected $P$-value $< 0.05$ controls the expected false discovery rate to be less than 5%. Diagnostic accuracy of selected genes was evaluated by determining the area under the receiver operating characteristic curve (AUC, pROC package). De Long’s method for calculating AUC variance was used for the calculation of AUC 95% confidence intervals.
Ethical considerations
All patients gave informed consent before participation. The study protocol was approved by the Academic Medical Center’s ethical committee.

Results

Patient characteristics
As depicted in Figure 1, seventy-seven CD patients were recruited and gave informed consent. All patients had consulted their treating physician with symptoms of active disease. All patients completed v and IBDQ questionnaires and CRP and CDEIS were obtained. Only 70.1% of patients showed signs of active inflammation at colonoscopy (CDEIS > 3) or MRI. Twenty-nine patients started a novel therapy for remission induction such as prednisone or anti-TNF-α. The other 25 active patients were excluded due to other treatment strategies, including experimental treatment in clinical trials (3 patients), IBD-U (5 patients), need for surgery due to stricturing disease (2 patients), or the start of thiopurines or methotrexate (7 patients), or dose escalation of current treatment (2 patients). Six patients withdrew their consent after the first colonoscopy due to the prospect of the second colonoscopy that would be performed after 8 weeks of therapy. From the 29 patients starting remission induction, 7 patients started infliximab, 13 patients started adalimumab, and 9 patients started corticosteroids. Four patients had to discontinue therapy due to adverse events, including one patient suffering from serum sickness after the second infliximab infusion. After 8 weeks of remission induction therapy, all disease activity scores were repeated (CDAI, IBDQ, CDEIS and CRP) and individual fold changes were calculated. CDEIS was considered to be ‘gold standard’ reflecting response to therapy most accurately. Five patients showed some response to induction therapy however not fulfilling the criteria of response (improvement in all activity scores) or non-response (deterioration or no change in all activity scores). Therefore these patients were excluded from further analyses.

RNA was isolated from peripheral blood, which was obtained at 5 time points during induction therapy (before and at 3, 7, 14, and 56 days of therapy) from 10 responders and 10 non-responders. Most samples showed sufficient quality of the RNA according to the RIN, however RNA from one non-responder patient did not meet the RIN criteria at several time points and was therefore excluded from analyses. This led to a final cohort of 10 responders and 9 non-responders. No significant differences were present between the baseline characteristics of the two groups (Table 1). Individual fold changes in CD activity scores of the included patients are depicted in Figure 2. Patients were classified according to their fold change in CDEIS thus all responders showed improvement of the CDEIS, whereas non-responders remained equally or deteriorated. CRP levels at baseline varied widely, as shown in Table 1, which is one of the reasons that CRP fold changes did not differentiate responders from non-responders like the CDEIS. Similarly, the median CDAI and IBDQ were different between responders and non-responders (defined by CDEIS), however these subjective questionnaires did not differentiate the two groups like the CDEIS did.
Biomarkers for therapy response in Crohn’s patients

Gene expression between responders and non-responders before the start of therapy
At baseline, 133 probesets showed differential expression between responders and non-responders \( (P < 0.001) \). Table 2 lists the 10 most differentially expressed genes before the start of therapy, which differentiate responders from non-responders by at least 1.5 fold change.

![Flow chart of patient inclusion](image)

**Figure 1 | Flow chart of patient inclusion**
Figure 2 | Disease activity score fold changes in responders and non-responders. Division of the activity score after 8 weeks by the score at baseline resulted in fold changes. Thereby fold changes represent individual changes after 8 weeks of therapy. Patients were classified as responders when they showed CDEIS fold change < 1, whereas patients were classified non-responders when CDEIS fold change was ≥ 1. CDEIS = Crohn’s Disease Endoscopic Index of Severity; CRP = C-Reactive Protein; IBDQ = Inflammatory Bowel Disease Questionnaire; CDAI = Crohn’s Disease Activity Index.

