

University of Groningen

## Clinical and genetic factors associated with disease course in inflammatory bowel disease

Spekhorst, Lieke Maaïke

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Spekhorst, L. M. (2018). *Clinical and genetic factors associated with disease course in inflammatory bowel disease*. Rijksuniversiteit Groningen.

**Copyright**

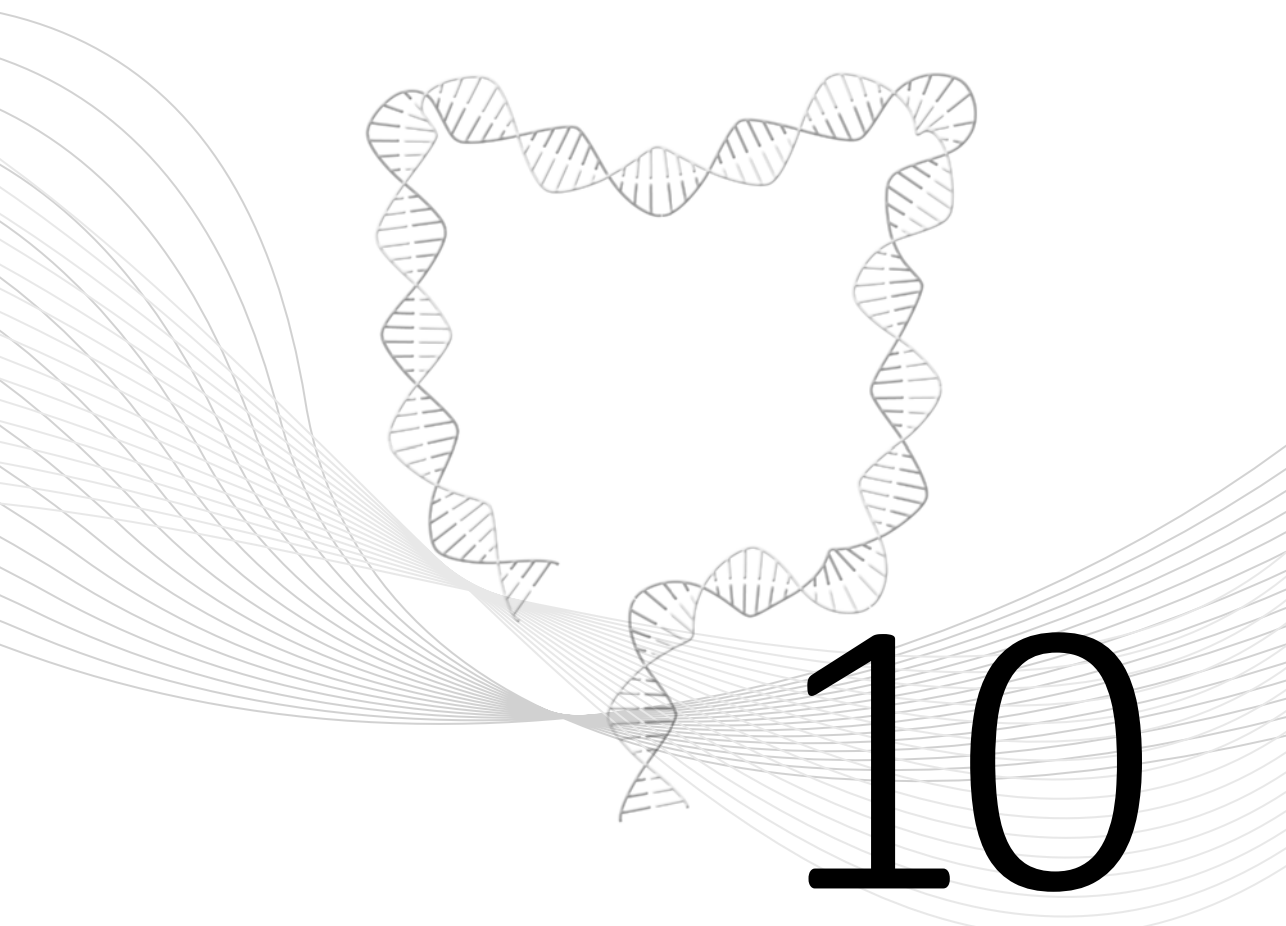
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



