Genomic and Expression Analyses Identify a Disease-Modifying Variant for Fibrostenotic Crohn's Disease

Visschedijk, M.C.; Spekhorst, L.M.; Cheng, S.C.; Loo, E.S. van; Jansen, B.H.D.; Blokzijl, T.; ... ; Parelsnoer Inst

Journal of Crohn's and Colitis, 2018, 582–588
doi:10.1093/ecco-jcc/jjy001
Advance Access publication January 19, 2018
Original Article

Original Article

Genomic and Expression Analyses Identify a Disease-Modifying Variant for Fibrostenotic Crohn's Disease

Marijn C. Visschedijk,[a,b] Lieke M. Spekhorst,[a,b] Shih-Chin Cheng,[c] Ellen S. van Loo,[d] B. H. Dianne Jansen,[a] Tjasso Blokzijl,[a,e] Hyunsuk Kil,[f] Dirk J. de Jong,[g,*] Marieke Pierik,[h,*] Jeroen P. W. J. Maljaars,[i,*] C. Janneke van der Woude,[j,*] Adriaan A. van Bodegraven,[k,*] Bas Oldenburg,[l,*] Mark Löwenberg,[m,*] Vincent B. Nieuwenhuijs,[n] Floris Imhann,[a,b] Suzanne van Sommeren,[a,b] Rudi Alberts,[a] Ramnik J. Xavier,[c] Gerard Dijkstra,[a,*] Klaas Nico Faber,[a,e] C. Marcelo Aldaz,[f] Rinse K. Weersma,[a,*,†] Eleonora A. M. Festen[a,b,†]

[a]Department of Gastroenterology and Hepatology, University of Groningen and University Medical Centre Groningen, Groningen, The Netherlands [b]Department of Genetics, University of Groningen and University Medical Centre Groningen, Groningen, The Netherlands [c]Broad Institute of Harvard and MIT, Boston, USA [d]Department of Surgery, University of Groningen and University Medical Centre Groningen, Groningen, The Netherlands [e]Department of Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [f]Department of Epigenetics and Molecular Carcinogenesis, Science Park, The University of Texas M.D. Anderson Cancer Centre, Smithville, USA [g]Department of Gastroenterology and Hepatology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands [h]Division of Gastroenterology and Hepatology, University Medical Centre Maastricht, Maastricht, The Netherlands [i]Department of Gastroenterology and Hepatology, Leiden University Medical Centre, Leiden, The Netherlands [j]Department of Gastroenterology and Hepatology, Erasmus Medical Centre, Rotterdam, The Netherlands [k]Department of Gastroenterology and Hepatology, VU University Medical Centre, Amsterdam, The Netherlands [l]Department of Gastroenterology and Hepatology, University Medical Centre Utrecht, Utrecht, The Netherlands [m]Department of Gastroenterology and Hepatology, Academic Medical Centre, Amsterdam, The Netherlands [n]Department of Surgery, Isala Clinics, Zwolle, The Netherlands

*On behalf of the Dutch Initiative on Crohn and Colitis [ICC] and the IBD pearl of the Parelsnoer Institute.
†These authors contributed equally to this work.

Corresponding author: Professor R. K. Weersma, MD, PhD, University of Groningen and University Medical Center Groningen, Gastroenterology and Hepatology Division, Hanzeplein 1, 9700 RB Groningen, PO Box 30.001, The Netherlands. Email: r.k.weersma@umcg.nl

Abstract

Background and Aims: Crohn's disease [CD] is a chronic inflammatory disease with unpredictable behaviour. More than half of CD patients eventually develop complications such as stenosis, for which they then require endoscopic dilatation or surgery, as no anti-fibrotic drugs are currently available. We aim to identify disease-modifying genes associated with fibrostenotic CD.

Methods: We performed a within-case analysis comparing 'extreme phenotypes' using the Immunochip and replication of the top single nucleotide polymorphisms [SNPs] with Agena Bioscience in two independent case-control cohorts totalling 322 cases with fibrostenotis [recurrent after surgery] and 619 cases with purely inflammatory CD.

Results: Combined meta-analysis resulted in a genome-wide significant signal for SNP rs11861007 [$p = 6.0910^{-11}$], located on chromosome 16, in lncRNA RP11-679B19.1, an lncRNA of unknown function, and close to exon 9 of the *WWOX* gene, which codes for WW domain-containing oxidoreductase. We analysed mRNA expression of *TGF*-β and downstream genes in ileocecal resection material from ten patients with and without the *WWOX* risk allele. Patients carrying the risk allele [A] showed enhanced colonic expression of *TGF*-β compared to patients homozygous for the wild-type [G] allele [$p = 0.0079$].

