

GENES SPECIFICALLY REGULATED IN INFLAMMATORY BOWEL DISEASE

Philip Zimmermann and Reto Baumann

Nebion AG, Switzerland

© NEBION AG, July 2015.

Inflammatory bowel disease (IBD) is a heterogeneous group of recurring or chronic long-term inflammatory conditions of the gastrointestinal (GI) system. The two main types of IBD are known as Crohn's disease (CD) and ulcerative colitis (UC). In this study, we used GENEVESTIGATOR tools to identify genes specifically regulated in CD and UC compared to over 3,000 other experimental conditions and diseases. We selected the most specific hits and characterized their regulation across diseases and tissues. With our analysis, we confirmed existing associations but also discovered novel genes previously uncharacterized and specifically regulated in response to IBD.

INTRODUCTION

IBD manifests itself most prominently in the colon and the small intestine, but other areas such as the esophagus might also be affected to a variable degree. It is assumed that a complex combination of environmental, genetic and microbial factors is the underlying cause of IBD, which shares common clinical features with typical autoimmune diseases (Baumgart and Carding, 2007; Xavier and Podolsky, 2007). Major differences between UC and CD are found in the location and the nature of the inflammatory lesions within the gastrointestinal tract. Both pathologies typically present with similar clinical symptoms such as abdominal pain, severe cramping, diarrhea, vomiting, weight loss and anemia in addition to inflammation-associated complaints that can affect the entire body. In extreme cases, surgery might be necessary to remove damaged parts of the intestines, subsequently often requiring colostomy or ileostomy. Therefore these diseases have a very negative impact on the quality of life of patients.

Treatment of IBD usually combines the actions of anti-inflammatory and immunosuppressive drugs to control the inflammatory process and thereby limit tissue damage (Baumgart and Carding, 2007; Baumgart and Sandborn, 2012; Ordas et al., 2012). Modern clinical treatments often focus on the inhibition of the pro-inflammatory cytokine TNF-alpha by delivering monoclonal antibodies targeted against this protein. These antibodies (e.g. Infliximab, Adalimumab, Golimumab and Certolizumab), marketed by different pharmaceutical companies, are uniquely targeting TNF-alpha and can also be used to treat other chronic inflammatory diseases such as psoriasis and rheumatoid arthritis.

It is of pivotal importance to understand inflammatory processes at the cellular and molecular level to design effective and specifically tailored therapies. Chronic tissue inflammation is also generally known to be an important risk factor for the development of various types of cancer, e.g. by promoting the generation of genetic mutations or stimulating cell proliferation (Rakoff-Nahoum, 2006). In fact, more than 20% of IBD patients develop colitis-associated cancer (CAC) within 30 years of disease onset and more than 50% of them will die from this aggressive type of cancer (Feagins et al., 2009; Lakatos and Lakatos, 2008).

IDENTIFICATION OF GENES SPECIFICALLY REGULATED IN IBD

The goal of this study was to identify genes specifically regulated in CD and UC using GENEVESTIGATOR. Prior to the analysis the NEBION curation team searched public repositories for relevant experiments and extensively curated 16 microarray studies of IBD, counting for 673 samples. Additionally, two experiments involving treatment with Infliximab (a drug used to treat IBD) and several colorectal cancer experiments were also curated. The so enriched content allowed measuring biological effects more robustly using multiple sample comparisons of diseased versus healthy or treated versus controls.

To identify IBD specific genes, we used GENEVESTIGATOR (Hruz et al., 2008) and selected a compendium of 49,191 samples profiled on the Affymetrix Human 133 Plus 2 platform. We then iteratively applied the Perturbations tool from the GENE SEARCH toolset to identify first the top 50 IBD specific up-regulated probesets. In a second step, we looked for both up- and down-regulated genes, keeping the top 25 identified probesets, respectively. After removing redundant probesets (targeting the same transcript), our search yielded a set of 21 up-regulated and 21 down-regulated genes specific for IBD. Additionally we found some of them having either positive or negative regulation capabilities also in colorectal cancer (CC), although we did not select CC among the target conditions. For almost all other diseases, drugs and other types of perturbations, no regulation was observed for the selected genes.

