Journal of Crohn's and Colitis (2014) 8, 678–685

Available online at www.sciencedirect.com

ScienceDirect

Genetic susceptibility and genotype–phenotype association in 588 Danish children with inflammatory bowel disease

C. Jakobsen [a,*], I. Cleynen [b], P.S. Andersen [c], S. Vermeire [d], P. Munkholm [e], A. Paerregaard [a], V. Wewer [a]

[a] Department of Paediatrics, Hvidovre University Hospital, Copenhagen, Denmark
[b] Department of Clinical and Experimental Medicine, KU Leuven, Belgium
[c] Department of Microbiology and Infection Control, State Serum Institute, Copenhagen, Denmark
[d] Department of Gastroenterology, University Hospitals Leuven, Belgium
[e] Department of Gastroenterology, Medical Section, Herlev University Hospital, Copenhagen, Denmark

Received 21 August 2013; received in revised form 28 November 2013; accepted 15 December 2013

KEYWORDS
Crohn's disease;
Ulcerative colitis;
Genotype;
Phenotype;
Epidemiology

Abstract

Aim: To investigate the association between known inflammatory bowel disease (IBD)-associated genetic variants and development of paediatric IBD, and specific clinical sub-phenotypes.
Material and methods: In this case–control study we included IBD patients <18 years of age at diagnosis from the Danish National Patient Registry and healthy children <18 years of age were randomly selected from the Danish Central Office of Civil Registration. The latter had filled out a questionnaire regarding health status, and DNA was obtained from blood samples and the buccal mucosa. Patient files were retrieved and clinical information was extracted. DNA was obtained from Guthrie cards from the Danish National Neonatal Screening Biobank (PKU-biobanken) at Statens Serum Institut and from blood samples.
Results: A total of 588 IBD patients (244 Crohn's disease (CD), 318 ulcerative colitis (UC) and 26 IBD-unclassified (IBDU)) and 543 healthy controls were included. We found an association between CD and rs22411880 (*ATG16L1*, odds ratio (OR) = 1.7 [1.1–1.7], p = 0.003), rs5743289 (*NOD2*, OR = 1.4 [1.1–1.9], p = 0.009) and the paediatric specific rs1250550 (*ZMIZ1*, OR = 0.7 [0.5–0.9], p = 0.01). None of the investigated 41 SNPs were associated with disease localisation, medical treatment or surgery after correcting for multiple analyses.
Conclusion: We found an association between CD and three previously published genetic variants and replicated the association with the paediatric specific *ZMIZ1* gene. No Bonferroni corrected significant genotype–phenotype associations were found. For future studies aimed at

* Corresponding author at: Department of Paediatrics, Hvidovre Hospital, Kettegaard Alle 30, 2650 Hvidovre, Denmark. Tel.: +45 3862386 fax: +45 28623770.
E-mail address: Christian.jakobsen.03@regionh.dk (C. Jakobsen).

1873-9946/$ - see front matter
http://dx.doi.org/10.1016/j.crohns.2013.12.010

finding predictors for disease course in (paediatric) IBD, it will be worthwhile to include a combination of genetic, clinical and serological markers.

1. Introduction

Inflammatory bowel disease (IBD) includes Crohn's disease (CD), ulcerative colitis (UC) and indeterminate colitis/IBD unclassified (IBDU). The aetiology of IBD is unknown, but it is thought to originate from a complex interplay between environmental- and genetic factors.[1] Over the past decade our understanding of the genetics of IBD has improved using genome wide associations studies (GWA studies). Several GWA studies in adult IBD patients have been published[2–6] and subsequent meta-analyses have identified 163 genetic loci associated with IBD.[7–10] However, the loci discovered so far only cover around 25–30% of the genetic heritability to IBD.[8] To date two GWA studies have been conducted using paediatric IBD patients.[11,12] In the study by Kugathasan et al.,[12] two not previously published loci were identified: one on chromosome 20q13 (rs2315008 and rs4809330) and one on21q22 (rs2836878). Both loci were in non-coding regions; however the region on 20q13 was in linkage disequilibrium (LD) with a region harbouring the gene encoding *TNFRSF6B* (tumour necrosis factor receptor superfamily member 6b). Further analysis showed that *TNFRSF6B* mRNA expression was higher in colonic biopsies from IBD patients compared to healthy controls. No gene-containing regions were in LD with the locus on 21q22. The multicentre GWA study by Imielinski et al.[11] included patients from Italy, Scotland, the United States and Canada. A subset of these patients was also included in Ref. 12 In this study, five new regions were shown to be associated with early onset IBD: 16p11 (rs8049439) near the *IL27* gene, 22q12 (rs2412973), 10q22 (rs1250550), 2q37 (rs4676410) and 19q13 (rs10500264). Subsequent meta-analysis of the paediatric cohorts confirmed 21 of the previously reported 32 susceptibility loci associated with adult CD.[11] In UC, 8 of the 17 susceptibility loci were also found in paediatric UC.[11] These results indicate a shared genetic susceptibility among adult- and paediatric IBD patients.