Conclusion: We have identified a variant in *WWOX* and in lncRNA RP11-679B19.1 as a disease-modifying genetic variant associated with recurrent fibrostenotic CD and replicated this association in an independent cohort. *WWOX* can potentially play a crucial role in fibrostenosis in CD, being positioned at the crossroads of inflammation and fibrosis.

Key Words: Crohn's disease; fibrosis; genetics

1. Introduction

Crohn's sisease [CD] is a chronic inflammatory bowel disease [IBD] with unpredictable disease behaviour; more than half of CD patients eventually develop complications such as fistulae and stenosis.[1] Patients with fibrosis-induced stenosis require endoscopic dilatation or surgery, as no anti-fibrotic drugs are currently available. The estimated lifetime risk for surgery in CD patients is 70%, with fibrostenosis as the most common indication for surgery.[2] Within 18 months after surgery, the endoscopic recurrence rate of fibrostenosis is 40–60%.[3]

Fibrostenosis of the intestine is a result of chronic inflammation: the permanent scarring and the consequent luminal narrowing is induced by continuous events of injury and healing. Finally, the intestine is no longer able to restore the normal histological configuration and excessive deposition of extracellular matrix prevails. This leads to thickening of the intestinal wall, which in turn leads to luminal stricturing, finally causing stenosis and obstruction.

Genetic studies of IBD have identified over 240 risk loci associated with disease development.[4,5] A large genotype–phenotype association study was performed, including 34 819 IBD patients, which identified three loci [*NOD2*, *MHC* and *MST1* 3p21] associated with disease location.[6] However, only a few risk variants are known that influence disease behaviour, such as *SMAD3* with recurrent surgery in CD, *MAGI1* with complicated structuring CD, and *FOXO3*, *XACT* and *IGFBP1-3* with severe disease course in CD.[7–10] Variants associated with disease onset seem not to be associated with disease behaviour, which suggests that the biological pathways that underlie disease behaviour are distinct from those that underlie disease onset.[10] In this study, we aim to identify disease-modifying genes associated with the fibrostenotic phenotype in CD. This will help us gain insight into the biological mechanisms underlying intestinal fibrosis development, which might open new avenues for the treatment of CD.

We performed a within-case analysis comparing 'extreme phenotypes', an approach that provides power to detect associations within disease cohorts that has previously been successfully adopted for other traits, including obesity.[11,12] We used an immune-focused fine mapping chip [Immunochip] with 166 251 single nucleotide polymorphisms [SNPs], in two independent case-control cohorts totalling 322 individuals with multiple [two or more] resections due to confirmed ileal stenosis, which we used as cases, and which we compared to patients with 619 purely inflammatory CD [non-penetrating/non-fibrostenotic with disease duration > 5 years]. Following this approach, we identified and replicated a variant within lncRNA RP11-679B19.1 and in *WWOX*, as a disease-modifier, associated with recurrent fibrostenotic CD. We then studied the potential functional consequences of this variant by studying resection material of CD patients carrying the risk allele.

2. Methods

2.1. Genetic analysis [cohort description, quality control and statistical analysis]

2.1.1. Subjects

The discovery cohort consisted of 521 Dutch CD patients, selected out of ~ 1500 Dutch CD patients. Selection was based on reviewing the medical records of the total cohort of 1500 patients. We selected patients on extreme ends of the fibrostenotic phenotype: a case group and a control group. The case group consisted of 242 patients with 'recurrent fibrostenotic' CD, which was defined as the need for multiple [two or more] resections due to confirmed ileal stenosis. The control group consisted of 279 patients with purely inflammatory disease behaviour CD [Montreal classification B1[13]] defined as non-penetrating, non-fibrostenotic CD with a disease duration of > 5 years, without the need for any kind of surgery. All patients of the discovery cohort were also included in the IIBDGC core-phenotype study.[6]

For the replication phase, patients were selected from the Parelsnoer Institute database [http://www.parelsnoer.org/]. This database consists of 3394 IBD patients [CD: $n = 2118$] collected from the IBD Centres in all eight University Medical Centres in the Netherlands. Patients already included in the discovery phase were excluded. Eighty CD patients with recurrent fibrostenotic disease and 340 patients with purely inflammatory disease were selected with similar criteria. While the phenotypes of the discovery cohort could be checked against the patients' medical charts, regulations prevented us from doing the same with the replication cohort. Hence, the replication cohort might have some non-cases and/or non-controls admixed. All patients were of Central European descent.