## Immunogenicity to anti-TNF $\alpha$ : a non-HLA association

LM Spekhorst, DJ Buurman, T Billiet, A Vich Villa, EAM Festen, G Dijkstra\*,  
S Vermeire\*, RK Weersma\*

\*Authors contributed equally

*Manuscript in preparation*

## Abstract

**Introduction:** Some patients treated with anti-TNF $\alpha$  develop anti-drug antibodies (ADAs), and these can result in a loss of response. Because immunogenicity to anti-TNF $\alpha$  plays a role in the loss of response to anti-TNF $\alpha$  therapy, our aim was to replicate known HLA regions and identify novel (non-HLA) genetic regions associated with the development of ADAs in patients with inflammatory bowel disease (IBD).

**Methods:** We used two independent cohorts. The UMCG cohort (Netherlands) consisted of 61 IBD patients with ADAs and 163 patients without ADAs. The Leuven cohort (Belgium) consisted of 77 IBD patients with ADAs and 115 patients without ADAs. ADAs were measured using two different in-house-developed ELISA assays. Imputation of the HLA-region was performed with the HIBAG package. ImmunoChip genome-wide imputation was done using the Michigan imputation server. For the association analysis a cut-off P value of  $< 0.05$  was used. For the genome-wide association meta-analysis a P value of  $< 1.0 \times 10^{-8}$  was used.

**Results:** The HLA-DQA1\*05 allele was replicated ( $P = 1.0 \times 10^{-3}$ ) in 121 IBD patients positive for ADAs to anti-TNF $\alpha$  and 239 IBD patients negative for ADAs. We were not able to replicate a previously documented association with the HLA-DRB1\*03 allele ( $P = 0.40$ ). Our genome-wide association meta-analysis for 121 IBD patients with ADAs to anti-TNF $\alpha$  and 239 IBD patients negative for ADAs showed eight suggestive association signals ( $P$  values  $< 10 \times 10^{-5}$ ).

**Conclusion:** We replicated the HLA-DQA1\*05 allele associated with anti-TNF $\alpha$  ADAs and identified eight suggestive association signals in non-HLA regions that need to be replicated in larger cohorts. Our results suggest that immunogenicity studies might be useful as biomarkers to predict individual treatment response to anti-TNF $\alpha$  therapy.

## Introduction

Inflammatory Bowel Disease (IBD) is a chronic inflammatory disease of the gastrointestinal tract. Patients with a severe disease phenotype—one which is unresponsive to immunomodulators—often require therapy with so-called biologicals: drug products derived from biological sources. Anti-tumour necrosis factor alpha (anti-TNF $\alpha$ ) agents such as infliximab and adalimumab are highly effective in most patients, but they are expensive and their side-effects can be severe. Furthermore, approximately 25% of patients do not show response after induction therapy (primary non-responders) and 23-46% of patients, after initially responding to induction therapy, experience loss of response within one year (secondary non-responders).<sup>1</sup> While there are several causes of this loss of response to anti-TNF $\alpha$  treatment, the primary cause is formation of anti-drug antibodies (ADAs).<sup>2</sup> Strategies to prevent ADAs are scheduled dosing to maintain stable trough drug levels and co-administration of an immunomodulator (e.g. thiopurines).<sup>3</sup> However, we are currently unable to predict or differentiate which patients are at risk of developing ADAs to anti-TNF $\alpha$ .

As the adaptive immune system plays an important role in the development of ADAs, it is feasible that genetic factors play a role in this immunogenicity; however data on genetic associations with ADA development are scarce. A large study from the United Kingdom assessed immunogenicity in 1284 CD patients who had been using anti-TNF $\alpha$  for at least 12 months. They identified a genome-wide association signal on chromosome 6 and demonstrated that this signal was driven by HLA-DQA1\*05.<sup>4</sup> Another study from Leuven, Belgium assessed immunogenicity in 76 IBD patients who developed antibodies to infliximab and compared them to 116 matched IBD controls who did not develop antibodies during at least two years of infliximab maintenance therapy. These authors showed that HLA-DRB1\*03 was associated with ADAs in IBD patients treated with infliximab.<sup>5</sup> Interestingly the HLA-DQA1\*05 allele is correlated with the HLA-DRB1\*03 allele. The HLA-DQA1\*05 signal could thus be driving the HLA-DRB1\*03 association signal. This highlights the need for additional studies to increase our power to discover associated variants. At the same time, we need better predictors of response to anti-TNF $\alpha$  therapy in order to enable tailored therapy for individual patients, which would both improve patient care and lower costs. The aim of this study was thus twofold: 1) to identify and confirm HLA regions (HLA-DQA1\*05 and HLA-DRB1\*03) associated with development of ADAs to anti-TNF $\alpha$  in IBD patients, and 2) to identify novel (non-HLA) genetic regions associated with the development of ADAs.

