Drug repositioning in inflammatory bowel disease and primary sclerosing cholangitis by using genetic information

Valerie Collij (s2116464)
Februari 26th 2015

Prof. Dr. R.K. Weersma and Dr. E.A.M. Festen
Department of Gastroenterology and Hepatology & Department of Genetics
University Medical Centre Groningen
Summary

Our aim was to use the knowledge on the genetic background of inflammatory bowel disease (IBD) and primary sclerosing cholangitis (PSC) to identify new drug targets for IBD and/or PSC and identify biologicals or small molecules targeting these genes. We used the candidate genes of IBD and PSC discovered from previous studies and ran them through the Drugbank by using the tool we developed in R. We studied direct protein-protein interactions (PPIs) to gain insight in the networks in which the proteins encoded by IBD and PSC risk genes function (DAVID, KEGG). Finally we did a thorough literature search to select the most promising drugs for IBD and PSC based on the evidence from phase I/II/III RCTs or animal studies (PubMed, Clinicaltrial.gov). First we validated our method by showing that known IBD drugs target IBD risk genes, either directly or through one PPI. Secondly, we identified 46 drugs targeting IBD candidate genes, which have already been investigated in IBD. Thirdly, we identified 25 drugs targeting IBD risk genes, which are already used or investigated in other inflammatory disorders. Fourthly, we identified 26 experimental or investigational drugs whose mechanism looks promising for IBD. Finally we identified 16 drugs targeting PSC candidate genes, either directly or through one PPI.

Conclusion: In this study we showed that genes associated with IBD are targeted by approved therapies for IBD and we identified drugs that can possibly be repositioned or further developed for the treatment of IBD and PSC.

Ons doel was om de kennis van de genetische achtergrond van inflammatoire darmziekten (IBD) en primaire scleroserende cholangitis (PSC) te gebruiken om nieuwe aangrijpingspunten van medicijnen voor IBD en/of PSC te identificeren en het identificeren van medicijnen die op deze genen aangrijpen. We hebben de kandidaat genen van IBD en PSC gebruikt die ontdekt zijn in voorgaande onderzoeken. We hebben deze kandidaat genen gelinkt aan medicijnen uit de Drugbank door gebruik te maken van een programma in R die we zelf ontwikkeld hebben. We hebben directe eiwit-eiwit interacties (PPIs) bestudeerd om inzicht te krijgen in de netwerken waarin de eiwitten die gecodeerd worden door de IBD en PSC kandidaat genen functioneren (DAVID, KEGG). Als laatste hebben we een grondige literatuurstudie verricht om veelbelovende medicijnen voor IBD en PSC te selecteren op basis van bewijs van fase I/II/III RCTs of dierstudies (PubMed, ClinicalTrial.gov). Allereerst hebben we onze methode gevalideerd door aan te tonen dat medicijnen die voor IBD gebruikt worden aangrijpen op de IBD kandidaat genen, ofwel rechtstreeks of via een PPI. Daarna hebben we 46 medicijnen geïdentificeerd die aangrijpen op IBD kandidaat genen, al onderzocht zijn in IBD. Verder hebben we 25 medicijnen geïdentificeerd die al gebruikt worden of onderzocht zijn in andere inflammatoire aandoeningen die aangrijpen op IBD kandidaat genen. Als vierde hebben we 26 experimentele of onderzochte medicijnen geïdentificeerd die een veelbelovend werkingsevenement hebben voor IBD. Als laatste hebben we 16 medicijnen geïdentificeerd die aangrijpen op PSC kandidaat genen, ofwel rechtstreeks of via een directe PPI.

Conclusie: In dit onderzoek hebben we aangetoond dat genen die geassocieerd zijn met IBD aangrijpen op goedgekeurde medicijnen voor IBD en we hebben medicijnen geïdentificeerd die werken op andere inflammatoire aandoeningen die mogelijk ook geïndiceerd kunnen worden voor IBD en/of PSC. Ten slotte hebben we medicijnen geïdentificeerd die verder ontwikkeld kunnen worden voor IBD en/of PSC patiënten.
Introduction

Inflammatory bowel diseases (IBDs) are chronic inflammatory disorders of the gastrointestinal tract. IBD consists of two diseases, namely ulcerative colitis (UC) and Crohn’s disease (CD). Both conditions involve periods of inflammation alternated by periods of remission (1). In UC, inflammation is only present in the mucosal layer and limited to the colon (2). In CD inflammation can be present from the oral cavity till the anus. This inflammation is not limited to the mucosal layer as in UC, but can be transmural. This means that inflammation can occur through the mucosal layer to the skin or other organs (3). IBD is a common disease in Europe: approximately 2.5-3 million people in Europe suffer from it (4). The occurrence of IBD is at a relatively young age. In UC the peak age of onset lies between the ages of 30 and 40 years, while the peak of onset in CD lies between the ages of 20 and 30 years (5).

As mentioned earlier, IBD is a chronic disorder and the age of onset is relatively young. Therefore the impact on the daily life of IBD patients can be major. First of all IBD patients have 1.7 times more risk for developing colorectal cancer (CRC) compared to the general population. The cumulative risk of CRC in CD is 8.3 per cent after 30 years of disease and in UC this risk is 5 per cent (6). Secondly, many IBD patients have to undergo surgery relatively early in life. Within 10 years after diagnosis 50 per cent of the CD patients need surgery (7). In UC 20-30 per cent of the patients eventually undergo surgery (8). Thirdly, many IBD patients are suffering from their symptoms while they are active in the labor market. The consequence of this is that 19-22 per cent of the IBD patients experience long-term work disability and are dependent on a disability pension (9). All in all, IBD influences daily life on many levels and can result in a severe decrease in quality of life (10).

Although much research has been performed in IBD, the exact pathology of the disease still remains unknown (11). Researchers have discovered through twin studies that IBD has a hereditary component. The concordance rate of monozygotic twins in CD is 40-50 per cent and in UC this concordance rate is about 20 per cent. Although monozygotic twins are genetically identical, the concordance rate for developing IBD is not 100 per cent (12). These numbers show that IBD is a result of genetic predisposition in combination with environmental factors, a so-called multifactorial disease. The estimated influence of genetic predisposition in developing IBD is < 25 per cent (13).

A disease associated with IBD is primary sclerosing cholangitis (PSC). PSC is a progressive cholestatic liver disease in which the small and large bile ducts of the liver are affected by inflammation and fibrosis (14). PSC is a rare disease, the overall prevalence ranges from 0.22-16.2 per 100,000 persons and the peak age of diagnosis is 35-47 years. Complaints associated with PSC can vary from no symptoms to complaints of fatigue, lethargy and pruritis. Ultimately, PSC can lead to cirrhosis of the liver, liver failure and finally the need for liver transplantation or death (15). PSC patients have an increased risk of developing hepatobiliary malignancies: 5-15 per cent of the PSC patients will develop cholangiocarcinoma during their lives (16). PSC is associated with IBD: 21-80 per cent of the PSC patients have IBD. This large range is mostly due to differences in screening programs and different rates of performing rectal and sigmoid biopsies in PSC patients. 85-90 per cent of PSC patients with IBD
have UC. PSC-IBD patients have a significantly higher risk of developing CRC than patients who only have IBD. The odds ratio for developing CRC in PSC-IBD patients is 4.26, with a 95 per cent confidence interval of 2.80-6.48 (16). So, PSC and IBD often co-occur within the same patient, and genetic variants associated to IBD also seem to play a role in PSC (17).