2.2.2. Discovery phase

2.1.2.1. Immunochip genotyping

In the discovery phase of the study we used genotype data of ~1500 CD patients described by Jostins *et al*.[14] In brief, DNA was extracted from whole blood. An Illumina Immunochip was used to genotype

the DNA samples. The Immunochip is a custom Illumina Infinium immune-focused fine-mapping chip comprising 196 524 SNPs selected primarily based on genome wide association study [GWAS] analysis of 12 immune-mediated diseases.

2.1.2.2. Quality control

Quality control [QC] was performed using PLINKv1.07 software.[15] QC removed SNPs with a minor allele frequency [MAF] less than 0.001, as well as SNPs with more than 10% missing genotypes. In total, 547 patients were selected [248 fibrostenotic CD patients and 299 CD patients with purely inflammatory disease behaviour]. After removing duplicate individuals [n = 6] and individuals with more than 15% missing genotype data [n = 20], 521 individuals remained [no differential missingness, χ^2 p = 0.67]. After QC, the dataset consisted of 521 individuals and 166 251 SNPs with a genotyping call rate of 99%.

2.1.2.3. Statistical analysis and prioritization of SNPs for replication

After allele association analysis [χ^2 test] and cluster plot inspection, 34 SNPs with a p-value < 1.0×10^{-3} remained. To adequately test each locus for association in the replication phase we selected two SNPs per locus where possible. In addition, we decided to test the SNP that previously showed the strongest association with severe CD disease course [rs12212067, *FOXO3*], although this SNP was not associated with fibrostenotic disease in our discovery cohort.[10,16] In total, 43 SNPs passed the design of two plexes with the Assay Design suite of Agena Bioscience [https://seqpws1.agenacx.com/AssayDesignerSuite.html].

2.1.3. Replication phase

2.1.3.1. Genotyping

Forty-three SNPs were genotyped using the Agena Bioscience Massarray in the replication phase [http://agenabio.com]. During QC we excluded SNPs with a Hardy–Weinberg equilibrium p-value < 0.0001 in controls and an overall call rate < 90% and individuals with < 85% of SNPs confidently genotyped. After QC, the replication cohort consisted of 80 CD cases with recurrent fibrostenotic disease behaviour, 340 CD patients with purely inflammatory disease and 39 SNPs with a genotype call rate of 99%. Since rs11861007 failed the Agena Bioscience design and it was the only Immunochip SNP in the *WWOX* region, genotyping of SNP rs11861007 was performed using Taqman technology [Applied Biosystems].

2.1.3.2. Statistical analysis

Allelic association analysis [χ^2 test] and meta-analysis of the complete cohort was performed with PLINKv1.07 software.[15] To test if disease localization could be a cofounding factor, we additionally performed an allelic association analysis for rs11861007 [χ^2 test] in ileal disease patients [n = 86] vs colonic disease patients [n = 36]. Similarly, allelic association analysis for rs11861007 [χ^2 test] in Montreal Class B2 phenotype patients [n = 165] vs Montreal Class B1 and B3 phenotype patients [776] was performed to determine whether rs11861007 was associated with the Montreal B2 phenotype [and not recurrent fibrostenotic disease]. Logistic regression analyses with disease as a covariate was performed with R statistics.

Given that the prevalence of recurrent stenotic disease in CD was 0.04 in our replication cohort, that the allele frequency of our rs11861007 risk allele A is 0.06 in Caucasian patients of Western and Central European descent (CEU), and that we expected the risk variant to have an odds ratio [OR] of 4.3, we expected to have > 80% power to discover this association in our replication cohort. Ultimately, we only had 60% power to detect the association that we did detect for rs11861007 in our replication cohort at an OR of 2.1 [http://zzz.bwh.harvard.edu/gpc/cc2.html].