## Methods

### Cohort description and sample selection

In this study we used two independent cohorts. The first was collected at the University Medical Center Groningen (UMCG), the Netherlands, and consisted of 61 IBD patients positive for ADAs to infliximab or adalimumab and 163 patients negative for ADAs. ADAs were measured at one time-point during treatment with an in-house-developed ELISA (Sanquin, Amsterdam).

The second cohort was collected at the University Hospitals of Leuven, Belgium and consisted of 77 IBD patients positive for ADAs to infliximab and 115 patients negative for ADAs. Patients positive for ADAs developed ADAs during infliximab treatment. Patients negative for ADAs had been using infliximab for a minimum of two years and had not developed ADAs. ADAs were measured with an in-house-developed ELISA (KU, Leuven).

### Genotyping

All patients were genotyped using the ImmunoChip: a custom-made genotyping array containing ~200,000 genetic variants for 12 auto-immune mediated diseases derived from genome-wide association analysis.<sup>6</sup>

### Genotype calling and quality control

Genotype calling was performed per batch separately with OptiCall using default settings.<sup>7</sup> Quality control was performed using PLINK.<sup>8</sup> Samples with a call rate  $\leq 0.99$  were excluded from the datasets. Variants with a call rate  $\leq 0.99$  or a minor allele frequency (MAF)  $\leq 0.001$  and variants with a Hardy-Weinberg equilibrium P value  $\leq 0.0001$  were also excluded. To identify ancestry outliers, samples were merged with 1000 Genomes Project data and outliers identified using multidimensional scaling plots. These ancestry outliers were then removed from the dataset. To prepare the data for imputation, we used GenotypeHarmonizer<sup>9</sup> and VCFtools.<sup>10</sup> Strands were aligned and identifiers updated to the 1000 Genomes data in genome build GRCh37. The datasets were then combined. Duplicates and related samples were identified and removed.

### HLA imputation and association analysis

Imputation of the HLA-region was performed using the HIBAG package<sup>11</sup> with Illumina ImmunoChip hg19 (2-digit resolution) as a reference panel. Association analysis was performed by ANOVA test using an additive model. A P value  $< 0.05$  was considered to be statistically significant.

For the replication of HLA-DQA1\*05 both cohorts (UMCG and Leuven) were combined, with a total of 121 IBD patients positive for ADAs and 239 IBD patients negative for ADAs to anti-TNF $\alpha$ .

For the replication of HLA-DRB1\*03 only the UMCG cohort was assessed, with 56 IBD patients positive for ADAs and 143 IBD patients negative for ADAs to anti-TNF $\alpha$ .

### **ImmunoChip imputation and quality control**

ImmunoChip genome-wide imputation was done using the Michigan imputation server.<sup>12</sup> Data were first pre-phased with SHAPEIT2,<sup>13</sup> then imputed using IMPUTE2<sup>14</sup> with Haplotype Reference Consortium (r1.1) as the reference panel. An imputation info score > 0.80 was applied.

Stringent quality control was performed. Using default settings in VCFtools, all variants with a call rate of  $\leq 0.99$  or MAF  $\leq 0.001$  and variants with a Hardy-Weinberg equilibrium P value  $\leq 0.0001$  were removed. Duplicated and related samples were also removed.

After these filtering steps, the final UMCG cohort contained 56 samples positive for ADAs and 143 samples negative for ADAs and genotype information for 1,332,743 variants. The final Leuven cohort contained 65 samples positive for ADAs and 96 samples negative for ADAs and genotype information for 1,189,042 variants.