In attempt to discover the exact pathology of IBD and PSC, the genetic epidemiology of both diseases has been studied through genome wide association studies (GWAS). GWAS use single nucleotide polymorphisms (SNPs). SNPs are single base variants in DNA that are highly variable between individuals (figure 1). These SNPs are compared between patients and healthy controls, searching for significant differences to identify areas of the genome (risk loci) associated with disease (18).

![Figure 1: An example of a SNP, a difference in a single base sequence (18)](image)

If a certain SNP occurs significantly more often in the disease group than in the control group, this SNP is likely to be associated with disease. At this moment, GWAS analyse up to 2,5 million SNPs per individual. Because of this large amount of multiple testing, the chance of identifying false-positive association of a SNP is high if you stick to a p value of < 0.05. In GWAS, the appropriate significance threshold is therefore determined at a p value of < $5 \times 10^{-8}$ (0.05/1000.000) (19). At this moment, 233 risk loci in IBD (20, unpublished confidential data) and 16 risk loci in PSC (21) have been identified, that meet this statistical genome wide significance thresholds.

In these risk loci genes are located which encode proteins. These proteins could be of influence for the diseases IBD and PSC. Previous studies identified genes within these loci, which are candidate genes for IBD and PSC. The discovery of the gene STAT1 is an example of an IBD associated candidate gene and is located at chromosome 2 (20). Another example is candidate gene IL2, associated with PSC and located at chromosome 4 (21). The identification of candidate genes within the susceptibility loci in IBD and PSC is very important for gaining better understanding of the pathogenesis of both diseases. In addition, previous research discovered that almost 75 per cent of the IBD risk loci overlap with genetic risk loci for other immune mediated diseases, like rheumatoid arthritis (RA) (22). This implies that underlying disease mechanisms are similar between different immune mediated diseases.
One thing to consider is that the proteins encoded by candidate risk genes function within networks by interacting with other proteins, the so-called protein-protein interaction (PPI). An example of a PPI is the interleukin 23 (IL23) pathway (see figure 2). The gene IL23R is proven to be associated with CD but IL23 is not. As seen in figure 2, IL23 is directly needed in order to activate IL23R and they therefore have a direct PPI (23).

Figure 2: IL23 pathway, genes marked with an asterisk are proven to be associated with CD (23)

In order to get more insight into disease mechanisms Okada et al. took these PPIs into account and developed a systemic strategy to integrate genetic risk variants with diverse genomic and biological data sets, building networks of the pathways in which the risk genes interact. This strategy provides insight into the pathogenesis of RA (24). They also showed how these findings can be translated into clinical practice by guiding drug discovery for complex traits in RA. They demonstrated that putative RA risk genes are targets of approved therapies for RA. They also showed that drugs approved for other indications could possibly be repurposed for the treatment of RA (24). Their study provides a good example of developing targeted therapy for a complex disease for which genetic risk variants have been identified (see figure 3).

Figure 3: An overview of RA candidate genes found in the RA risk loci and drugs which target these genes. Drugs used for other diseases than RA might also be useful for RA (24).
Discovering targeted therapy for IBD and PSC is very important, because both conditions do not have optimal treatment yet. As mentioned earlier, IBD patients suffer from periods of inflammation alternated by periods of remission. IBD patients receive two types of treatment, namely therapy to bring active disease in remission and maintenance therapy to keep disease in remission. Medical treatment consists of drugs like mesalazines, anti-TNFs and corticosteroids. If these medical treatments fail, patients need to undergo surgery with resection of the inflamed intestine or colon (25).

In PSC the only available drug treatment is Ursodeoxycholic acid (UDCA). PSC patients treated with UDCA show a trend towards a better survival than PSC patients treated with placebo. However, the difference in survival has not been shown to be significant so far, partly due to a lack in power, meaning more research has to be done on a larger sample of patients (26). Up until this point liver transplantation is the only treatment available for advanced stage PSC (27). This indicates the importance of better medical treatment in PSC.

However, the development of drugs is extremely expensive, approximately 60 billion US dollar for the development of around 20 molecular entities per year. The reason that these costs are so high is because of failure of the invented molecular entities in later testing. Only 5 per cent of the potential drugs, which reach Phase I clinical trials are approved for clinical use by the Food and Drug Administration (FDA) down the line (28). This is mostly due to the fact that one cannot predict the working mechanism of the investigational drug in the specific test setting.

A way to make drug development more efficient is the use of genetic information. Even if a certain gene does not contribute that much in the occurrence of the disease, targeting that gene could have a very large impact. A great example of this is HMG-CoA reductase inhibitors (statins), which have a major impact on LDL-cholesterol levels in blood (29). In the general population however, the genetic variants in the gene encoding HMG-CoA reductase have only very small effects on the levels of LDL-cholesterol (30). So, by knowing which proteins are targeted by drugs, and which genes encode these proteins, we can link drugs to protein targets associated with candidate risk genes for disease.

The discovery of new drugs is important for the treatment of both IBD and PSC as mentioned earlier. By using the candidate genes found in previous studies, we can link candidate genes to drugs and identify promising new drugs for both diseases. These drugs might already be in use for or being investigated in IBD and/or PSC, but it could also be that these drugs have been developed for other diseases. Because they target IBD disease pathways these drugs could possibly be repurposed for IBD and/or PSC (31, 32). Given the fact that PSC is a rare disorder, this way of developing new drugs is important. The development of drugs for rare disorders is relatively expensive and the studies are limited to small populations. Therefore it is hard to conduct clinical research in this disease group and to get funding for it (33). Using genetic information as a starting point for medical research in rare disorders can make this research easier and more cost-efficient.
Research question

At this moment, there is no systematic strategy for providing a list of drugs which can be used to target pathways of the specific risk genes of IBD and PSC. The aim of this project is to develop a list of registered and investigational drugs, which can be used to target specific proteins, as risk genes for IBD and PSC, identified through genetic studies.

This leads to the following research question: Can we identify registered and/or investigational small molecules or biological therapies that could be used for the treatment of IBD and/or PSC, based on genetic risk variants?

The developed list of drugs will contain drugs which specifically target proteins from IBD and/or PSC associated genes or genes that are in the same biological pathway or have a direct PPI with IBD and/or PSC risk genes. These drugs could be drugs that have already been developed for IBD patients, drugs developed for other indications, or investigational drugs that are not yet on the market.
Materials and methods

The basis of this project is a list of candidate risk genes for IBD and PSC derived from the genetic risk variants identified through GWAS. These candidate genes have been derived from the associated genetic risk variants in previous research (20, 21 and unpublished confidential data). Before extracting the candidate risk genes of IBD and PSC from these papers, I first had to understand how these candidate genes had been identified. The method for doing this as described in the paper of Jostins et al. is the method that is most widely used and therefore this is the method I describe in this report (20, 21).