Table 1 | Baseline patient characteristics of responders and non-responders

<table>
<thead>
<tr>
<th></th>
<th>Responders</th>
<th>Non-Responders</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at inclusion (Mean (SD))</td>
<td>38.8 (9.5)</td>
<td>36.4 (13.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Gender (n Female)</td>
<td>3</td>
<td>5</td>
<td>0.26</td>
</tr>
<tr>
<td>Disease duration (months)</td>
<td>99.5 (54–208)</td>
<td>121 (22–215)</td>
<td>1.00</td>
</tr>
<tr>
<td>CDEIS (median, IQR)</td>
<td>9.4 (7.7–14.9)</td>
<td>8.4 (3.2–18.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>CRP (median, IQR)</td>
<td>8.9 (3.7–24.1)</td>
<td>5.7 (2.8–31.5)</td>
<td>1.00</td>
</tr>
<tr>
<td>Leucocytes (10E9/L) (median, IQR)</td>
<td>8.9 (7.1–10.4)</td>
<td>7.7 (6.1–8.8)</td>
<td>0.17</td>
</tr>
<tr>
<td>ESR (median, IQR)</td>
<td>26 (6–32)</td>
<td>12 (4–32.5)</td>
<td>1.00</td>
</tr>
<tr>
<td>CDAI (median, IQR)</td>
<td>182 (114–247)</td>
<td>169 (112–238)</td>
<td>1.00</td>
</tr>
<tr>
<td>IBDQ (median, IQR)</td>
<td>162 (133–170)</td>
<td>150 (125–168)</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Induction therapy (n)
- Anti-TNF-α (ADA/IFX) | 7 (6/1) | 5 (3/2) | 0.66 |
- Prednisone            | 3      | 4      |      |

CDEIS = Crohn’s Disease Endoscopic Index of Severity; CRP = C-Reactive Protein; ESR = Erythrocyte sedimentation rate; CDAI = Crohn’s Disease Activity Index; IBDQ = Inflammatory Bowel Disease Questionnaire; Anti-TNF-α = anti-tumor necrosis factor-alpha; ADA = adalimumab; IFX = infliximab.
Biomarkers for therapy response in Crohn’s patients

<table>
<thead>
<tr>
<th>Genes</th>
<th>Fold change</th>
<th>P-value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATG16L2</td>
<td>1.64</td>
<td>9.09E-06</td>
<td>0.17</td>
</tr>
<tr>
<td>ARHGAP9</td>
<td>1.59</td>
<td>6.79E-05</td>
<td>0.21</td>
</tr>
<tr>
<td>MAP3K2</td>
<td>1.68</td>
<td>7.12E-05</td>
<td>0.21</td>
</tr>
<tr>
<td>GMFG</td>
<td>1.55</td>
<td>7.39E-05</td>
<td>0.21</td>
</tr>
<tr>
<td>RICTOR</td>
<td>1.87</td>
<td>8.09E-05</td>
<td>0.21</td>
</tr>
<tr>
<td>ARHGDIB</td>
<td>1.65</td>
<td>9.04E-05</td>
<td>0.21</td>
</tr>
<tr>
<td>ANK1</td>
<td>-1.67</td>
<td>1.69E-04</td>
<td>0.21</td>
</tr>
<tr>
<td>SLC44A2</td>
<td>1.52</td>
<td>2.04E-04</td>
<td>0.21</td>
</tr>
<tr>
<td>FOSL2</td>
<td>1.61</td>
<td>2.07E-04</td>
<td>0.21</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>2.00</td>
<td>2.12E-04</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Fold change: fold changes (responders/non-responders); FDR: false discovery rate.

Genes differentially expressed between responders and non-responders during induction therapy

Testing for differences in gene expression between responders and non-responders during the first eight weeks of remission induction therapy revealed 705 differentially expressed probesets ($P < 0.05$). Figure 3 depicts a heatmap of the differences in fold changes between responders and non-responders of the 62 differentially expressed genes ($P < 0.01$).