### **Genome-wide association meta-analysis**

A genome-wide association meta-analysis using inverse-variance weighting was carried out, combining the two cohorts. Association testing was performed using PLINK. Genetic variants were selected if they had a P value <  $10^{-5}$  in the meta-analysis, an association signal in the same direction, and a statistically significant P value of < 0.05 in both cohorts. Lead single nucleotide polymorphisms (SNPs) were determined by pairwise linkage disequilibrium using SNP Annotation and Proxy Search (SNAP).<sup>15</sup> A P value <  $10^{-8}$  was considered genome-wide significant and a P value <  $10^{-5}$  was considered suggestive.

### **Functional annotation of SNPs and genes**

Locuszoom was used to construct regional association plots.<sup>16</sup> Lead SNPs (P value <  $10^{-5}$ ) identified with SNAP were selected for further characterization by looking into their association with complex traits and effect on gene expression (eQTL). The Genome Aggregation Database (gnomAD) was used for the functional annotation of SNPs,<sup>17</sup> as were the results from the Genotype-Tissue Expression (GTEx) Consortium, for which multiple tissue eQTL are available.<sup>18</sup> In addition, we used an in-house-developed tool called GeneNetwork.<sup>19</sup> GeneNetwork uses results from the Gene Expression Omnibus microarray to predict pathways and function against other biological databases such as the Gene Ontology and Reactome. Ensembl was used to annotate genes and to predict regulatory functions.<sup>20</sup>

## Results

### HLA regions associated with ADAs to anti-TNF $\alpha$

The genome-wide association meta-analysis was carried out by combining the two cohorts, with a total of 121 IBD patients positive for ADAs and 239 IBD patients negative for anti-TNF $\alpha$  ADAs. This meta-analysis showed a suggestive association signal for variant 6:32608610 located in the *HLA-DQA1* locus ( $P = 2.3 \times 10^{-5}$ , Odds ratio (OR) = 2.09). Furthermore, there was a suggestive association signal for variant 6:32566021 ( $P = 5.5 \times 10^{-5}$ , OR = 1.98), which is located in an mRNA, AF522251, that is involved in the expression of *HLA-DQB1/DRB1* (Table 1A).

For the replication of HLA-DQA1\*05, both cohorts (UMCG and Leuven) were combined for a total of 121 IBD patients positive for anti-TNF $\alpha$  ADAs and 239 IBD patients negative for these ADAs. The HLA-DQA1\*05 allele ( $P = 1.0 \times 10^{-3}$ ) was associated with the development of ADAs, thus replicating the HLA-DQA1\*05 association found by Sazonovs *et al* (Table 1B).

For the replication of HLA-DRB1\*03, only the UMCG cohort was assessed, with 56 IBD patients positive for anti-TNF $\alpha$  ADAs and 143 IBD patients negative for these ADAs. The HLA-DRB1\*03 allele was not associated with ADAs in IBD patients (UMCG patients) treated with infliximab or adalimumab ( $P = 0.40$ ), thus we were not able to replicate the HLA-DRB1\*03 association found by Billiet *et al* (Table 1C).

### Non-HLA regions associated with ADAs to anti-TNF $\alpha$

A genome-wide association meta-analysis was performed for 121 IBD patients positive for anti-TNF $\alpha$  ADAs and 239 IBD patients negative for these ADAs. We did not identify any signals at genome-wide significance level ( $P$  value  $< 1.0 \times 10^{-8}$ ), but did identify eight lead SNPs with a suggestive  $P$  values  $< 10 \times 10^{-5}$ , and these are listed along with their candidate genes in Table 2.

Genetic variant rs1984590 ( $P = 3.6 \times 10^{-5}$ , OR = 0.43) is located in the *FAM107B* gene (Figure 1), which is involved in Toll Like Receptor (TLR) cascades ( $P = 3.9 \times 10^{-4}$ ) and in the TGF- $\beta$  receptor signalling pathway ( $P = 1.3 \times 10^{-3}$ ).<sup>19</sup> Another interesting non-HLA signal is genetic variant rs11167828 ( $P = 7.6 \times 10^{-5}$ , OR = 2.18), which is located near the minor histocompatibility protein HB-1 (*HMH1*) gene (Figure 2). This protein is a precursor of the minor histocompatibility antigen HB-1 (mHag HB-1), an immunogenic peptide that is presented on the cell surface by MHC class I HLA-B44 and that can generate an immune response after recognition by specific T cells. A Locus plot of the other lead SNPs and their candidate genes can be found in **Supplementary File 1**.