Jostins et al. performed a meta-analysis in which they combined data from 15 GWAS of IBD. They tested 1.23 million SNPs in 25,683 IBD patients (cases) and 10,920 healthy individuals (controls) to study whether certain SNPs occurred significantly more often in cases than in controls. Because of this large amount of cases and controls, the study had a very good statistical power to identify new genetic risk variants. The researchers also stratified their analysis by study population, meaning that they corrected for differences within the study samples. The reason for this is to prevent the identification of false-positive IBD associated SNPs, which are actually associated with other factors, such as differences in ancestry within the study population.

Eventually the researchers discovered 163 genetic risk loci associated with IBD. In these 163 risk loci genes are located and several steps were made to identify candidate genes for IBD within these regions on the genome (20). First of all the researchers identified which SNPs were in linkage disequilibrium (LD). LD means that variants in the DNA (like SNPs or whole genes) inherit together because they are on the same string of DNA that does not break during cell division. It could therefore be that the associated SNP found in previous analyses is not causal to disease, but that the associated SNP is in LD with the causal SNP (34). The next step was to analyze which genes were located in or nearby the SNPs associated with IBD by using expression quantitative trait locus (eQTL) analyses. An eQTL is a genetic variation that strongly correlates with the expression levels of messenger RNA (mRNA). By detecting which mRNAs were influenced by the IBD associated genetic risk variants, the genes regulated by the IBD associated risk SNPs were identified (35).

Furthermore, the researchers analyzed the relations between putative risk genes by using the GRAIL tool. GRAIL uses a statistical framework that assesses the significance of relatedness between genes in disease regions. GRAIL also uses a text-based similarity measure that scores two genes for relatedness to each other based on text in PubMed abstracts (36). The researchers also used DAPPLE; a tool that analyzes whether proteins encoded by the risk genes interact with genes in other loci, either directly or indirectly, and could therefore contribute to disease (37). Finally, Jostins et al. performed co-expression analysis. Co-expression is a phenomenon in which certain genes encode proteins that are expressed simultaneously and are therefore likely to be involved in the same pathway. By analysing these pathways, additional candidate genes of disease were discovered (38).

After following all these steps, previously conducted research identified 362 candidate genes for IBD and 16 candidate genes for PSC, which we use in this
The smaller amount of PSC candidate genes is mainly due to the rareness of PSC, resulting in less power for detecting PSC risk loci than in IBD risk loci (20, 21, unpublished confidential data). In order to link these candidate genes to drugs, we downloaded the Drugbank. The Drugbank is the largest publicly available database containing 7737 drugs and additional information about these drugs like their protein targets (39). With the help of dr. R. Alberts I developed a tool linking putative IBD genetic drug targets directly to drugs by using R. R is a language and environment for statistical computing and the development of graphics (40). I followed an online R course to learn about the basics of working with R (41). The function of R in this project was to develop a script (see appendix 1) in which we could upload a file with candidate genes of IBD and PSC in R. The next step was to run the Drugbank download through the script, producing a list of drugs that target the IBD and PSC candidate genes. In order to do so, several steps were needed.

The first step was to study where the information we needed, in this case the drugs and their gene targets, were stored in the Drugbank download. The Drugbank download is an XML file and therefore the XML package of R was needed to open it in R (42). An XML file consists of nodes and inside these nodes more information is stored. We therefore analysed in which nodes the drugs and their gene targets were stored in the download. After establishing the location of this information, we wrote a script that lists all the drugs and their associated gene targets available in the Drugbank in a table. In order to do this we had to consider the fact that many drugs do not have a known gene target and this information is therefore not available in the Drugbank. A drug with no gene target did not get listed in the table.

The next step was to extract the candidate genes of IBD and PSC and their associated drugs from this table. Before we could do this, we had to consider the fact that genes have synonyms. In order to take the synonyms into account, we had to use the R package org.Hs.eg.db. The org.Hs.eg.db package contains the abbreviations, full names and synonyms of human genes (43). For example, a candidate gene of IBD is ALDH2. Table 1 contains the information from the org.Hs.eg.db package about the gene name ALDH2. The synonyms were described in the third column, as seen in table 1. We therefore developed a script in which the synonyms of the candidate genes of IBD and PSC were extracted automatically from the org.Hs.eg.db package (see appendix 2). By using the synonyms, we prevent missing drugs due to different names of gene-targets listed in the Drugbank.

Table 1: the org.Hs.eg.db package contains the id number, synonyms and full gene names of human genes, for example the human gene ALDH2

<table>
<thead>
<tr>
<th>Row</th>
<th>Id</th>
<th>Synonyms</th>
<th>Gene name</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>702</td>
<td>184</td>
<td>ALDH2</td>
<td>aldehyde dehydrogenase 2 family (mitochondrial)</td>
<td>ALDH2</td>
</tr>
<tr>
<td>703</td>
<td>184</td>
<td>ALDH2</td>
<td>aldehyde dehydrogenase 2 family (mitochondrial)</td>
<td>ALDH2</td>
</tr>
<tr>
<td>704</td>
<td>184</td>
<td>ALDHI</td>
<td>aldehyde dehydrogenase 2 family (mitochondrial)</td>
<td>ALDH2</td>
</tr>
<tr>
<td>705</td>
<td>184</td>
<td>ALDM</td>
<td>aldehyde dehydrogenase 2 family (mitochondrial)</td>
<td>ALDH2</td>
</tr>
</tbody>
</table>
After taking these steps, we could run the candidate genes of IBD and PSC, including their synonyms, through the Drugbank download and get the drugs targeting these candidate genes. We developed an additional script, using a similar method, to extract more extensive information on the drugs from the Drugbank: information like indication of the drug (see appendix 3). The development of the tool in R was very important for this project, because this tool allowed us to extract all the drugs, which target the candidate genes of IBD and PSC, automatically. This tool will also be very useful for repeating this process to identify targeted drugs for other diseases based on genetic risk variants.

The next step of this project was the analysis of PPIs to gain insight in the networks in which the proteins encoded by the IBD and PSC risk genes function. For IBD we analysed the PPIs in approved drugs used for the treatment of IBD. We did this by looking up the gene targets of approved IBD drugs in the Drugbank. For PSC we studied the PPIs for all PSC candidate genes. The reason for this different method in IBD and PSC is that the highest amount of risk loci has been identified in IBD and therefore IBD is the most successfully studied complex disease (31). In PSC however, not much risk loci have been identified due to the rareness of the disease. Therefore we did in PSC for all 16 candidate genes PPI analysis to gain more insight in the pathogenesis of PSC (21).

After we got the synonyms of these genes by using the script described in appendix 2, we put these genes in the DAVID database. DAVID is a web-based database that provides PPIs and protein pathways for uploaded gene names. DAVID collects its information from other databases, for example the KEGG database. After submitting the genes, we got a list of genes which interact either directly or indirectly with our submitted genes (44). For this research we only used the direct PPIs, because these PPIs influence the candidate genes directly. In order to figure out which genes were in direct PPI, pictures from the KEGG database were studied. KEGG provides maps in which protein pathways are visualized (see figure 4) (45).