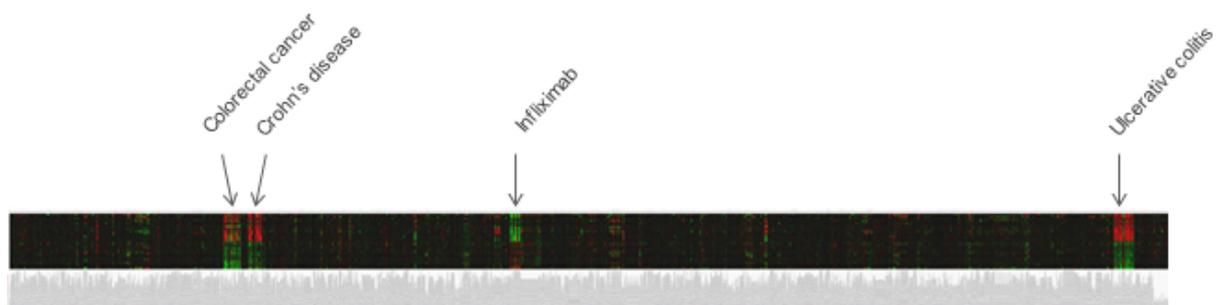


Figure 1. Identification of the top 42 transcripts most specifically regulated in CD and UC (selected as "target") against a reference set of over 3000 experimental conditions and diseases (defined as "base"). The image was re-dimensioned to better visualize it.

In this list of the top upregulated probesets, several genes have previously been associated with IBD, for example:

- REG genes: the regenerating islet-derived (REG) gene family comprises several genes up-regulated in several pathologies related to epithelial inflammation (see van Beelen Granlund et al. (2013) for a review). In our study, REG1A, REG1B, REG3A and REG4 are consistently upregulated.
- Other inflammation related genes NOS2, CXCL1, CXCL5, CXCL6, S100A8, etc.
- Intestinal immune response: DEFA5, DEFA6 (see Fahlgren et al. (2003)), and CCRL1.

Several of the genes identified as being most specific for IBD have not yet been associated with the disease in the literature. A complete list of our top hits is available for free on demand (contact us at sales@nebion.com).

IDENTIFYING RESPONSE CLUSTERS

The above search revealed genes specifically or preferentially regulated in IBD, as compared to their response to other diseases. To explore their fine-tuned regulation within IBD, we performed a hierarchical clustering in GENEVESTIGATOR across a subset of perturbations relevant to this study, i.e. perturbations from experiments related to CD, UC, CC and Infliximab. Figure 2 shows the clustering results for the top 21 up-regulated and top 21 down-regulated IBD specific genes.