**Table 1** HLA regions associated with anti-drug antibodies to anti-TNF $\alpha$ .

**A:** Results from the genome-wide association meta-analysis by combining UMCG and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNF $\alpha$  antibodies.

Chr	BP Position	SNP	A1	Meta-analysis P value	OR	Q	I	UMCG P value	OR	Leuven P value	OR	INFO score	Candidate Genes
6	32608610	6:32608610	C	$2.31 \times 10^{-5}$	2.0939	0.473	0	0.01081	1.853	0.0004482	2.381	0.91	HLA-DQA1
6	32566021	6:32566021	G	$5.53 \times 10^{-5}$	1.9422	0.9092	0	0.004688	1.907	0.003573	1.98	0.93	mRNA AF522251

Candidate genes are identified by one of the gene prioritization methods (GnomAd, GTEX, GeneNetwork and Ensemble).

BP: Basepair; Chr: chromosome; OR: odds ratio; SNP: single nucleotide polymorphisms; A1: the effect (OR) with respect to the A1 allele; I: Heterogeneity I<sup>2</sup> percentage; Q: P value for Cochran's Q statistic; INFO score: imputation score, cut off > 0.80.

**B:** Results for the association analyses for the HLA-DQA1\*05 allele by combining UMCG and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNF $\alpha$  antibodies.

HLA-DQA1*05	Number of allele carriers (ADA +) (%)	Number of allele carriers (ADA -) (%)	P value	Coefficient estimate
	64 (53%)	82 (34%)	$1.0 \times 10^{-3}$	0.15

Heterozygous allele carriers occur twice, and homozygous allele carriers occur only once.

ADA: anti-drug antibody; IBD: inflammatory bowel disease.

**C:** Results for the association analyses for the HLA-DRB1\*03 allele in het UMCG cohort, with 56 IBD patients positive for the development of antibodies and 143 IBD patients negative for anti-TNF $\alpha$  antibodies.

HLA-DRB1*03	Number of allele carriers (ADA +) (%)	Number of allele carriers (ADA -) (%)	P value	Coefficient estimate
	15 (21%)	30 (27%)	0.40	0.08

Heterozygous allele carriers occur twice, and homozygous allele carriers occur only once.

ADA: anti-drug antibody; IBD: inflammatory bowel disease.

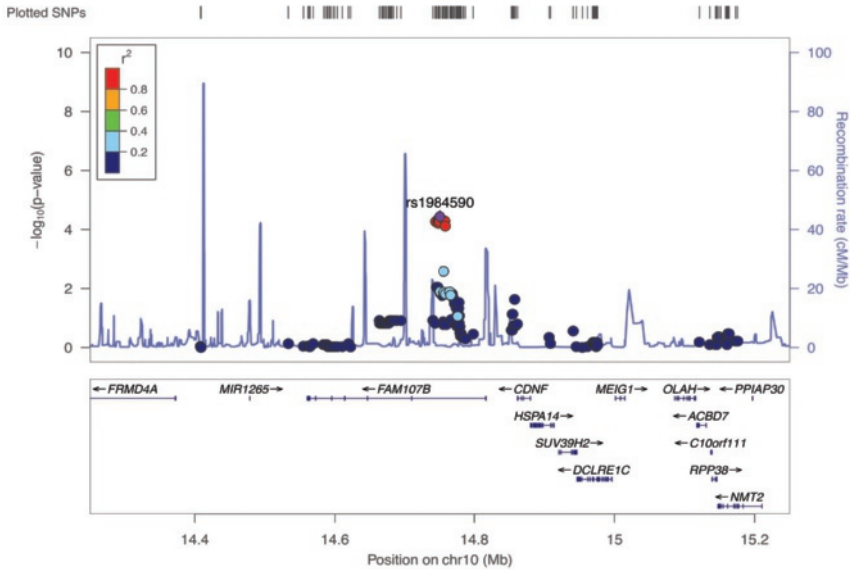


**Table 2** Results from the genome-wide non-HLA association meta-analysis by combining UMCg and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNF $\alpha$  antibodies.