![Diagram](image)

Figure 4: The direct PPI of for example TNF/TNFR (45).
The last step of this project was a thorough literature search to assess which drugs found in the previous steps were the most promising new drugs for IBD and PSC. This step was essential, because the drugs found in the previous steps were only selected based on an interaction between a certain gene and a drug, but what kind of interaction is unknown. It could for example be that a drug could work very well for IBD patients but it could also be that a drug gives colitis as a side effect and thus increases the symptoms of IBD.

We used PubMed to analyse the previously found drugs. First we studied whether a drug was associated with IBD by using the following query: ("medicine") AND ("IBD" OR "crohn" OR "ulcerative colitis" OR "colitis"). For PSC we used this query: ("medicine") AND ("PSC" OR "primary sclerosing cholangitis"). These queries allowed us to analyse whether a drug was used in the treatment of IBD, investigated in IBD or PSC or if the drug causes colitis-like side effects. If there were no associations with IBD or PSC in the literature, we studied whether there was an association with other autoimmune disorders by using the following query: ("medicine") AND ("auto" OR "auto immune" OR "autoimmune" OR "inflammation" OR "inflammatory"). If there still was no association, we studied whether a drug had an interesting working mechanism for IBD or PSC. Drugs were selected based on evidence from phase I/II/III randomized clinical trials (RTCs) or animal studies.

The queries used for PubMed were also used for ClinicalTrials.gov. ClinicalTrials.gov is a web-based database in which clinical trials are listed. Since 2008 researchers are obligated to register their clinical trial on this website. This is important, because failed clinical trials are also listed here. Even if researchers decided not to publish their failed trial, one can find information about it on ClinicalTrials.gov (46). This database also contains information about clinical trails that are currently recruiting. We took ClinicalTrials.gov into account to make sure that a drug that was reported to be effective in in an article PubMed was not reported to have i.e. severe side effects in trials registered in ClinicalTrials.gov, but never published.
Results

Through the method described above we identified a total of 103 drugs associated with IBD and 16 drugs associated with PSC (see figure 5). We identified these drugs by running 1631 candidate gene-names associated with IBD through the Drugbank. For PSC we ran 140 candidate gene-names through the Drugbank, this number was based on the 16 PSC associated candidate genes, 13 direct PPIs with these candidate genes and the synonyms of these 29 genes. After running these gene-names for both conditions in the Drugbank with our script we identified 338 drugs linked to 83 candidate genes of IBD and 37 drugs linked to 7 PSC candidate genes. Of these drugs, we excluded 235 IBD linked drugs and 21 PSC linked drugs because of lack of evidence in animal studies and phase I/II/III RCTs or because of expected adverse effects of these drugs for IBD and/or PSC. We divided the remaining 103 IBD associated drugs in four groups based on their characteristics.

Figure 5: flowchart, which shows the amount of drugs we found

First of all we validated our method by showing that known IBD drugs target IBD risk genes, either directly or through one PPI (see figure 6). For example, the drugs Mesalazine and Sulfasalazine are both 5-aminosalicylic acid derivatives and are directly linked to the IBD candidate gene PTGS2. These drugs also target the genes CHUK and IKBKB, which are in direct PPI of the IBD candidate genes NFKB1 and RELA. The tumor necrosis factor blocking agent (anti-TNF) Adalimumab is directly linked to the IBD candidate genes FCGR2A, FCGR2B, FCGR3A and FCGR3B. Adalimumab also targets the gene TNF, which is also the gene target of the anti-TNF drugs Infliximab and Certoluzimab. TNF is linked to the IBD candidate genes NFKB1, TNFRSF1A and MAPKAPK2 through one PPI. Finally, the immunosuppressant agent Ciclosporin targets the gene PPP3R2, which is linked to the IBD candidate gene CALM3 through one PPI.
Figure 6: the link between the IBD genetic risk variants (SNPs) (red), IBD candidate genes associated with the risk variants (purple), genes from PPI (blue) and approved IBD drugs (orange).

In the second part of our study we identified 46 drugs targeting IBD candidate genes, which have already been investigated in IBD, either through RCTs or through animal studies (see appendix 5 for the complete figure of 46 drugs targeting IBD candidate genes). As seen in figure 7, examples of drugs we identified as potential IBD drugs that are being investigated for the treatment of IBD are vitamin D analogues. In clinical practice IBD patients get vitamin D analogues administered to prevent osteoporosis, which is common in IBD patients (47). However, animal studies have shown that vitamin D is a regulator of the immune system and could therefore be a potential drug for autoimmune disorders like IBD (48). A RCT in 108 CD patients revealed that CD patients receiving vitamin D3 treatment had a reduced risk of clinical relapse from 29 per cent to 13 per cent compared to the CD patients receiving placebo. This reduced risk was not significantly different, with a P value of 0.06 (49). Larger studies need to be done to investigate this effect further and information from our study could stimulate such research.
Another example of a drug we identified as potential IBD drug which is investigated in IBD is Thalidomide. Thalidomide is a controversial drug: it is not listed in the Dutch treatment protocol for IBD. In some cases however, physicians do administer it to IBD patients who do not respond to steroids or other drugs normally used to treat IBD. The controversy surrounding Thalidomide is due to its potential severe side effects, like polyneuropathy (25). An open label clinical trial has shown that Thalidomide can be effective in some patients with CD, but only 22 CD patients were enrolled in this trial (50). More research needs to be done to investigate this further and our study shows that genetic information supports the use of Thalidomide in the treatment of IBD.

![Figure 7: Ten examples of drugs which are investigated in IBD, but the drugs are not (yet) listed in the Dutch protocol of gastroenterology for usage in IBD patients](image)

In the third part of our study we identified 25 drugs targeting IBD risk genes, which are used or investigated in other inflammatory disorders (see appendix 5 for the complete list of 25 drugs targeting IBD candidate genes). Examples of drugs used or investigated in the inflammatory disorders RA, multiple sclerosis (MS) or psoriasis are listed in figure 8. Ruxolitinib for example is a JAK1/2 inhibitor and is approved by the FDA for use in myelofibrosis, but the drug has also been tested in RA in a double blind phase IIa RCT. This study showed a high efficacy of Ruxolitinib in RA patients (51).

Another example is the drug Muromonab, an immunosuppressive agent. Muromonab is an anti-CD3, meaning that it is directed against the CD3 receptor, located at T lymphocytes. Muromonab has been approved for use in patients with acute transplant rejection. Animal models have shown that Muromonab also has an effect on autoimmune reactions and research in healthy individuals has shown a decrease in certain inflammatory cytokines, like IL17 after administration of Muromonab. A phase II placebo-controlled trial showed efficacy of this drug in patients with autoimmune diabetes (52).
The last example in this group is Interferon beta 1b, which has been approved for use by the FDA in 1993 for the autoimmune disorder MS (39). The therapeutic actions of this drug are not (yet) clear, but research has found that Interferon beta 1b has an anti-inflammatory effect (53). Based on research performed in other inflammatory disorders and based on the genetic information in our study, these drugs look very promising for IBD.