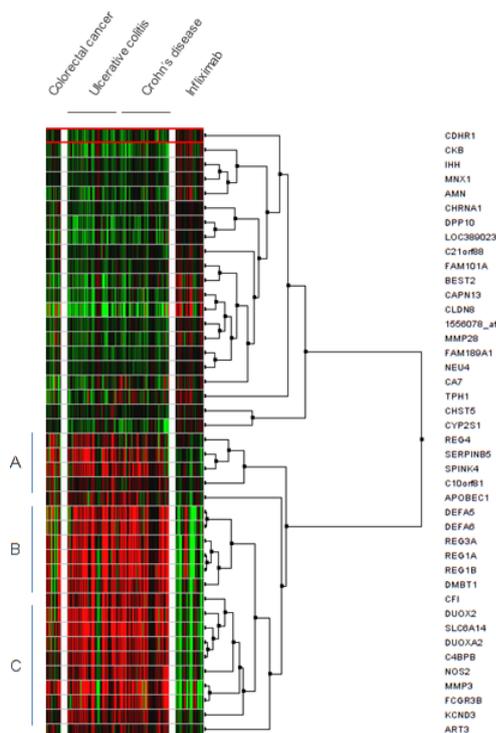


Figure 2. Hierarchical clustering of the top 42 (21 up- and 21 down-regulated) IBD-specific regulated transcripts. Among those up-regulated in CD and UC, three main clusters can be found (labeled A, B and C). Cluster B contains genes highly upregulated in IBD combined with the strongest response to (DEFA5, DEFA6 DMBT1 and the REG family of genes, with the exception of REG4 which is present in cluster A). Cluster C is characterized by NOS2 and other genes involved in the inflammatory response.

EXPLORATION OF THE REGULATORY NETWORK OF REG4

From the above search, only the top 21 specifically up-regulated candidates were retained. To explore the broader gene regulatory network around each cluster, a co-expression analysis for a chosen target gene can be performed. As an example, we chose REG4 and searched for genes having similar regulation. To obtain biologically relevant results, we selected only conditions in which REG4 is significantly regulated by using the Perturbations tool. The top 100 most correlated candidate genes are shown in Figure 3. Interestingly, from the three other genes found in cluster A (Figure 2), only SPINK4 is tightly regulated with REG4. The other two genes, namely SERPINB5 and C10orf81, were not found among the 100 most closely regulated genes for REG4. Possibly the chosen subset of perturbations not only contained CD, UC, CC and Infliximab, but also a dozen of other conditions in which REG4 is regulated. It seems that REG4 is also strongly involved in a few other processes, in which it is tightly regulated with the genes found in this analysis. Among the co-expressed genes obtained, a network analysis was done by visualizing co-expression relationships beyond a chosen threshold and circular clustering (the latter is done automatically in the Co-Expression tool). This analysis revealed the presence of several clusters, two of which are marked in Figure 3. The cluster in blue contains the second probe measuring REG4, as well as SPINK4, BACE2, TSTA3, ZG16B and KRT7 (see also Figure 4). The cluster marked in red contains genes from very different pathways, as indicated by the PANTHER database (<http://pantherdb.org/>). This list contains the genes MLPH, AGR2, CTSE, KDELR1, ANG, CREB3L1, HGD, TFF2, C11orf9, RSPH1, TMC5I, and KCNE3.