To our knowledge, only one subsequent study[13] has replicated the findings by Imielinski et al.[11]The relationship between genotype and phenotype is of special clinical interest as it may improve our understanding of disease pathways (for instance the association between *NOD2* and ileal disease localisation in CD), and aid in predicting disease course and treatment response.[14] Several studies on genotype–phenotype associations in adult and paediatric IBD patients have been published but results are inconsistent. (*ATG16L1*,[15,16] *CARD9*,[17] *CDKAL1*,[18] *DLG5_e26*,[19,20] *ECM1*, *IBD5*,[21–23] *ICAM-1*,[24,25] *ICOSLG*,[8,26] *IL2/IL21*,[27] *IL27*,[11,13] *IRGM*,[15,28,29] *MAGI-2*,[30] *MDR1*,[31,32] *Myosin IXB*,[33] *NELL-1*,[34] *NOD 1*,[35] *NOD 2*,[28,36,37] *PTPN22*,[38] *SLC7A10*,[11] *TLR 4*,[39] *TNFSF15*,[40,41] *XBP-1*,[42] *ZMIZ-1*,[11,13] *ZNF365*[8]).

The aim of our study was to investigate if known genetic variants are associated with IBD in the Danish paediatric population; and if these variants are associated with specific sub-phenotypes.

2. Material and methods

2.1. Patients

In our case–control study we included IBD patients below 18 years of age at diagnosis. Patients treated in a hospital setting in Denmark are prospectively registered using the diagnostic codes from the WHO International Classification of Diseases, version 10 (ICD-10). They are registered in the electronic register of the local hospitals and in the National Patient Registry (NPR) using the Civil Registration Number (CPR-number). The CPR-number is a unique number given to each person at birth. We extracted data from the NPR using CPR-numbers and ICD-10 codes DK500–519 of all patients below 18 years of age in Eastern Denmark. Since 1982 all parents and their newborn children have been offered the opportunity to participate in the Danish national neonatal screening programme for congenital diseases and more than 90% of families have participated. All blood samples are kept on Guthrie cards and frozen to −20 °C in a biobank at Statens Serum Institut in Denmark.

All files from patients registered in NPR and who had DNA in the neonatal screening biobank at Statens Serum Institut or had agreed to have blood samples taken were retrieved and reviewed by one investigator (CJ). Clinical data at the last date of follow-up were recorded and included gender, date of diagnosis, height and weight at diagnosis, disease localization at diagnosis, disease behaviour at last follow-up (CD only) according to the Montreal classification,[43] perianal disease (yes/no), medical treatment (systemic steroids, azathioprine (AZA)/6-mercaptopurine (6MP), methotrexate (MTX), infliximab (IFX), adalimumab (ADA)) and date of surgery. Only bowel resections were included in the analysis.

Disease localization in CD and UC was classified according to the Paris classification[44] and for analysis subdivided into extensive and non-extensive diseases. In UC, extensive and non-extensive diseases included E3/E4 and E1/E2, respectively. In CD, extensive disease included L3 and all patients with upper GI disease (L4) and non-extensive disease included L2 and L1.

2.2. Controls

Controls were selected as part of another study.[45] In short, controls <18 years of age were randomly selected from the Danish Central Person Registry from Eastern Denmark, Funen and Aarhus. Only controls without an ICD-10 diagnosis of DK500–519 in the National Patient Registry were selected. As part of another questionnaire study, controls filled out data regarding their health status and controls with gastrointestinal- or autoimmune diseases were excluded.

2.3. DNA extraction

The DNA in this study was extracted from EDTA-stabilized blood samples, from Guthrie-cards or from cotton-swabs