![Image of genetic information and drug targets]

**Figure 8:** Ten examples of drugs which are investigated in other inflammatory disorders and could also be effective in IBD.

In the fourth part of our study we identified 26 experimental or investigational drugs whose mechanism looks promising for the treatment of IBD (see appendix 6 for the complete figure of 26 drugs targeting IBD candidate genes). As seen in figure 9, an example of this is the experimental drug INCB3284. INCB3284 is a CCR2 antagonist (39). The role of CCR2 in IBD has been investigated by comparing resection specimens of the ileum of CD patients with resection specimens of the ileum of patients who had non-IBD related conditions. The researchers found a 30 times higher CCR2 expression in the ileum of CD patients than in the ileum of the patients with non-IBD related conditions. The researchers concluded therefore that a CCR2 antagonist could be an effective treatment of CD, especially the CD patients with disease activity in the small bowel. The investigational drug INCB3284 should therefore be further investigated for the treatment of IBD (54).
Another example is the investigational biotech drug SGN-30, which has a high affinity for the CD30 receptor (39). Researchers studied the role of CD30 in mice studies for IBD, because CD30 is increased in the serum of IBD patients. In this mice study they found that mice, which did not have the CD30-ligand and therefore could not produce CD30, did not develop IBD while the mice with the CD30-ligand did develop IBD. Therefore the researchers suggested that a drug which decreases the CD30 level in the serum of IBD patients could be very beneficial for IBD patients. An example of this drug could be SCN-30. Because of this working mechanism we selected SCN-30 as a potential drug for IBD (55).

**Figure 9: Fifteen examples of drugs which have a promising working mechanism for IBD, which makes them potential IBD drugs.**

In the final part of our study we identified 16 potential new drugs for PSC (see for the complete figure appendix 7). Figure 10 for example shows the approved biotech drug Aldesleukin. Aldesleukin is an IL2 analogue and is used in the treatment of metastatic renal cell carcinoma (39). Currently a clinical trial is recruiting patients with autoimmune disorders like IBD, RA but also PSC, to study the effect of administering low dosages of Aldesleukin (56). This study is being performed because previous research has shown that IL2 is an important cytokine in controlling the immune system and could therefore be an important target for drugs in autoimmune disorders (57).

For PSC we also identified 11 drugs used or investigated in other inflammatory disorders. Etanercept for example is approved for use in RA and Daclizumab is investigated in MS (39). We finally identified 4 potential new PSC drugs that are not (yet) investigated in other inflammatory disorders, but do have an interesting working mechanism that makes them interesting for use in PSC. An example is the experimental small molecule LFA703. Previous research has shown that LFA703 inhibits IL18 and this cytokine plays a role in inflammation and autoimmune disorders. LFA703 could therefore be very interesting for the treatment of PSC (58).
Figure 10: Nine examples of potential new PSC drugs
Discussion

The aim of my study was to use the knowledge on the genetic background of IBD and PSC to identify new drug targets for IBD and/or PSC and to identify biologicals or small molecules targeting these genes. As a starting point we used the candidate genes for IBD and PSC that have been discovered by previous studies. We searched the Drugbank for drugs targeting these risk genes using the tool we developed in R. We also studied PPIs of IBD and PSC risk genes by using DAVID and KEGG in order to find more drug targets for both disorders. Finally we did a thorough literature search to select the most promising drugs for IBD and PSC based on the evidence from phase I/II/III RCTs or animal studies.

We identified a total of 103 potential new drugs which target IBD candidate genes and 16 potential new drugs that target the candidate genes of PSC. We divided these drugs into five groups. The first group is the group in which we validated our method by showing that known IBD drugs target IBD risk genes. The second group is the group of 46 drugs that we identified which target IBD candidate genes and which have already been investigated in IBD, either through RCTs or through animal studies. Thirdly, we identified 26 drugs targeting IBD risk genes, which are already used or investigated in other inflammatory disorders. These drugs could possibly be repositioned for treatment of IBD. In the fourth group are 26 experimental or investigational drugs which we identified and whose mechanism looks promising for IBD. Finally we identified 16 drugs targeting PSC candidate genes, either directly or through one PPI.

These findings are very important for both IBD and PSC, because both diseases currently lack optimal medical treatment. If medical treatment in IBD patients fails, patients need to undergo surgery with potentially invalidating resection of the inflamed intestine or colon (25). For PSC no medical treatment is available at all (26). Currently, liver transplantation is the only treatment available for advanced stage PSC (27). The limited treatment options for both IBD and PSC influence daily life on many levels and can result in a severe decrease in quality of life in these patients (10). However, research to discover new drugs in order to improve the medical treatment options for IBD and PSC is extremely expensive. This is mostly due to the fact that one cannot predict the working mechanism of the investigational drug in the specific primary test settings and the subsequential clinical trials (28). Especially in rare disorders like PSC this is a problem due to low funding and a small sample size (33). Using genetic information as a starting point for medical research in (rare) disorders can make this research easier and more cost-efficient.

Although we identified many potential new drugs for IBD and PSC patients, it is possible that we missed potential new drugs that could be identified by searching for drugs targeting genetic risk variants due to a lack of information of certain factors vital for this research. Our research suffers from two main caveats in the basic information. First of all: our knowledge of the effect of common genetic variants on most mechanisms is unknown. We generally know whether genetic variants influence gene expression or not, but whether they influence the immune system or a disease phenotype in any other way is generally unknown. The first step in our study, the step from disease associated genetic variants to candidate gene, is hence based on incomplete knowledge. Also, we do not know how many risk loci and candidate
genes are not discovered (yet) in both IBD and PSC. In IBD much research has been performed to identify these risk loci with very well powered studies. In PSC however, much less risk loci have been discovered (21). This is mostly due to the rareness of the disease, which renders the studies looking for new genetic risk loci underpowered (32). If we are missing risk loci and candidate genes of both disorders, we are missing the possibility of discovering drugs for these diseases based on genetic information. Furthermore, we do not know what the biological effects are of all these genetic risk variants and putative candidate genes because the exact pathology of both IBD and PSC is not yet known (11). Examples of drugs we missed were the thiopurines azathioprine and 6-mercaptopurine. Reason of this could be the factors mentioned above.

Another problem, related to the previous problem of our incomplete understanding of disease mechanisms, is that we do not know what the effect of a potential new drug linked by our method to a candidate gene would be. Questions concerning this subject could be: do we need up-regulation or down-regulation of the proteins encoded by the candidate genes? On top of that we do not know the working mechanisms of all drugs; do they up-regulate or down-regulate the proteins encoded by the candidate genes? We tried to answer some of these questions by performing the thorough literature search to predict the function of the drug based on its effect in other (inflammatory) disorders.