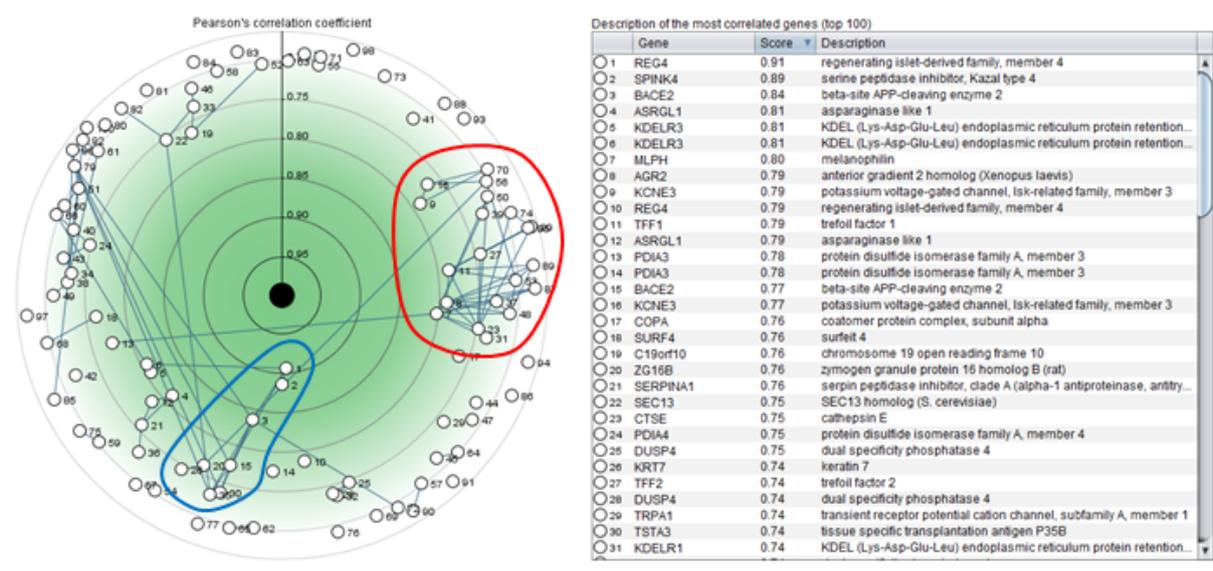


Figure 3. Co-expression analysis for REG4. The search for co-expressed genes were performed on a data matrix containing only perturbations in which REG4 is significantly regulated (77 from 3230 conditions tested).

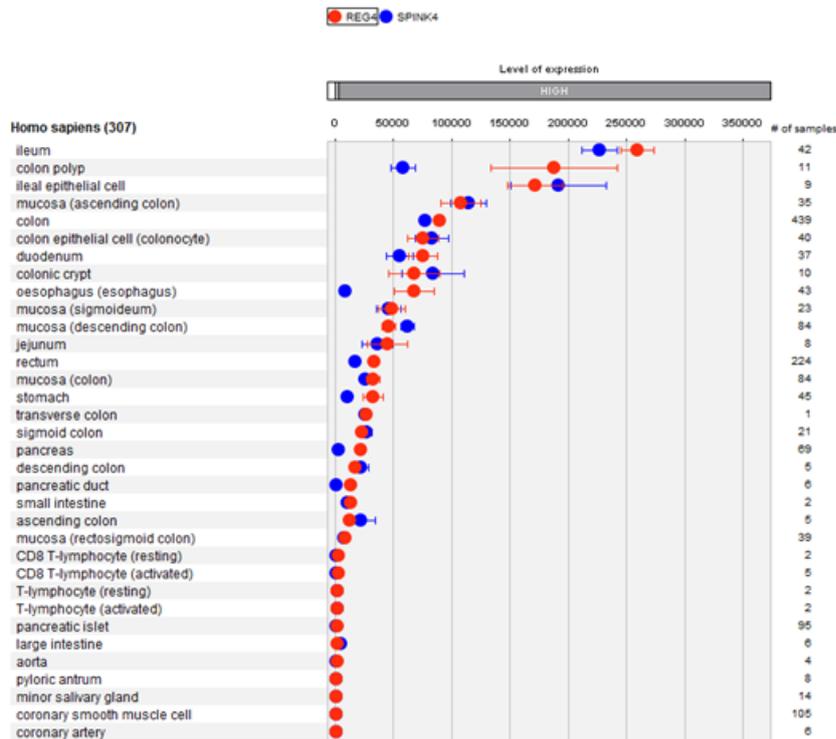


Figure 4. Expression of REG4 and SPINK4 gene across tissues and cell types. Only the top expressing tissues are represented (from 307 tissues tested). These two genes seem to be specifically expressed in the mucous or epithelial lining of the gastrointestinal tract. Besides being very similarly regulated in response to diseases and other perturbations, they are also expressed in the same set of tissues.