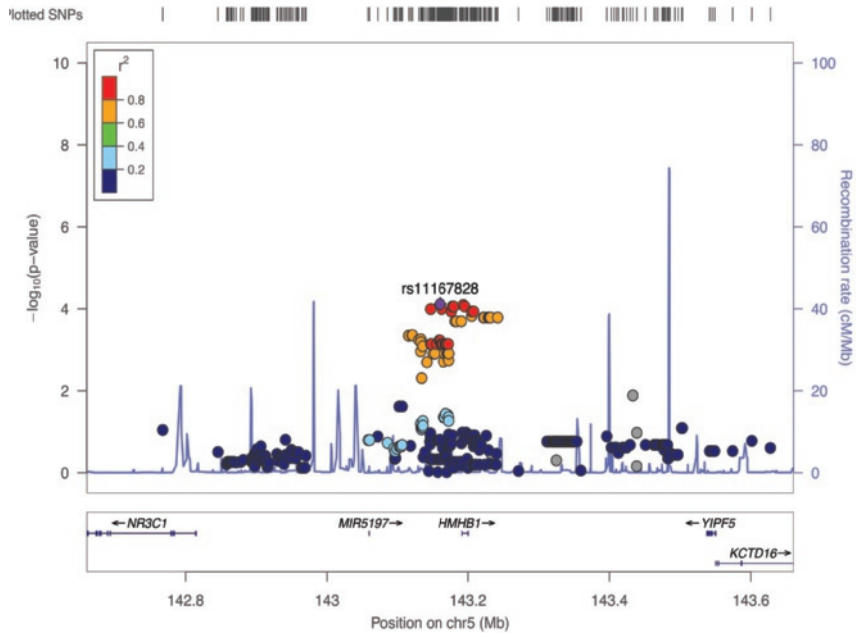
Chr.	BP Position	SNP	A1	Meta-analysis P value	OR	Q	I	UMCG P value	OR	Leuven P value	OR	INFO score	Candidate Genes
11	114847733	rs10736477	G	$1.16 \times 10^{-5}$	0.4627	0.9525	0	0.001219	0.458	0.002431	0.4677	0.82	-
3	37620481	rs6419833	A	$1.69 \times 10^{-5}$	2.0545	0.6335	0	0.004063	1.91	0.001009	2.242	0.883	<i>ITGA9</i>
10	14750252	rs1984590	A	$3.65 \times 10^{-5}$	0.4318	0.2372	28.44	0.02199	0.5367	0.0001761	0.331	0.91	<i>FAM107B</i>
2	120043214	rs3795891	G	$4.57 \times 10^{-5}$	0.5012	0.4958	0	0.000833	0.4437	0.01238	0.5591	0.98	<i>C2orf76</i>
3	179795773	rs6771471	T	$5.07 \times 10^{-5}$	2.0417	0.7538	0	0.007005	1.934	0.001983	2.16	0.81	-
5	140780321	rs6880273	A	$5.71 \times 10^{-5}$	2.0527	0.9403	0	0.004337	2.026	0.003784	2.081	0.94	<i>PCDHG</i> and subfamily
11	114785756	rs2513473	A	$6.14 \times 10^{-5}$	0.5008	0.9365	0	0.002899	0.4943	0.006283	0.5081	0.88	-
5	143160225	rs11167828	C	$7.65 \times 10^{-5}$	2.1757	0.4381	0	0.01426	1.898	0.001029	2.579	0.90	<i>HMH1</i>

Candidate genes are identified by one of the gene prioritization methods (GnomAd, GTEX, GeneNetwork and Ensemble).

BP: Basepair; Chr: chromosome; OR: odds ratio; SNP: single nucleotide polymorphisms; A1: the effect (OR) with respect to the A1 allele; I: Heterogeneity I<sup>2</sup> percentage; Q: P value for Cochran's Q statistic, INFO score: imputation score, cut off > 0.80.