Moreover we experienced that for certain drugs their gene-target was not known or we had to exclude drugs due to the fact that they were too experimental and therefore not listed in PubMed. Therefore it could be that we missed promising drugs for the treatment of IBD and/or PSC. There is also the possibility that we excluded drugs, which were not tested in other inflammatory disorders but could be promising for IBD and/or PSC. A reason for this is that drugs are not selected for clinical trials based on genetic information but more based on trial and error in more prevalent diseases with similar phenotypes. Therefore it could be that more drugs are available that could be used for the treatment of for IBD and PSC but they are not being tested in any inflammatory disease.

To keep our study manageable we had to employ a different method for IBD and PSC for including candidate genes based on PPIs. For IBD we only studied the PPIs in the gene targets of approved therapies for IBD patients, whereas for PSC we studied the PPIs for all PSC candidate genes. IBD is a very well investigated disease, due to its high prevalence in western society. PSC however is not as thoroughly investigated and not as many genetic risk variants have been identified. Therefore, we choose to add the PPIs for all PSC candidate genes; we hope this provides additional insight in the pathological background of PSC.

Finally we need to consider the value of the potential new drugs, which we have selected as promising. How much influence can a single genetic variant have on the pathogenesis of a disease and how much influence does targeting a gene have on the symptoms patients experience? From literature we know that even if a certain genetic variant does not contribute that much in the total risk for disease, targeting that gene can have a very large impact. A great example of this is HMG-CoA reductase inhibitors, which have a major impact on LDL-cholesterol levels in blood (29). In the general population however, the genetic variants in the gene encoding HMG-CoA
reductase have only very small effects on the levels of LDL-cholesterol (30). We also selected drugs based on evidence of phase I/II/III RCTs and animal studies, not only in IBD or PSC but also in other inflammatory disorders. A reason for this is that almost 75 per cent of the IBD risk loci overlap with genetic risk loci for other immune mediated diseases, like RA (22). So if a drug works for RA patients, it could therefore work for IBD or PSC patients.

Overall, in this research we have shown that genes associated with IBD are targeted by approved therapies for IBD and we identified drugs that can possibly be repositioned or further developed for the treatment of IBD and/or PSC. The drugs we selected as most promising should therefore be investigated in these diseases in the future. These findings could lead to better treatment for IBD and/or PSC patients by using already existing drugs and this tool can be used for other diseases based on their genetic background.
References


(18) Visschedijk MC, Festen AEM, Wijmenga C, Weersma RK. De genetische achtergrond van inflammatoire darmziekten ontrafeld door genoomwijde associaties studies. Ned Tijdschr Geneeskd 2009;153;A402

(19) Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet 2014 May;15(5):335-46


(34) Li YR, Keating BJ. Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. Genome Med 2014 Oct;6(10):91


(41) Peng R, Data Science Team. Coursera course R Programming by John Hopkins University. Available at: https://www.coursera.org/course/rprog


(46) Clinicaltrials.gov Background. Available at: https://clinicaltrials.gov/ct2/about-site/background

(47) O'Sullivan M. Vitamin D as a novel therapy in inflammatory bowel disease: new hope or false dawn? Proc Nutr Soc. 2015 Feb;74(1):5-12


(53) Freedman MS. Evidence for the efficacy of interferon beta-1b in delaying the onset of clinically definite multiple sclerosis in individuals with clinically isolated syndrome. Ther Adv Neurol Disord 2014 Nov;7(6):279-288


(56) Induction of Regulatory t Cells by Low Dose il2 in Autoimmune and Inflammatory Diseases (TRANSREG). Available at: https://clinicaltrials.gov/ct2/show/NCT01988506?term=Aldesleukin+AND+PSC&rank=4


Appendix 1: script of listing all the drugs and their gene targets in a table

genes <- read.table("PPIpscgenes.txt", sep="\t", header=F, quote="")
genes1 <- as.character( read.table("PPIpscgenes.txt", sep="\t", header=F, quote="")[,1] )
aantalgenes <- length(genes1)

library (XML)
doc <- xmlTreeParse("xaa.xml", getDTD = F)
r <- xmlRoot(doc)
aantaldrugs <-xmlSize(r)
drugs <- NULL
genes <- NULL
werking1 <- NULL
for (i in 1:aantaldrugs) {

drugname <- xmlValue(r[[i]]"name"[1])
targets <- r[[i]]"targets"
kinderen <- xmlChildren (targets[1])
aantalkinderen <- as.numeric (length(kinderen))
if (aantalkinderen > 0 ) {
    for ( j in 1:aantalkinderen ) {
        if ( ! is.na(xmlValue(kinderen [[j]]"polypeptide"[1][1])) ) { #hoort nog iets te staan, zie onderaan
            ###if ( ! is.na(xmlValue(kinderen [[j]]"polypeptide"[1][1][1])) ) )
            {
                if ( length ( xmlValue(kinderen [[j]]"polypeptide"[1][1][1])"gene-name"[1][1]) )
                >0 ) { #length is zero tegengaan door length in te voeren

                    gen <- xmlValue ( kinderen [[j]] [[7]] "gene-name")[1])
                    werking <- xmlValue ( kinderen [[j]] [[6]])
                    drugs <- c ( drugs , drugname )
                    genes <- c ( genes, gen )
                    werking1 <- c (werking1, werking)

                }
            }
        }
    }
}
result <- data.frame(drugs=drugs, genes=genes, werking1= werking1)
Appendix 2: script for extracting the synonyms of the IBD and PSC candidate genes from the org.Hs.eg.db package

library(org.Hs.eg.db)
dbCon <- org.Hs.eg_dbconn()
sqlQuery <- 'SELECT * FROM alias, gene_info WHERE alias._id == gene_info._id;' aliasSymbol <- dbGetQuery(dbCon, sqlQuery)

totaal_gensyn <- data.frame()
for (v in 1:aantalgenes) {
  gennammeind <- genes1 [v]
  endresultsyn <- aliasSymbol[aliasSymbol$alias_symbol==(genes1 [v]), ]
  totaal_gensyn <- as.data.frame (rbind(totaal_gensyn, endresultsyn))
}

aantalendresultsyn <- nrow (totaal_gensyn)
edresultsynid <- data.frame ()
totaal_gensynid <- data.frame ()
columid <- totaal_gensyn [ ,1]
columid2 <- aliasSymbol [,1]
for (p in 1:aantalendresultsyn) {
  tussenstap <- as.character (columid [p])
  endresultsynid <- subset(aliasSymbol, columid2 == tussenstap)
  totaal_gensynid <- as.data.frame (rbind(totaal_gensynid, endresultsynid))
}

write.csv(totaal_gensynid, "synoniemenPPIpsc.csv")
Appendix 3: script for additional information about the drugs

drugfound <- read.table("drugs.txt", sep="\t", header=F, quote="")
drugfound1 <- as.character ( read.table("drugs.txt", sep="\t", header=F, quote="")[,1] )

aantaldrugsfound <- length(drugfound1)

doc <- xmlTreeParse("xan", getDTD = F)
r <- xmlRoot(doc)
aantaldrugs <- xmlSize(r)