Of the 62 differentially expressed genes, 15 genes were differentially expressed between responders and non-responders after 3, 7 and 14 days of therapy, 47 genes were differentially expressed at two consecutive time points: 14 genes at 3 and 7 days and 33 genes at 7 and 14 days (Figure 4). To illustrate the individual profiles of gene expression and the fold changes during treatment, the most differentially expressed genes are depicted per time frame (Figure 5).
Figure 3 | Heatmap of the 62 differentially expressed interaction genes between responders and non-responders ($P < 0.01$). To calculate fold changes, baseline differences were set to zero and differences in fold changes after 3, 7 and 14 days between responders and non-responders are shown. The upper panel shows increased expression in responders as compared to non-responders. The lower panel reflects a decreased expression in responders versus non-responders.
Biomarkers for therapy response in Crohn’s patients

Figure 4 | Venn diagram of the 62 differentially expressed interaction genes between responders and non-responders (P < 0.01). This Venn diagram shows differentially expressed interaction genes between responders and non-responders (P < 0.01). Biomarkers relevant for clinical practice should have a similar profile in at least two consecutive time points. Therefore three groups were differentiated: genes differentially expressed after 3 and 7 days; genes differentially expressed after 3, 7 and 14 days; genes differentially expressed after 7 and 14 days.

Fillagrin-2 (FLG2) and carbon catabolite repression (CCR) 4-negative on TATA-less transcription complex, subunit 4 (CNOT-4) are most differentially expressed after 3 and 7 days of therapy (P = 5.2E-04 and P = 6.4E-04, respectively). Non-responders start with higher expression levels of FLG-2 at baseline (Figure 4a) and can be identified by a decrease in fold change after 3 and 7 days of therapy (Figure 4b). On the contrary, non-responders start at lower expression of CNOT4 and increase after 3 and 7 days of therapy. After 8 weeks of therapy (56 days) no difference is present (Figure 5a).

After 3, 7, and 14 days ankyrin 1, erythrocytic (ANK1) and microtubule associated mono-oxygenase, calponin and LIM domain containing 2 (MICAL2) were most differentially expressed between responders and non-responders (P = 2.2E-04 and P = 6.7E-04, respectively). As depicted in Figure 5b, this was partly due to a significant difference at baseline (both genes were expressed at higher levels in non-responders). Furthermore, four-and-a-half LIM domain protein 2 (FHL2) and Insulin-like growth factor 2 mRNA-binding protein 2 or p62 (IGF2BP2) show differential profiles. Non-responders mostly start at higher levels than responders and then their expression of FHL2 and IGF2BP2 is lowered (P = 6.8E-04 and P = 3.7E-03, respectively) at 3, 7, and 14 days. No difference is present after 8 weeks of therapy (Figure 5c).

Calcyclin (S100A6) is differentially expressed after 7 and 14 days of therapy. Non-responders show lower expression at baseline and increase their S100A6 expression after 7 and 14 days, as compared to responders, who lower their expression (P = 9.0E-04). Also mitogen-activated protein kinase 8 interacting protein 3 (MAPK8IP3) increases in non-responders after 7 and 14 days of therapy (P = 9.9E-04). These differences were still present after 8 weeks of therapy (Figure 5d).

Since FLG2 is a protein-coding transcript, we measured levels of FLG2 in plasma samples of responders and non-responders. In analogy with decreased transcript levels after 7 days of therapy, non-responders show a decrease in their FLG2 protein levels in plasma (P < 0.02). Responders show no differential levels of FLG2 in plasma after 7 days (data not shown).
Sensitivity and specificity of differentially expressed genes
To identify a biomarker with diagnostic accuracy, the true positive rate (sensitivity) and false positive rate (specificity) should be taken into account. The area under the curve (AUC) is a measure of how well a parameter distinguishes non-responders from responders. The top 15 AUC’s of the 62 differentially expressed genes ($P<0.01$) are listed in Table 3. At 7 days of therapy, MICAL2 appears to have the best diagnostic accuracy with an AUC of 0.96 (0.88 - 1.00), followed by FLG2 (AUC of 0.95 (0.84 - 1.00)), IGF2BP2 (AUC of 0.93 (0.80 0 1.00)), and FHL2 (AUC of 0.91 (0.76 -1.00)). Supplementary Figure 1 shows Receiver Operating Characteristic (ROC) curves of the sensitivity plotted in function of 1 minus the specificity for different cut-off points: 3 days, 7 days, and 14 days.