TREATMENT SPECIFICITY OF INFLIXIMAB

Infliximab, the first monoclonal TNF-alpha neutralizing antibody approved by the FDA for the treatment of IBD, works by reducing inflammation in the GI tract of affected individuals. In the previous part of this case study, we included by default also datasets from treated disease states (often in comparison to untreated or placebo controls, and in groups of responders versus non-responders). The analysis was done by comparing expression response in CD and UC (as "targets") against all other diseases and conditions (defined as "base"). To get a deeper insight into the mode of action of the immune-modulatory effects of Infliximab, we performed the complement analysis: we searched for genes down-regulated in response to Infliximab (as "target" category) and minimally regulated across all other conditions (defined as "base" categories), including CD and UC. The search instantly gave a list of genes specifically down-regulated upon Infliximab treatment (and minimally changed in all other conditions), as displayed in Figure 5. Moreover, the results allowed distinguishing responders (mostly green horizontal lines under Infliximab) from non-responders (mostly red horizontal lines under Infliximab). Although IBD was not selected in the target conditions, it appeared as top condition in which these genes were up-regulated. Among several hundred other indications tested, only CC comparisons yielded similar results across the entire set of genes. A few indications, such as rheumatoid arthritis, showed similar responses in a subset of these genes, suggesting that Infliximab could potentially be used for other indications. The genes obtained showed minimal response to all other 3000 tested conditions and diseases, indicating that they are possibly involved in specific downstream biological processes controlled by the inhibition through Infliximab.