drugs <- NULL
druggroups <- NULL

for (i in 1:aantaldrugs) {

    drugname <- xmlValue(r[[i]]['name'][1])
    druggroup2 <- r[[i]]['groups']
    kinderen <- xmlChildren (druggroup2[[1]])
    aantalkinderen <- as.numeric (length(kinderen))
    if (aantalkinderen > 0 ) {
        for (j in 1:aantalkinderen) {

            if (! is.na(xmlValue(kinderen [[j]][1]))) {
                if ( length (xmlValue(kinderen [[j]]["groups"][1]["group"][1][1])) >0 ) {
                    groups <- xmlValue ( kinderen [[j]] [1])
                    drugs <- c ( drugs , drugname )
                    druggroups <- c ( druggroups, groups )
                }
            }
        }
    }

resultgroups <- data.frame(drugs=drugs, druggroups=druggroups)
totaaldruggroups <- data.frame()

for (v in 1:aantaldrugsfound) {
    drugsnaam <- drugfound1 [v]
    endresultdrugs <- resultgroups[resultgroups$drugs==(drugfound1 [v]),]
    totaaldruggroups <- as.data.frame (rbind(totaaldruggroups, endresultdrugs))
}

write.csv(totaaldruggroups, "druggroupsxan.csv")
Appendix 4: 46 drugs which are investigated in IBD, but the drugs are not (yet) listed in the Dutch protocol of gastroenterology for usage in IBD patients

<table>
<thead>
<tr>
<th>IBD risk SNPS</th>
<th>IBD candidate gene</th>
<th>Target drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11168249</td>
<td>VDR</td>
<td>Vitamin D analogues (8 drugs)</td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2A, FCGR2B, FCGR3A, FCGR3B, IFGAL</td>
<td>Efilizumab</td>
</tr>
<tr>
<td>rs6025</td>
<td>SELP</td>
<td>Heparin</td>
</tr>
<tr>
<td>rs7134599</td>
<td>PTGS2</td>
<td>Thalidomide</td>
</tr>
<tr>
<td>rs10798069</td>
<td>IFNG</td>
<td>Glucosamine</td>
</tr>
<tr>
<td>rs3740415</td>
<td>IFNGR2</td>
<td>Pentostatin</td>
</tr>
<tr>
<td>rs6017342</td>
<td>IFNGR2</td>
<td>Fontolizumab</td>
</tr>
<tr>
<td>rs1077773</td>
<td>IL12B</td>
<td>Interferon gamma-1b</td>
</tr>
<tr>
<td>rs4246215</td>
<td>Ustekinumab</td>
<td>Ustekinumab</td>
</tr>
<tr>
<td>rs4899554</td>
<td>ADA</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs6871626</td>
<td>AHR</td>
<td>Aldesleukin</td>
</tr>
<tr>
<td>rs9557195</td>
<td>FADS1</td>
<td>Interferon Alfa-2a, Recombinant</td>
</tr>
<tr>
<td>rs6142618</td>
<td>FOS</td>
<td>Thalidomide</td>
</tr>
<tr>
<td>rs6927022</td>
<td>GPR18</td>
<td>Interferon gamma-1b</td>
</tr>
<tr>
<td>rs11879191</td>
<td>HLA-DRB1</td>
<td>Glucosamine</td>
</tr>
<tr>
<td>rs1077773</td>
<td>ICAM1</td>
<td>Interferon Alfa-2a, Recombinant</td>
</tr>
<tr>
<td>rs7404095</td>
<td>PTK2B</td>
<td>Thalidomide</td>
</tr>
<tr>
<td>rs2284553</td>
<td>IFNAR1, IFNAR2</td>
<td>Fontolizumab</td>
</tr>
<tr>
<td>rs7134599</td>
<td>IFNG</td>
<td>Interferon gamma-1b</td>
</tr>
<tr>
<td>rs2284553</td>
<td>IFNGR2</td>
<td>Ustekinumab</td>
</tr>
<tr>
<td>rs6871626</td>
<td>IL12B</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs12722515</td>
<td>IL2RA</td>
<td>Thalidomide</td>
</tr>
<tr>
<td>rs2472649</td>
<td>IL8</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs10758669</td>
<td>JAK2</td>
<td>Thalidomide</td>
</tr>
<tr>
<td>rs1569723</td>
<td>MMP9</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs3774959</td>
<td>NFKB1, NFKB2</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs2945412</td>
<td>NOS2</td>
<td>Efalizumab</td>
</tr>
<tr>
<td>rs10798069</td>
<td>PLA2G4A</td>
<td>Quinacrine</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>rs11064881</td>
<td>PRKAB1</td>
<td>Metformin</td>
</tr>
<tr>
<td>rs7404095</td>
<td>PRKCB</td>
<td>Ellagic Acid, Vitamin E</td>
</tr>
<tr>
<td>rs10798069</td>
<td>PTGS2</td>
<td>6 different drugs</td>
</tr>
<tr>
<td>rs6025</td>
<td>Selp</td>
<td>Nadroparin, Dalteparin</td>
</tr>
</tbody>
</table>
Appendix 5: 25 drugs which are used for treatment or investigated in other inflammatory disorders and could also be effective in IBD

<table>
<thead>
<tr>
<th>IBD risk SNPs</th>
<th>IBD candidate gene</th>
<th>Target drugs</th>
<th>Diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10758669</td>
<td>JAK2</td>
<td>Ruxolitinib</td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>rs10798069</td>
<td>PTGS2</td>
<td>Etoricoxib</td>
<td></td>
</tr>
<tr>
<td>rs917997</td>
<td>IL1R1</td>
<td>Anakinra</td>
<td></td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2A, FCGR2B, FCGR3A, FCGR3B</td>
<td>Intravenous Immunglobulin</td>
<td></td>
</tr>
<tr>
<td>rs2284553</td>
<td>IFNAR1</td>
<td>Alemtuzumab, Muromonab</td>
<td></td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2A, FCGR2B, FCGR3A, FCGR3B</td>
<td>Interferon beta-1b, Natural alpha interferon</td>
<td></td>
</tr>
<tr>
<td>rs653178</td>
<td>ALDH2</td>
<td>Alefacept</td>
<td>Psoriasis</td>
</tr>
<tr>
<td>rs2188962</td>
<td>IL5</td>
<td>Disulfiram</td>
<td></td>
</tr>
<tr>
<td>rs3774959</td>
<td>NFKB1</td>
<td>Pranlukast</td>
<td>Asthma</td>
</tr>
<tr>
<td>rs2266959</td>
<td>MAPK1</td>
<td>Arsenic trioxide</td>
<td></td>
</tr>
<tr>
<td>rs3091316</td>
<td>CCL11</td>
<td>CAT-213</td>
<td></td>
</tr>
<tr>
<td>rs3091316</td>
<td>CCL2</td>
<td>Danazol</td>
<td></td>
</tr>
<tr>
<td>rs2188962</td>
<td>IL3</td>
<td>Amlexanox</td>
<td></td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2B</td>
<td>Antithymocyte globulin</td>
<td></td>
</tr>
<tr>
<td>rs11150589</td>
<td>ITGAL</td>
<td>Interferon Alfa-2b, Recombinant</td>
<td></td>
</tr>
<tr>
<td>rs2284553</td>
<td>IFNAR1, IFNAR2</td>
<td>Alfalcacidol</td>
<td></td>
</tr>
<tr>
<td>rs11168249</td>
<td>VDR</td>
<td>Ibudilast</td>
<td></td>
</tr>
<tr>
<td>rs3024505</td>
<td>IL10</td>
<td>Triflusal</td>
<td></td>
</tr>
<tr>
<td>rs3774959</td>
<td>NFKB1</td>
<td>Carvedilol</td>
<td></td>
</tr>
<tr>
<td>rs2945412</td>
<td>NOS2</td>
<td>INCB9471, AMD-070</td>
<td></td>
</tr>
<tr>
<td>rs6025</td>
<td>SELE</td>
<td>O-Phosphoethanolamine</td>
<td></td>
</tr>
<tr>
<td>rs113010081</td>
<td>CCR5</td>
<td>Abciximab</td>
<td></td>
</tr>
<tr>
<td>rs7404095</td>
<td>PRKCB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2A, FCGR2B, FCGR3A, FCGR3B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 6: 26 drugs which have a promising working mechanism for IBD, which makes them potential IBD drugs.