**Figure 1** Regional association plot of genetic variant rs1984590 showing a suggestive genetic association for ADA formation (purple triangle). Filled in circles are genotyped or imputed SNP's from the ImmunoChip. The colour illustrates linkage disequilibrium with the associated SNP



**Figure 2** Regional association plot of genetic variant rs11167828 showing a suggestive genetic association for ADA formation (purple triangle). Filled in circles are genotyped or imputed SNP's from the ImmunoChip. The colour illustrates linkage disequilibrium with the associated SNP

## Discussion

In this study we aimed to replicate known HLA regions (HLA-DQA1\*05 and HLA-DRB1\*03) associated with the development of ADAs to anti-TNF $\alpha$  in IBD patients, as well as to identify novel (non-HLA) genetic regions associated with the development of these ADAs. We were able to replicate the association of the HLA-DQA1\*05 allele with ADA formation to anti-TNF $\alpha$  in 121 IBD patients positive for ADAs and 239 IBD patients negative for anti-TNF $\alpha$  ADAs, but were not able to replicate the association with the HLA-DRB1\*03 allele described by Billiet *et al.*<sup>5</sup> Our non-HLA genome-wide association meta-analysis in 121 IBD patients with anti-TNF $\alpha$  ADAs and 239 IBD patients negative for these ADAs did not show any signals at genome-wide significance level, but did show eight suggestive association signals in non-HLA regions.

We were able to replicate the association of the HLA-DQA1\*05 allele with ADA formation in our cohort, but not the association of the HLA-DRB1\*03 allele. Although the HLA-DRB1\*03 allele has been associated with ADA formation, it also has been shown to be correlated with the HLA-DQA1\*05 allele. We were not able to distinguish whether the HLA-DQA1\*05 allele might have driven the signal of the HLA-DRB1\*03 allele association, has an independent effect or that they drive the same phenotypic association. Data on genetic associations predicting the development of ADAs to anti-TNF $\alpha$  in IBD patients are scarce. However, a study in multiple sclerosis identified an protective association of the HLA-DRB1\*03:01-DQA1\*05:01-DQB1\*02:01 haplotype associated with the development of ADAs in patients treated with anti-interferon-beta (anti-IFN $\beta$ ).<sup>21</sup>

This is the first study to assess non-HLA genetic regions associated with the development of ADAs to anti-TNF $\alpha$  in IBD patients. Our genome-wide association meta-analysis showed eight suggestive association signals in non-HLA regions. Genetic variant rs1984590 is located in the *FAM107B* gene, which is involved in the pathway of Toll Like Receptor (TLR) cascades and TGF- $\beta$  receptor signalling.<sup>19</sup> Active TGF- $\beta$  binds to its receptor and regulates mucosal immune reactions through the TGF- $\beta$  signalling pathway, playing a role in the pathogenesis of IBD.<sup>22</sup> Another interesting non-HLA signal is genetic variant rs11167828, which is located near the minor histocompatibility protein HB-1 (*HMHBI*) gene. Minor histocompatibility antigens play a role in the adaptive immune system, as has been shown in graft-versus-host disease in HLA-matched allogeneic hematopoietic cell transplantation.<sup>23,24</sup>

Literature on the genetic background of the development of ADAs in IBD patients treated with anti-TNF $\alpha$  is scarce and conflicting. One of the main reasons for this are the small sample sizes caused by a lack of genotype data, missing information about antibody status, and the fact that only a small subgroup of patients develop ADAs. These small sample sizes make it more difficult to identify signals at genome-wide significance level. As the current cohort was underpowered, we also could not identify any signals at a genome-wide significance level. In our study, HLA-haplotypes

were derived from SNP-based imputation, but more information could have been obtained by using next generation sequence technology or by HLA typing through polymerase chain reaction with sequence-specific primers (PCR-SSP). Differences in allele frequencies between populations in the HLA region could contribute to the discrepancy in genetic association signals in the HLA between different populations. Unfortunately, information on differences in allele frequencies between the Belgian and Dutch populations regarding HLA-DRB1\*03 and HLA-DQA1\*05 are not available through <http://www.allelefreqencies.net>. Furthermore, this study also used two different assays and two different time points for measuring ADAs, and this could also have influenced our results.

Understanding the underlying factors that contribute to the development of ADAs to anti-TNF $\alpha$  is important because it would allow us to pre-select those patients who will benefit from treatment. Studies on genetic association with the development of ADAs are very relevant, as their findings could be applied to the development of immunogenicity assays based on patient HLA haplotype. These assays could be used in clinical practice as predictive biomarkers for the development of ADAs, thus allowing personalized treatment for individual patients.