2.2.2. SNP annotation

LocusZoom was used to construct regional association plots[17] [Supplementary Figure S1]. Exploration of the linkage disequilibrium [r^2 > 0.8] was performed with Haploview.[18] We assessed whether associated SNPs had known functional consequences or regulatory features. The Encyclopedia of DNA Elements [ENCODE][19] was searched using the UCSC Genome Browser.[20] Specifically, SNPs located in the following regulatory features were searched: DNaseI-hypersensitivity sites, transcription factor binding sites, histone modification and DNA-polymerase sites. We tested whether associated variants [or SNPs in high linkage disequilibrium [r^2 > 0.8]] showed an effect on gene expression levels of genes. Expression quantitative trait loci [eQTL] analysis was performed with the eQTL browser [http://genenetwork.nl/bloodeqtlbrowser/], based on non-transformed peripheral blood in 5311 individuals.[21] Additional enhancer analyses were performed with the Fantom5 enhancer atlas.[22]

2.2.3. *In silico* analysis [rs11861007]

The strongest association for fibrostenotic disease was found in the RP11-679B19.1 lncRNA, and in the *WWOX* gene. The function of the RP11-679B19.1 lncRNA is unknown. To predict in which biological or cellular process and molecular function the *WWOX* gene is involved, we used an in-house developed RNA network tool [http://www.genenetwork.nl].[23] The RNA network uses a method, based on principal component analysis, to build transcriptional profiles for biological pathways, which can be used to predict gene functions. A more detailed description of the method of the RNA network can be found in the paper by Fehrmann *et al.*[23]

2.4. Expression analysis

2.4.1. Resection material of CD patients

2.4.1.1. Subjects

The resection material was collected during surgical resection procedures [samples of the small bowel and colon] from patients for whom we had extensive phenotype data [University Medical Center Groningen]. Twenty-nine patients with fibrostenotic CD behaviour were included. The resection material of CD patients was preserved [storage at −80°C] in three parts: the ileum [proximal of the stenosis], the stenosis itself [medial part] and the colon [distal of the stenosis]. As stenotic tissue is a final stage in the process of fibrosis formation, we considered the stenotic tissue, medial part of the resection, as not representative. Therefore, we included only samples of the ileum and colon, respectively proximal and distal from the stenosis. Pathology records of the resection material were checked and revealed no severe inflammation or fibrosis at the end of the resection sides, proximal and distal, of the stenosis.

2.4.1.2. Genotyping and patient selection

After DNA isolation, genotyping of the *WWOX* SNP [rs11861007] in 71 patients was performed using Taqman technology with a call rate of 96% [Life Technologies]. We selected the five CD patients who carried the *WWOX*-genotype [risk allele, A] as cases [we

found no patients homozygous for the risk allele in our cohort]. We selected five matched CD patients with the GG genotype [wild-type] as controls [matched based on age, disease duration, inflammation and stenosis].

Quantitative polymerase chain reaction [qPCR] was performed of *WWOX*, *TGF-β*, *iNOS*, *IL1-B*, *TNF-α*, *FOXP3*, *α-SMA*, *Collagen Type 1* and downstream genes *PAI-I [SERPINE]* and *CTGF* of the ileocolonic tissue of five CD patients carrying the *WWOX* risk allele and five CD patients homozygous for the *WWOX* wild-type allele [see Online Methods].

3. Results

In this study we included two independent case-control cohorts totalling 322 recurrent fibrostenotic and 619 purely inflammatory CD cases; 60% of the cohort were female with a mean age at diagnosis of 25 years. The cohorts were selected based on 'extreme phenotypes', resulting in a statistically significant difference in disease location and behaviour between patients [Table 1]. After analysis of 166 251 SNPs in 242 fibrostenotic and 279 purely inflammatory CD cases and replication of 34 selected SNPs in an independent cohort of 80 fibrostenotic and 340 purely inflammatory CD cases, the combined meta-analysis resulted in a genome-wide significant signal for SNP rs11861007 [$p = 6.09 \times 10^{-11}$, OR = 3.2, heterogeneity [I^2] < 75%] [Table 2, Supplementary Table S1]. The minor heterogeneity for SNP rs11861007 between the discovery and the replication cohort might be caused by a slight admixture of non-cases and/or non-controls in the replication cohort, as described in the Methods section. To assess whether the association between rs11861007 and fibrostenotic disease was not mainly due to ileal disease localization we tested this marker for association with ileal disease localization. We found no association between rs11861007 and ileal disease localization [$p = 0.27$]. Additional logistic regression analyses for the association between SNP rs11861007 and re-fibrosis, with disease location as a covariate, still showed a significant association [$p = 0.004$]. We also tested for an association between rs11861007 and the Montreal Class B2 phenotype [any stricuring disease] in our cohort, but did not find an association [$p = 0.25$].