<table>
<thead>
<tr>
<th>IBD risk SNPs</th>
<th>IBD candidate gene</th>
<th>Target drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11168249</td>
<td>VDR</td>
<td>Vitamin D analogues (3 drugs)</td>
</tr>
<tr>
<td>rs2945412</td>
<td>NOS2</td>
<td>Nitric oxide synthase pathway (8 drugs)</td>
</tr>
<tr>
<td>rs113010081</td>
<td>CCR2</td>
<td>INCB3284, MLN-120</td>
</tr>
<tr>
<td>rs3774959</td>
<td>NFKB1</td>
<td>CC-8490, SGN-30</td>
</tr>
<tr>
<td>rs1077773</td>
<td>AHR</td>
<td>Nimodipine</td>
</tr>
<tr>
<td>rs1801274</td>
<td>FCGR2A, FCGR2B, FCGR3A, FCGR3B</td>
<td>Cetuximab</td>
</tr>
<tr>
<td>rs6142618</td>
<td>HCK</td>
<td>Bosutinib</td>
</tr>
<tr>
<td>rs2284553</td>
<td>IFNAR1, IFNAR2</td>
<td>Interferon alfa-n1</td>
</tr>
<tr>
<td>rs6871626</td>
<td>IL12B</td>
<td>humanized SMART Anti-IL-12 Antibody</td>
</tr>
<tr>
<td>rs2188962</td>
<td>IL13</td>
<td>Cintredekin Besudotox</td>
</tr>
<tr>
<td>rs7657746</td>
<td>IL2</td>
<td>WX-G250</td>
</tr>
<tr>
<td>rs12722515</td>
<td>IL2RA</td>
<td>Denileukin diftitox</td>
</tr>
<tr>
<td>rs3740415</td>
<td>NFKB2</td>
<td>Adenosine monophosphate</td>
</tr>
<tr>
<td>rs11064881</td>
<td>PRKAB1</td>
<td>Pomalidomide, Licofelone</td>
</tr>
<tr>
<td>rs10798069</td>
<td>PTGS2</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 7: Potential new drugs for PSC patients

**Diseases investigated in PSC or other inflammatory disorders:**
- Liver transplantation
- T cell lymphoma
- Rhinitis
- Renal cell carcinoma
- Rheumatoid arthritis
- Psoriasis
- Prevention renal rejection

**Target drugs:**
- Aldesleukin, Daclizumab
- Basiliximab, Denileukin difitox, Pseudoephedrine, WX-250, Etanercept, Eflizumab
- Antithymocyte globulin, LFAV03 (experimental), Abatacept, Galiximab, Belatacept
- Basiliximab, Bevacizumab, Denileukin difitox, Efalizumab, Anaplastic lymphoma, Galiximab, Belatacept, Ruxolitinib, Tolacitinib, Tolactinib, Myelofibrosis, Aphthous stomatitis

**Genes in direct PPI:**
- IL2RA
- IL2
- TNFRSF14
- CD226
- CD80/CD86
- JAK1
- CD95/CD95L
- Etanercept
- Galiximab
- Tofacitinib
- Belatacept
- Abatacept, Galiximab
- Denileukin difitox

**PSC candidate genes:**
- IL2RA
- IL2
- TNFRSF14
- CD226
- CD80/CD86
- JAK1
- CD95/CD95L
- Etanercept
- Galiximab
- Tofacitinib
- Belatacept
- Abatacept, Galiximab
- Denileukin difitox

**PSC candidate gene:**
- IL2RA
- IL2
- TNFRSF14
- CD226
- CD80/CD86
- JAK1
- CD95/CD95L
- Etanercept
- Galiximab
- Tofacitinib

**PSC risk SNPs:**
- rs4147359
- rs1340064
- rs3748816
- rs3788087
- rs2260966
- rs4147359
- rs1340064
### Meetings:

<table>
<thead>
<tr>
<th><strong>Meeting</strong></th>
<th><strong>Explanation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday Morning Meeting</td>
<td>Weekly meeting of the Department of Genetics in which the research section gave presentations of their research progress.</td>
</tr>
<tr>
<td>IBD group meeting</td>
<td>Weekly meeting within the research group I was part of to discuss the progress of my team members.</td>
</tr>
<tr>
<td>Journal Club</td>
<td>Weekly meeting in which an article was explained by a PhD student.</td>
</tr>
<tr>
<td>Genetics broad meeting</td>
<td>Weekly meeting of the entire Department of Genetics.</td>
</tr>
<tr>
<td>Molecular Medicine Seminars</td>
<td>Weekly meetings of fundamental research projects.</td>
</tr>
<tr>
<td>Clinical IBD meeting</td>
<td>Weekly meeting in which IBD patients were discussed within the clinical IBD team.</td>
</tr>
<tr>
<td>Clinical endoscopic meeting</td>
<td>Weekly meeting in which endoscopic movies of patients were analysed.</td>
</tr>
<tr>
<td>3GI meeting</td>
<td>Monthly meeting in which PhD students gave an overview of their PhD.</td>
</tr>
</tbody>
</table>

### Presentations

<table>
<thead>
<tr>
<th><strong>Presentation</strong></th>
<th><strong>Explanation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday morning meeting</td>
<td>I presented my final results during this meeting.</td>
</tr>
<tr>
<td>NVGE congress</td>
<td>I submitted an abstract to this congress in Veldhoven and the abstract is admitted and I have to present during this congress on the 19\textsuperscript{th} of March.</td>
</tr>
</tbody>
</table>

### Other projects

<table>
<thead>
<tr>
<th><strong>Project</strong></th>
<th><strong>Explanation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Designing protocols for collecting feces from Iran, Japan and India</td>
<td>During my scientific internship I developed a protocol in which doctors can write information about their patients who are participating in research.</td>
</tr>
</tbody>
</table>