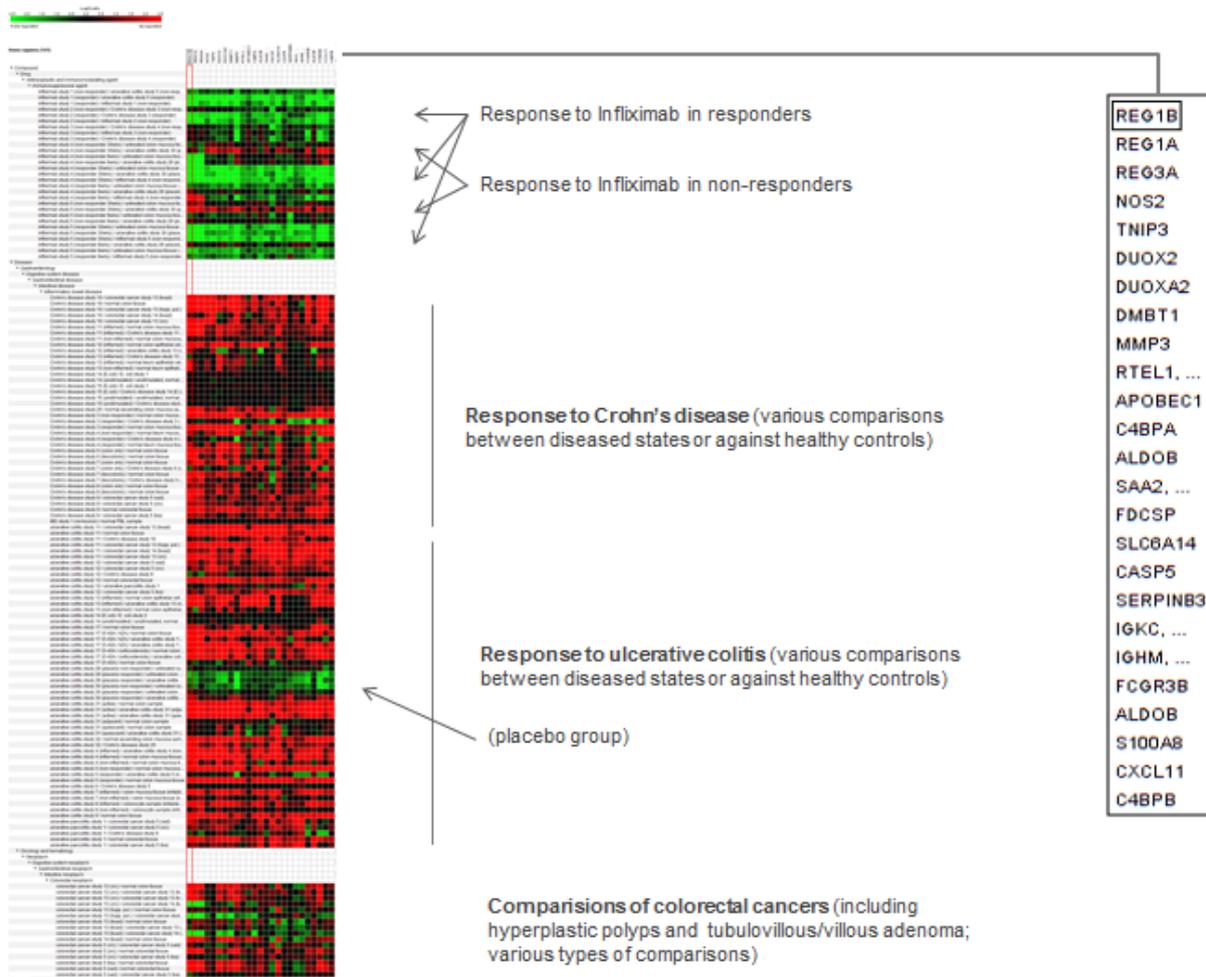


Figure 5. Identification of genes specifically down-regulated by Infliximab. Interestingly, there is a large overlap with the list of genes most specifically up-regulated in IBD.

CONCLUSIONS

The above analysis illustrates how GENEVESTIGATOR can be used to identify genes specifically regulated in a chosen disease or in selected conditions, as compared to their regulation in all other conditions (in this example, over 3000 other conditions were tested simultaneously). The results yielded genes known to be involved in CD and UC, but also brought up several new genes. The co-regulation network analysis suggests that REG4 is not only involved in the inflammatory response to CD or UC, but is also regulated in other conditions, such as response to estrogen, diabetes, lung adenocarcinoma and prostate. Among all genes co-regulated with REG4 (as based on responses to perturbations), SPINK4 stands out as being additionally expressed in the same set of tissues. It is likely that these two genes are tightly co-regulated, possibly by the same transcriptional regulator. Our analysis also gave insight into the gene networks specifically regulated by Infliximab.

REFERENCES

Baumgart DC and Carding SR (2007) Inflammatory bowel disease: cause and immunobiology Lancet 369, 1627-1640. [[Abstract](#)]

Baumgart DC and Sandborn WJ (2012) Crohn's disease Lancet 380, 1590-1605. [[Abstract](#)]

*Fahlgren A, Hammarström S, Danielsson A and Hammarström ML (2003) Increased expression of antimicrobial peptides and lysozyme in colonic epithelial cells of patients with ulcerative colitis.*Clin Exp Immunol. 2003 Jan;131(1):90-101. [[Full Text](#)]

Feagins LA, Souza RF and Spechler SJ (2009) Carcinogenesis in IBD: potential targets for the prevention of colorectal cancer. Nat Rev Gastroenterol Hepatol. 2009 May;6(5):297-305. [[Abstract](#)]

*Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W and P Zimmermann (2008) Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes.*Advances in Bioinformatics 2008, 420747 [[Full Text](#)]

Lakatos PL, and Lakatos L. (2008). Risk for colorectal cancer in ulcerative colitis: changes, causes and management strategies World journal of gastroenterology : WJG 14, 3937-3947. [[Abstract](#)]

Ordas I, Eckmann L, Talamini M, Baumgart DC, and WJ Sandborn (2012) Ulcerative colitis The Lancet, Vol. 380, No. 9853, p1606-1619 [[Abstract](#)]

Xavier RJ and DK Podolsky (2007) Unravelling the pathogenesis of inflammatory bowel disease Nature 448, 427-434 [[Abstract](#)]

van Beelen Granlund A, Ostvik AE, Brenna O, Torp SH, Gustafsson BI, and A Kristian Sandvik. (2013) REG gene expression in inflamed and healthy colon mucosa explored by in situ hybridisation. Cell Tissue Res. 2013 Jun; 352(3): 639-646. [[Full Text](#)]

CONTACT INFORMATION

For any questions or to learn more about GENEVESTIGATOR®, please write to info@nebion.com or visit genevestigator.com.