Table 3 | Fifteen most accurate diagnostic genes to identify non-responders after 7 days of therapy

<table>
<thead>
<tr>
<th>Genes</th>
<th>$P$-value</th>
<th>AUC day 3 (95% CI)</th>
<th>AUC day 7 (95% CI)</th>
<th>AUC day 14 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MICAL2</td>
<td>6.7E-04</td>
<td>0.89 (0.73 -1.00)</td>
<td>0.96 (0.88 -1.00)</td>
<td>0.83 (0.60 -1.00)</td>
</tr>
<tr>
<td>FLG2</td>
<td>5.2E-04</td>
<td>0.83 (0.60 -1.00)</td>
<td>0.95 (0.84 -1.00)</td>
<td>0.83 (0.63 -1.00)</td>
</tr>
<tr>
<td>ZC3H3</td>
<td>1.1E-02</td>
<td>0.87 (0.66 -1.00)</td>
<td>0.95 (0.83 -1.00)</td>
<td>0.83 (0.63 -1.00)</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>3.7E-03</td>
<td>0.81 (0.55 -1.00)</td>
<td>0.93 (0.80 -1.00)</td>
<td>0.83 (0.61 -1.00)</td>
</tr>
<tr>
<td>FHL2</td>
<td>6.8E-04</td>
<td>0.90 (0.71 -1.00)</td>
<td>0.91 (0.76 -1.00)</td>
<td>0.90 (0.76 -1.00)</td>
</tr>
<tr>
<td>ZFAS1</td>
<td>1.4E-02</td>
<td>0.89 (0.72 -1.00)</td>
<td>0.91 (0.76 -1.00)</td>
<td>0.90 (0.74 -1.00)</td>
</tr>
<tr>
<td>SCAF11</td>
<td>2.2E-02</td>
<td>0.84 (0.63 -1.00)</td>
<td>0.91 (0.75 -1.00)</td>
<td>0.75 (0.48 -1.00)</td>
</tr>
<tr>
<td>LINC00273</td>
<td>6.7E-04</td>
<td>0.90 (0.75 -1.00)</td>
<td>0.91 (0.73 -1.00)</td>
<td>0.81 (0.61 -1.00)</td>
</tr>
<tr>
<td>ARHGDIB</td>
<td>2.3E-03</td>
<td>0.78 (0.51 -1.00)</td>
<td>0.89 (0.72 -1.00)</td>
<td>0.79 (0.55 -1.00)</td>
</tr>
<tr>
<td>FECH</td>
<td>1.7E-03</td>
<td>0.90 (0.75 -1.00)</td>
<td>0.89 (0.71 -1.00)</td>
<td>0.75 (0.50 -1.00)</td>
</tr>
<tr>
<td>ANK1</td>
<td>3.1E-03</td>
<td>0.84 (0.64 -1.00)</td>
<td>0.89 (0.72 -1.00)</td>
<td>0.79 (0.56 -1.00)</td>
</tr>
<tr>
<td>CNOT4</td>
<td>6.4E-04</td>
<td>0.90 (0.75 -1.00)</td>
<td>0.89 (0.73 -1.00)</td>
<td>0.64 (0.36 -1.00)</td>
</tr>
<tr>
<td>FECH</td>
<td>3.2E-02</td>
<td>0.84 (0.63 -1.00)</td>
<td>0.89 (0.68 -1.00)</td>
<td>0.75 (0.49 -1.00)</td>
</tr>
<tr>
<td>RPL35A</td>
<td>1.9E-02</td>
<td>0.90 (0.74 -1.00)</td>
<td>0.89 (0.71 -1.00)</td>
<td>0.84 (0.61 -1.00)</td>
</tr>
<tr>
<td>ADIPOR1</td>
<td>5.0E-03</td>
<td>0.81 (0.55 -1.00)</td>
<td>0.89 (0.70 -1.00)</td>
<td>0.83 (0.60 -1.00)</td>
</tr>
</tbody>
</table>

AUC: area under the ROC curve, after 3, 7 or 14 days of therapy; 95% CI: 95% confidence interval.