3.1. SNP annotation and *in silico* analyses

rs11861007 is located on chromosome 16, in an intron close to exon 9 of the *WWOX* gene, which codes for WW domain-containing oxidoreductase. Neither rs11861007 nor SNPs in high linkage disequilibrium [$r^2 > 0.8$] have known functional or regulatory features [eQTL or regulatory elements assessed in the ENCODE Encyclopedia of DNA Elements or FANTOM5 enhancers].[22] *WWOX* is the only coding gene in the locus [defined as 250 kb on either side of our top hit] making it the most likely positional candidate gene. The SNP is also located in the lncRNA RP11-679B19.1, in which it might affect folding, but for this lnRNA there is also no eQTL effect, and no function is known for RP11-679B19.1.

Co-transcriptional pathway analysis using an in-house-developed RNA network tool[23] predicts *WWOX* involvement with cellular components in the *extracellular matrix compartment* [$p = 1.41 \times 10^{-3}$] and an association with *collagen binding* [$p = 9.78 \times 10^{-5}$]. See Supplementary Table S2 for involvement in other pathways.

3.2. Expression analyses

Normal ileal and colonic tissue residing proximal to and distal from the stenotic part was sampled from five CD patients heterozygous for the *WWOX* risk allele [A] and five CD patients homozygous for the *WWOX* wild-type allele [G]. *WWOX* mRNA levels were similar in both the ileal and the colonic tissue of patients with or without the risk allele [Supplementary Figure S2]. In contrast, *TGF-β* expression was significantly higher in the colonic tissue of risk-allele-carrying patients compared to patients homozygous for the wild-type allele [Mann–Whitney U-test, $p = 0.0079$] [Figure 1A]. Similarly, downstream targets involved in fibrosis, such as Connective Tissue Growth Factor [*CTGF*] and matricellular *PAI-1* [*SERPINE1*], showed a trend towards increased expression in the risk-allele-carrying CD patients compared to the patients homozygous for the wild-type allele [Figure 1B; see Supplementary Figure S2 for all qPCR results].

4. Discussion

In this study, we have identified a variant in the *WWOX* gene as a disease-modifier associated with a recurrent fibrostenotic phenotype in CD, and we have replicated this association in an independent

Table 1. Clinical characteristics of patients in the combined cohort

	Fibrostenotic CD	Purely inflammatory CD	*p*-value
n [%]	322 [100%]	619 [100%]	
Patient characteristics			
Female, *n* [%]	180 [56%]	371 [60%]	0.23
Median age of onset, years [IQR 25–75]	24 [18–30]	26 [19–36]	0.38
Disease location, *n* [%]			
Montreal-L1, ileal	86 [27%]	93 [15%]	0.06
Montreal-L2, colonic	36 [11%]	291 [47%]	<0.0001
Montreal-L3 ileocolonic	200 [62%]	235 [38%]	0.0011
Montreal-L4, additional upper disease localization	32 [10%]	55 [9%]	1
Disease behaviour, *n* [%]			
Montreal-B1, inflammatory	0	619 [100%]	
Montreal-B2, stricturing	165 [51%]	0	
Montreal-B3, penetrating	157 [49%]	0	
Time until surgery			
Disease duration in years from diagnosis until first surgery, mean [SD]	4.1 [6.5]		

This table provides the clinical characteristics of patients with recurrent fibrostenotic and purely inflammatory CD in the combined cohort. Disease location and behaviour is based on the Montreal classification for CD. Chi-squared and Mann-Whitney U-tests [only for age of onset] are used to calculate the *p*-value. CD, Crohn's disease; IQR, interquartile range.

Table 2. Allelic association analyses

					Discovery cohort				Replication cohort				Meta-analysis	
Chromosome	SNP	Position [Hg19]	Candidate gene	Risk allele	Risk allele frequency in cases	Risk allele frequency in controls	*p*-value	OR	Risk allele frequency in cases	Risk allele frequency in controls	*p*-value	OR	*p*-value	OR
16	rs11861007	79238685	*WWOX*	A	0.19	0.05	1.26×10^{-11}	4.3	0.13	0.07	0.01	2.1	6.09×10^{-11}	3.2
22	rs371513	21988599	*CCDC116*	A	0.22	0.13	1.02×10^{-4}	1.9	0.20	0.15	1.47×10^{-1}	1.4	7.97×10^{-5}	1.7
5	rs55965691	593812	*CEP72*	A	0.008	0.06	1.60×10^{-5}	0.14	0.025	0.06	5.97×10^{-2}	0.38	9.37×10^{-5}	0.23
5	rs6883704	599390	*CEP72*	C	0.01	0.06	5.04×10^{-5}	0.18	0.025	0.06	5.97×10^{-2}	0.38	1.01×10^{-4}	0.25
3	rs17033143	10630258	*ATP2B2*	A	0.14	0.07	1.84×10^{-4}	2.2	0.09	0.07	3.24×10^{-1}	1.36	3.51×10^{-4}	1.9