In conclusion, we were able to replicate the HLA-DQA1\*05 allele associated with ADA formation to anti-TNF $\alpha$  in IBD patients. We also identified eight suggestive association signals in non-HLA regions that need to be replicated in larger cohorts. As the field of IBD is moving towards personalised medicine, it is important to conduct studies that focus on immunogenicity because future immunogenicity assays might serve as biomarkers to predict treatment response to anti-TNF $\alpha$  therapy.

## References

1. Pouillon L, Bossuyt P, Peyrin-Biroulet L. Considerations, challenges and future of anti-TNF therapy in treating inflammatory bowel disease. *Expert Opin Biol Ther* 2016;16:1277-90.
2. Danese S, Fiorino G, Reinisch W. Review article: causative factors and the clinical management of patients with Crohn's disease who lose response to anti-TNF $\alpha$  therapy. *Aliment Pharmacol Ther* 2011;34:1-10.
3. Hindryckx P, Novak G, Vande Casteele N, *et al.* Incidence, Prevention and Management of Anti-Drug Antibodies Against Therapeutic Antibodies in Inflammatory Bowel Disease: A Practical Overview. *Drugs* 2017;77:363-77.
4. Sazonovs A, Kennedy NA, Bewshea C, *et al.*; PANTS Investigator Consortium AT. OP013 HLA-DQA1 contributes to the development of antibodies to anti-TNF therapy in Crohn's disease. 2018. <https://www.ecco-ibd.eu/publications/congress-abstract-s/abstracts-2018/item/op013-hla-dqa1-contributes-to-the-development-of-antibodies-to-anti-tnf-therapy-in-crohn-x2019-s-disease-2.html>
5. Billiet T, Vande Casteele N, Van Stappen T, *et al.* Immunogenicity to infliximab is associated with HLA-DRB1. *Gut* 2015;64:1344-5.
6. Jostins L, Ripke S, Weersma RK, *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-24.
7. Shah TS, Liu JZ, Floyd JAB, *et al.* optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* 2012;28:1598-603.
8. Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007;81:559-75.
9. Deelen P, Bonder MJ, van der Velde KJ, *et al.* Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. *BMC Res Notes* 2014;7:901.
10. Danecek P, Auton A, Abecasis G, *et al.* The variant call format and VCFtools. *Bioinformatics* 2011;27:2156-8.
11. Zheng X, Shen J, Cox C, *et al.* HIBAG-HLA genotype imputation with attribute bagging. *Pharmacogenomics J* 2014;14:192-200.
12. McCarthy S, Das S, Kretzschmar W, *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;48:1279-83.
13. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2012;10:5-6.
14. Howie B, Marchini J, Stephens M. Genotype Imputation with Thousands of Genomes. Chakravarti A, ed. *G3: Genes/Genomes/Genetics* 2011;1(6):457-470.
15. Johnson AD, Handsaker RE, Pulit SL, *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;24:2938-9.
16. Pruim RJ, Welch RP, Sanna S, *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336-7.
17. Lek M, Karczewski KJ, Minikel E V, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285-91.
18. GTEx Consortium J, Thomas J, Salvatore M, *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580-5.
19. Gene Network. <http://www.genenetwork.nl/>
20. Zerbino DR, Achuthan P, Akanni W, *et al.* Ensembl 2018. *Nucleic Acids Res* 2018;46:D754-61.
21. Núñez C, Cénit MC, Alvarez-Lafuente R, *et al.* HLA alleles as biomarkers of high-titre neutralising antibodies to interferon- $\beta$  therapy in multiple sclerosis. *J Med Genet* 2014;51:395-400.
22. Ihara S, Hirata Y, Koike K. TGF- $\beta$  in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota. *J Gastroenterol* 2017;52:777-87.

23. Larsen ME, Kornblit B, Larsen MV, *et al.* Degree of Predicted Minor Histocompatibility Antigen Mismatch Correlates with Poorer Clinical Outcomes in Nonmyeloablative Allogeneic Hematopoietic Cell Transplantation. *Biol Blood Marrow Transplant* 2010;16:1370-81.
24. Heinold A, Opelz G, Scherer S, *et al.* Role of Minor Histocompatibility Antigens in Renal Transplantation. *Am J Transplant* 2007;8:95-102.