Figure 5 (as shown on the right) | Expression profiles and fold change profiles of most differentially expressed genes. a. Expression profile and fold changes of FLG2 and CNOT4 which are most differentially expressed after 3 and 7 days of therapy ($P = 5.2E-04$ and $P = 6.4E-04$, respectively). b. Expression profile and fold changes of ANK-1 and MICAL2 which are most differentially expressed after 3, 7 and 14 days of therapy between responders and non-responders ($P=2.2E-04$ and $P = 6.7E-04$, respectively). c. Expression profile and fold changes of FHL2 and IGF2BP2 which are differentially expressed between responders and non-responders after 3, 7, and 14 days of therapy ($P=6.8E-04$ and $P = 3.7E-03$, respectively). d. Expression profile and fold changes of S100A6 and MAPK8IP3 which are differentially expressed at 7 and 14 days of therapy between responders and non-responders. ($P=9.0E-03$ and $P = 9.9E-03$, respectively).
Biomarkers for therapy response in Crohn’s patients

Gene expression

**FLG2 (1569410_at)**

- Log2(expression)
- Time (days)

**CNOT4 (210867_at)**

- Log2(expression)
- Time (days)

**ANK1 (208353_x_at)**

- Log2(fold change)
- Time (days)

**MICAL2 (212473_s_at)**

- Log2(expression)
- Time (days)

Fold changes

**FLG2 (1569410_at)**

- Log2(fold change)
- Time (days)

**CNOT4 (210867_at)**

- Log2(fold change)
- Time (days)

**ANK1 (208353_x_at)**

- Log2(fold change)
- Time (days)

**MICAL2 (212473_s_at)**

- Log2(fold change)
- Time (days)
**Gene expression**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression (log2)</th>
<th>Fold Change (log2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHL2</strong> (202949_s_at)</td>
<td>-0.5 to 0.0</td>
<td>8.0 to 9.0</td>
</tr>
<tr>
<td><strong>IGF2BP2</strong> (218847_at)</td>
<td>0.5 to 1.0</td>
<td>6.2 to 6.6</td>
</tr>
<tr>
<td><strong>S100A6</strong> (217728_at)</td>
<td>0.0 to 0.5</td>
<td>0.25 to 0.50</td>
</tr>
<tr>
<td><strong>MAPK8IP3</strong> (213178_s_at)</td>
<td>0.0 to 0.5</td>
<td>-0.25 to 0.00</td>
</tr>
</tbody>
</table>

**Fold changes**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold Change (log2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHL2</strong> (202949_s_at)</td>
<td>8.0 to 9.0</td>
</tr>
<tr>
<td><strong>IGF2BP2</strong> (218847_at)</td>
<td>6.2 to 6.6</td>
</tr>
<tr>
<td><strong>S100A6</strong> (217728_at)</td>
<td>0.25 to 0.50</td>
</tr>
<tr>
<td><strong>MAPK8IP3</strong> (213178_s_at)</td>
<td>-0.25 to 0.00</td>
</tr>
</tbody>
</table>
Discussion

In the present study we investigated the value of gene expression profiles in peripheral blood as biomarkers for therapy response to remission induction therapy. Early after the start of therapy individual genes show fold change profiles that differentiate responders from non-responders. These fold change differences were often accompanied by differential expression between responders and non-responders before the start of therapy, but pre-treatment differential gene expression was not able to predict non-response to induction therapy, although this may be due to the limited number of patients in this cohort. Subsequently, after 7 days of induction therapy MICAL2, FLG2, IGF2BP2, and FHL2 seem to be promising biomarkers to early identify patients with non-response.

This is the first study focussing on gene expression changes in peripheral blood during CD induction therapy. At baseline a several genes were differentially expressed between responders and non-responders. Among them is ATG16L2 that is expression at higher levels in responders versus non-responders before the start of therapy. As a homologue of ATG16L1, it was recently shown to be a novel candidate for CD in an Asian population. Furthermore, in pancreas homeostasis, ATG16L2 knockdown models showed increased susceptibility to ER stress. The precise function of ATG16L2 is not elucidated, but it has been suggested to function as an adaptor protein involved in autophagy completion. ATG16L1 is a well-known mediator of autophagy processes and has been associated with CD in (meta-analyses) of genome wide association studies. When correcting for multiple testing, ATG16L2 expression levels at baseline would result in 17% false positives. No individual gene or combination of genes at baseline was able to predict response to induction therapy with 100% accuracy. Previously, microarray studies have been performed on biopsies to predict (non)response to infliximab treatment. Class prediction analysis revealed a top 5 differentially expressed genes predicting response with 100% accuracy: TNFAIP6, S100A8, IL-11, G0S2, and S100A9. These genes might reflect the patient’s inflammatory state rather than changes due to (non)response to therapy. The top 5 genes identified in biopsies were not differentially expressed in the present cohort focussing on peripheral blood samples of patients before or early during induction therapy. Biomarkers in peripheral blood would provide an attractive method for early diagnosis of non-response. Early transcriptional changes have been described in perturbation–response experiments, suggesting that most attention should be given to early time points since that is when the transcriptional response occurs. Nevertheless, for capturing gene expression patterns, a late time point is essential in order to differentiate between transient and sustained responses. Therefore, individual fold changes at early time points may predict response more accurately than baseline levels at itself. To include sustained response, gene expression after 8 weeks was included in this study.