The five most significant allelic association analysis results [χ^2 test] for the 166 251 SNPs in the discovery cohort [242 patients with recurrent fibrostenotic CD vs 279 patients with purely inflammatory CD] are presented. Replication was performed for 39 SNPs in a replication cohort [80 patients with recurrent fibrostenotic CD vs 340 patients with purely inflammatory CD]. Meta-analysis was performed in the combined cohort. CD, Crohn's disease; SNP, single nucleotide polymorphism; OR, odds ratio.

cohort. We found that patients carrying the risk allele show enhanced profibrotic signalling, through higher expression of tumour growth factor-β [TGF-β], in the colon. This suggests that the genetic variant in *WWOX* is associated with a decrease of WWOX function. Alternatively, the effect could be caused by a configuration change of the lnRNA RP11-679B19.1, which has recently been recorded in the same locus and is, yet, of unknown function.

The genetic variants that contribute to disease behaviour can be different from the variants that contribute to disease susceptibility. This has been shown for a variant affecting *FOXO3A* expression, which was found to be significantly associated with disease prognosis but which was not associated with CD susceptibility.[16] In this study, we confirm the concept that disease-modifying genes can differ from variants contributing to disease development. The SNP found to be most closely associated [rs11861007] has not been associated with disease development in previous studies.[4] Previous genotype–phenotype studies in IBD patients have not studied the specific phenotype in the present study: recurring ileal fibrostenosis in CD.[6] Previous studies focusing on genetic variants associated with a single fibrotic event in CD found associations with variants in *SMAD3* and *MAGI*, but due to the diverse coverage of the genotyping platforms we could not replicate these findings.[7,8] We could also not confirm a previously described association between *NOD2* variants and ileal fibrostenotic disease, which might be due to the fact that we studied recurrent fibrostenosis, not necessarily ileal, or because the *NOD2* variants are relatively rare in our cohort.

The SNP [rs11861007] is located in an intron close to exon 9 of the *WWOX* gene. Although we could not find an eQTL effect, regulatory features or enhancer activity, the *WWOX* gene is the only coding gene in the locus [defined as 250 kb on either side of our top hit] making it the most likely positional candidate gene. The SNP is also located in the lncRNA RP11-679B19.1, in which it might affect folding, but for this lnRNA there is also no eQTL effect, and no function is known for RP11-679B19.1. *WWOX* is a known tumour suppressor gene and encodes a protein that contains two WW domains and a short-chain dehydrogenase/reductase domain [SRD]. The mechanism of tumour suppression of *WWOX* involves apoptosis, modulation of the extracellular matrix and modulation of cell bioenergetics[24] Genome-wide association studies have shown that *WWOX* also plays a role in the pathogenesis of pulmonary fibrosis.[25] Targeted deletion of *Wwox* in epithelium in the mammary gland increased fibronectin levels[24] and conditional deletion of *Wwox* in the mammary gland significantly upregulated multiple collagen genes.[26] Moreover, molecular functions predicted by our RNA network tool show that *WWOX* plays a role in fibrosis formation, in agreement with the previously described literature reports.

The main mediator between intestinal inflammation and fibrosis in IBD is TGF-β,[27] which is overexpressed in intestinal tissue in CD patients.[28] Upon TGF-β stimulation, *WWOX* acts as an inhibitor of SMAD3 transcriptional activity by sequestering it in the cytoplasm.[29] We show an enhanced *TGF-β* expression in CD patients carrying the *WWOX* risk allele. TGF-β stimulates downstream signalling pathways resulting in expression of several profibrotic genes, including *CTGF*. qPCR analysis of *CTGF* in this study showed a trend towards increased expression in the *WWOX* risk-allele-carrying CD patients compared to the patients homozygous for the *WWOX* wild-type allele. We conclude that *WWOX* risk-allele-carrying individuals have enhanced TGF-β expression with a trend towards elevated expression of profibrotic genes as a downstream effect.

There are some limitations to this study. First, the cohort size is relatively small. However, by using a within-case analysis comparing

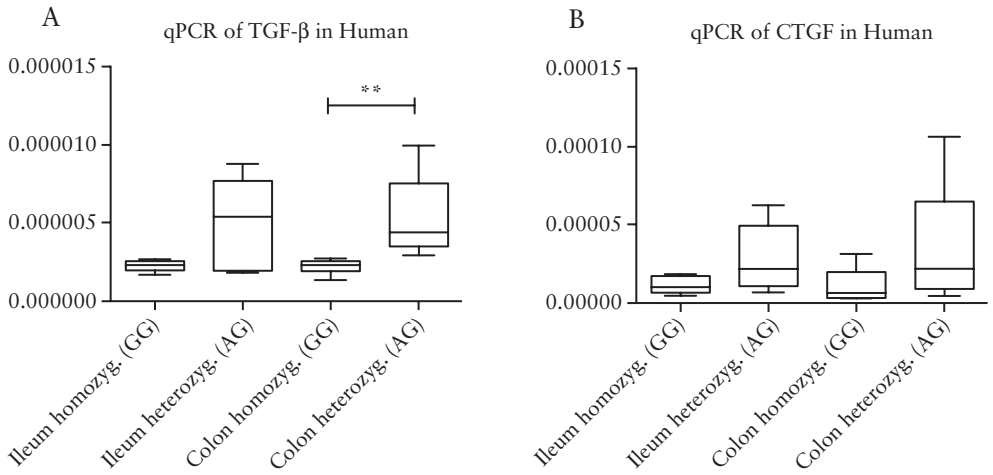

Figure 1. TGF-β + CTGF expression in human and macrophage polarization of *WWOX*. TGF-β is a crucial factor in the equilibrium between inflammation and fibrosis. [A, B] *TGF-β* and *CTGF* [profibrotic downstream gene of TGF-β] expression in the non-stenotic ileocecal resection tissue from five CD patients carrying the risk allele [AG] and five CD patients homozygous for the wild-type allele [GG]. CD, Crohn's disease; qPCR, quantification polymerase chain reaction; TGF-β, transforming growth factor; CTGF, connective tissue growth factor; homozyg, homozygous; heterozyg, heterozygous.

'extreme phenotypes' we increased the power of our analysis. This approach provides power to detect associations within disease cohorts and the approach has been successfully adopted for other traits.[11,12] Moreover, we replicated our findings in an independent cohort. Second, the upregulation of TGF-β in the ileal part of the resection material [upstream of the stenosis] in patients carrying the *WWOX* risk allele was a trend and not statistically significant. However, in the colonic part of the resection material [downstream of the stenosis] in patients carrying the *WWOX* risk allele we do show statistically significant enhanced *TGF-β* expression. All available *WWOX* risk-allele-carrying patients in our centre were included, but because the risk allele has a low frequency this number is quite small, which means our study is relatively underpowered. Increasing the sample size may turn the trend we observe into a significant association. Finally, the WWOX–SMAD3–TGF-β pathway has been described previously[29] and it has been proven that the proteins encoded by the genes interact, although the exact pathways through which they interact have not yet been elucidated, making it difficult to interpret the results from our study.

In conclusion, we have identified and replicated *WWOX* as a disease-modifying gene associated with the recurrent fibrostenotic phenotype in CD. Our expression analyses suggest a functional effect of the risk allele through enhanced expression of *TGF-β* in risk-allele carriers, indicating profibrotic expression. CD patients carrying the *WWOX* risk allele appear to have a profibrotic profile. To avoid fibrotic complications, it might be advisable to refrain from prescribing anti-inflammatory medication that enhances TGF-β signalling in intestinal fibroblasts in these patients.

Funding

This work was supported by several grants. MCV is supported by AGIKO grant (92.003.577) from the Netherlands Organization for Scientific Research (NWO). RKW is supported by a NWO VIDI grant [016.136.308]. EAMF is supported by a Career Development Grant [CD14-04] from the Dutch Digestive Foundation [Maag Lever Darm Stichting]. CMA is supported by NIH/NCI grant R01 CA102444.

Conflict of Interest

The authors declare that they have no competing interests.

Acknowledgments

This study was carried out under the auspices of the Parelsnoer Institute [PSI, http://www.parelsnoer.org], which is part of and funded by the Dutch Federation of University Medical Centres and was initially funded by the Dutch Government [2007–2011]. We thank Dianne Jansen and Tjasso Blokzijl for their expert technical help and Kate McIntyre for editorial assistance.

Author Contributions

Genetic study designs: MCV, EAMF, RKW. Expression analysis design: MCV, EAMF, GD, KNF, MA, RKW. Clinical information collection: MCV, LMS, EVL. Sample collection: MCV, LMS, FI, DJD, MP, LMP, FI, PWJM, CJW, AAB, BO, ML, GD, CW, SS, RKW. Genotyping: MCV, SS, RA. Expression, wet lab: BHJ, TB, HK. Statistical analyses, figures: MCV, BHJ, SCC, FI. Writing: MCV, EAMF, RKW. All authors read, critically revised and approved the final manuscript.