After 3, 7, and 14 days of induction therapy, 62 genes showed differential expression profiles of which two genes belong to the same family. Genes differentially expressed in peripheral blood between responders and non-responders, are not necessarily expressed in the intestine of CD patients. These genes may reflect several mechanisms, such as general inflammatory state differences, metabolic changes or mechanisms of therapy response.
As early as 3 and 7 days after start of therapy, FLG2 and CNOT4 are differentially expressed compared to baseline. FLG2 proteins are important in the maintenance of epithelial homeostasis and barrier functions. Genetic FLG2 variations have been associated with persistent atopic dermatitis. The RING finger protein CNOT4 is a component of the CCR4-NOT complex, which is evolutionarily conserved. The CNOT4 complex is a regulator of transcription and is important for multiple cellular functions.

Including all three early time points (3, 7, and 14 days) 15 genes were differentially expressed between responders and non-responders, including a top 4: ANK1, MICAL2, FHL2 and IGF2BP2. These were all differentially expressed at baseline and thereby consecutive fold changes after 3, 7, and 14 days discriminated response from non-response.

The ANK1 protein belongs to the ankyrin family which contains structural adaptor proteins that form key components of the spectrin-actin cytoskeleton. They consist of three domains, an N-terminal binding, a spectrin binding and a C terminal regulatory domain. They link transmembrane proteins to the cytoskeleton and play a role in cell motility, activation and proliferation. ANK1 is involved in hereditary spherocytosis, where loss of membrane surface area leads to impaired deformability due to defects in the membrane proteins ankyrin.

MICAL2 belongs to a family of signal transduction proteins with domains that interact with the cytoskeleton. Similar to FHL2, MICAL2 contains a LIM domain that can perform various functions. The LIM domain is a cysteine-rich zinc-binding protein-protein interaction motif. It participates in cell survival, transcription and signal transduction. Recently MICALS and their antagonist have been shown to regulate actin disassembly and assembly respectively. MICAL enzymes perform redox enzymatic reactions, which are utilized by macrophages during cellular activation by stimulating MsrB1 expression, as a part of innate immunity.

FHL2 belongs to a large family of LIM domain-containing proteins that are involved in modulation of several signalling pathways, including cell differentiation and growth control. In blood, FHL-2 is expressed in mononuclear cells and platelets (source: www.genecards.org). It has been described to play a role in the immune response by enhancement of the cellular innate immune response to influenza A virus infection. FHL2 overexpression increases IL-6 secretion in skeletal myoblasts. FHL2 upregulation in human liver specimens showed significant association with increasing inflammation score and cirrhosis. Furthermore, stimulated by transforming growth factor (TGF)-β1, FHL2 was shown to be highly expressed in primary and metastatic colon cancer but not in normal tissues. FHL2 appeared critical for cancer cell adhesion to extracellular matrix, migration and invasion by stimulating vimentin and matrix metalloproteinase-9 (MMP9) expression and causing a loss of E-cadherin. Hence, FHL2 might be an important mediator for invasion and/or metastasis of colon cancer. Finally, FHL-2 regulates CCL-19 induced dendritic cell migration. Stimulation with TNF- downregulates FHL2 in synovial fibroblasts isolated from different RA-patients (personal communication Prof. D. Baeten). IGF2BP2 functions by binding to the 5' UTR of the insulin-like growth factor 2 (IGF2) mRNA and thereby regulates its translation. This gene was differentially expressed between responders and non-responders after 7 days of therapy. IGF2 is a mitogen for various cell types and indispensable for normal embryonic growth. In blood cells IGF2BP2
is expressed in mononuclear cells and platelets (source: www.genecards.org). Recently IGF2BP2 was suggested as a new prognostic marker in a hepatocellular carcinoma as it induces a steatotic phenotype when overexpressed in mouse livers. In this HCC model, IGF2BP2 leads to IGF2-independent anti-apoptotic actions, facilitated via phosphorylation of ERK1/2. Potentially, the decrease of IGF2BP2 in non-responders could reflect a different state of apoptosis early during therapy.

After 7 and 14 days of therapy 33 genes were differentially expressed between responders and non-responders. Both S100A6 and MAPK8IP3 are expressed at lower levels in non-responders (baseline) and increase during remission induction, whereas responders show a reversed profile. MAPK8IP3 (or JSAP1) functions as a scaffold protein in the JNK cascade modulating cell migration. In addition, it was shown to associate with TLR4 resulting in LPS mediated JNK activation.

S100A6, also known as calcyclin, codes for proteins of the S100 family. Also FLG2, which discriminated early response at 3 and 7 days, is part of the S100 family. In 1965, these proteins were named after their ability to dissolve in a 100% ammonium sulphate solution. S100 proteins bind calcium leading to several potential biological tasks, e.g. regulation of protein phosphorylation, modulation of enzyme activity, promotion of cell growth, differentiation or apoptosis, participation in calcium homeostasis, preservation of cell shape and motility, regulation of coagulation, and induction of pro- and anti-inflammatory responses. The S100 family includes calprotectin (S100A8/A9), which is generally accepted as a faecal biomarker to identify intestinal inflammation in IBD and especially to differentiate IBD from irritable bowel syndrome. After anti-TNF-α induction therapy normalization of faecal calprotectin predicts sustained clinical remission in the majority of patients. A decrease in faecal calprotectin has been seen as early as after 2 weeks, however the variation is large with increased levels before the next anti-TNF-α infusion. In serum, calprotectin was recently identified as a complementary marker next to CRP and faecal calprotectin to predict relapse after infliximab withdrawal, however serum calprotectin did not correlate with CDEIS. S100A6 can function via interaction with intracellular target proteins indirectly regulating cell proliferation, apoptosis, and motility. Studies with rat neonatal cardiac myocytes demonstrate that S100A6 is induced by TNF-α via an NF-kappaB-dependent mechanism, serving a role in homeostasis to limit TNF-α-induced apoptosis by regulating p53 phosphorylation. These two family members of calprotectin (FLG2 and S100A6) may serve as surrogate biomarkers for early therapy response.

A limitation of this study is the lack of an independent validation cohort that can truly establish the value of these newly found biomarkers of early response to induction treatment. Nevertheless, since we investigated time series and thereby identified gene expression profiles per patient, these potential biomarkers are of interest. Moreover, since response was defined by the most objective score available, the CDEIS and patient’s individual changes in CDEIS, we feel that these patients represent true response and non-response. In daily clinical practise obtaining blood after one week of induction therapy for RNA isolation, from e.g. Paxgene collection tubes, that conserve RNA at room temperature overnight, should be easily feasible. Furthermore, performing a qPCR on these biomarkers is a technique which is available in most hospitals.
In conclusion, 62 genes were differentially expressed in the first two weeks of induction therapy differentiating responders from non-responders. The best individual predictive value of increased fold changes after one week of treatment was observed for MICAL2 and FLG2, followed by IGF2BP2 and FHL2. Nevertheless, these biomarkers will only be feasible in clinical practice after prospective validation in an independent cohort.

Acknowledgments
We are grateful for the help and support of the IBD nurses: Florien Toxopeüs, Mirjam van der Spek, Maria de Jong and Suzan Stolte and treating physician’s: Mark Löwenberg, Gijs van den Brink, Geert D’Haens and Svend Rietdijk. Furthermore, we would like to thank Richard Volckman and Jan Koster for support with performance of the microarrays.
Supplementary figures

**Supplementary figure 1 | ROC curves for diagnostic accuracy.** a. ROC curves of FLG2, CNOT4, MAPK8IP3, and S100A6. Per gene, the area under the curve (AUC) after 3, 7 and 14 days are shown. b. ROC curves of ANK1, MICAL2, FHL2 and IGF2BP2. Per gene, the area under the curve (AUC) after 3, 7 and 14 days are shown.
References