Supplementary Data

Supplementary data are available at *Journal of Crohn's and Colitis* Online.

References

1. Cosnes J, Cattan S, Blain A, *et al*. Long-term evolution of disease behavior of Crohn's disease. *Inflamm Bowel Dis* 2002;**8**:244–50.
2. Munkholm P, Langholz E, Davidsen M, Binder V. Disease activity courses in a regional cohort of Crohn's disease patients. *Scand J Gastroenterol* 1995;**30**:699–706.
3. De Cruz P, Kamm MA, Hamilton AL, *et al*. Crohn's disease management after intestinal resection: a randomised trial. *Lancet* 2014;**6736**:1–11.
4. Liu JZ, van Sommeren S, Huang H, *et al*.; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;**47**:979–86.

5. de Lange KM, Moutsianas L, Lee JC, *et al*. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**:256–61.
6. Cleynen I, Boucher G, Jostins L, *et al*.; International Inflammatory Bowel Disease Genetics Consortium. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;**387**:156–67.
7. Fowler SA, Ananthakrishnan AN, Gardet A, *et al*. SMAD3 gene variant is a risk factor for recurrent surgery in patients with Crohn's disease. *J Crohns Colitis* 2014;**8**:845–51.
8. Alonso A, Domènech E, Julià A, *et al*. Identification of risk loci for Crohn's disease phenotypes using a genome-wide association study. *Gastroenterology* 2015;**148**:794–805.
9. Ananthakrishnan AN, Xavier RJ. How does genotype influence disease phenotype in inflammatory bowel disease? *Inflamm Bowel Dis* 2013;**19**:2021–30.
10. Lee JC, Biasci D, Roberts R, *et al*.; UK IBD Genetics Consortium. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017;**49**:262–8.
11. Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nat Rev Genet* 2009;**10**:872–8.
12. Wang K, Li WD, Zhang CK, *et al*. A genome-wide association study on obesity and obesity-related traits. *PLoS ONE* 2011;6:e18939.
13. Silverberg MS, Satsangi J, Ahmad T, *et al*. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;**19**[Suppl A]:5A–36A.
14. Jostins L, Ripke S, Weersma RK, *et al*.; International IBD Genetics Consortium [IIBDGC]. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**491**:119–24.
15. 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.
16. Lee JC, Espéli M, Anderson CA, *et al*.; UK IBD Genetics Consortium. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* 2013;**155**:57–69.
17. Pruim RJ, Welch RP, Sanna S, *et al*. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;**26**:2336–7.
18. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;**21**:263–5.
19. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) project. *Science* 2004;**306**:636–40.
20. Rosenbloom KR, Dreszer TR, Pheasant M, *et al*. ENCODE whole-genome data in the UCSC Genome Browser. *Nucleic Acids Res* 2010;**38**:D620–5.
21. Westra HJ, Peters MJ, Esko T, *et al*. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;**45**:1238–43.
22. Andersson R, Gebhard C, Miguel-Escalada I, *et al*. An atlas of active enhancers across human cell types and tissues. *Nature* 2014;**507**:455–61.
23. Fehrmann RS, Jansen RC, Veldink JH, *et al*. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet* 2011;**7**:e1002197.
24. Abdeen SK, Salah Z, Khawaled S, Aqeilan RI. Characterization of WWOX inactivation in murine mammary gland development. *J Cell Physiol* 2013;**228**:1391–6.
25. Loth DW, Soler Artigas M, Gharib SA, *et al*. Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nat Genet* 2014;**46**:669–77.
26. Ferguson BW, Gao X, Kil H, *et al*. Conditional Wwox deletion in mouse mammary gland by means of two Cre recombinase approaches. *PLoS One* 2012;**7**:e36618.
27. Rieder F, Fiocchi C. Intestinal fibrosis in inflammatory bowel disease - Current knowledge and future perspectives. *J Crohns Colitis* 2008;**2**:279–90.
28. Burke JP, Ferrante M, Dejaegher K, *et al*. Transcriptomic analysis of intestinal fibrosis-associated gene expression in response to medical therapy in Crohn's disease. *Inflamm Bowel Dis* 2008;**14**:1197–204.
29. Ferguson BW, Gao X, Zelazowski MJ, *et al*. The cancer gene WWOX behaves as an inhibitor of SMAD3 transcriptional activity via direct binding. *BMC Cancer* 2013;**13**:593.