Gastroenterology 2016;150:1196–1207

Variants in *TRIM22* That Affect NOD2 Signaling Are Associated With Very-Early-Onset Inflammatory Bowel Disease

Qi Li,[1,2,*] **Cheng Hiang Lee**,[3,4,*] **Lauren A. Peters**,[5,6,*] Lucas A. Mastropaolo,[2] Cornelia Thoeni,[1,2] Abdul Elkadri,[1,2,7] Tobias Schwerd,[8] Jun Zhu,[6] Bin Zhang,[6] Yongzhong Zhao,[6] Ke Hao,[6] Antonio Dinarzo,[6] Gabriel Hoffman,[6] Brian A. Kidd,[6] Ryan Murchie,[1,2] Ziad Al Adham,[1,2,7] Conghui Guo,[2] Daniel Kotlarz,[9] Ernest Cutz,[10] Thomas D. Walters,[1,2] Dror S. Shouval,[11] Mark Curran,[12] Radu Dobrin,[12] Carrie Brodmerkel,[12] Scott B. Snapper,[11,13] Christoph Klein,[9] John H. Brumell,[1,7,14] Mingjing Hu,[3,4] Ralph Nanan,[3,4] Brigitte Snanter-Nanan,[3,4] Melanie Wong,[15] Francoise Le Deist,[16] Elie Haddad,[17] Chaim M. Roifman,[18] Colette Deslandres,[19] Anne M. Griffiths,[1,2] Kevin J. Gaskin,[3,4] Holm H. Uhlig,[8] Eric E. Schadt,[6,§] and Aleixo M. Muise[1,2,7,§]

[1]SickKids Inflammatory Bowel Disease Center and Cell Biology Program, Research Institute, Hospital for Sick Children, Toronto, Ontario, Canada; [2]Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of Toronto, Hospital for Sick Children, Toronto, Ontario, Canada; [3]Gastroenterology Department, The Children's Hospital at Westmead, Westmead, New South Wales, Australia; [4]The James Fairfax Institute of Paediatric Nutrition, University of Sydney, New South Wales, Australia; [5]Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, New York; [6]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai and the Icahn Institute for Genomics and Multiscale Biology, New York, New York; [7]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada; [8]Translational Gastroenterology Unit, Nuffield Department of Clinical Medicine, Experimental Medicine Division, University of Oxford and Department of Pediatrics, John Radcliffe Hospital, Oxford, UK; [9]Department of Pediatrics, Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University, Munich, Germany; [10]Division of Pathology, The Hospital for Sick Children, Toronto, Ontario, Canada; [11]Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts; [12]Janssen R&D, LLC, Spring House, Pennsylvania; [13]Division of Gastroenterology and Hepatology, Brigham & Women's Hospital, Department of Medicine, Boston, Massachusetts; [14]Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; [15]Immunology Department, The Children's Hospital at Westmead, Westmead, New South Wales, Australia; [16]Department of Microbiology and Immunology, CHU Sainte Justine and Department of Microbiology, Infectiology and Immunology, University of Montreal, Quebec, Canada; [17]Department of Pediatrics, CHU Sainte-Justine, Department of Microbiology, Infectiology and Immunology, University of Montreal, Quebec, Canada; [18]Division of Immunology, Department of Pediatrics, University of Toronto, The Hospital for Sick Children, Toronto, Canada; and [19]Division of Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, CHU Sainte-Justine, Montreal, Quebec, Canada

See Covering the Cover synopsis on page 1050.

BACKGROUND & AIMS: Severe forms of inflammatory bowel disease (IBD) that develop in very young children can be caused by variants in a single gene. We performed whole-exome sequence (WES) analysis to identify genetic factors that might cause granulomatous colitis and severe perianal disease, with recurrent bacterial and viral infections, in an infant of consanguineous parents. **METHODS:** We performed targeted WES analysis of DNA collected from the patient and her parents. We validated our findings by a similar analysis of DNA from 150 patients with very-early-onset IBD not associated with known genetic factors analyzed in Toronto, Oxford, and Munich. We compared gene expression signatures in inflamed vs noninflamed intestinal and rectal tissues collected from patients with treatment-resistant Crohn's disease who participated in a trial of ustekinumab. We performed functional studies of identified variants in primary cells from patients and cell culture. **RESULTS:** We identified a homozygous variant in the tripartite motif containing 22 gene (*TRIM22*) of the patient, as well as in 2 patients with a disease similar phenotype. Functional studies showed that the variant disrupted the ability of TRIM22 to regulate nucleotide binding oligomerization domain containing 2 (NOD2)–dependent activation of interferon-beta signaling and nuclear factor–κB. Computational studies demonstrated a correlation between the TRIM22–NOD2 network and signaling pathways and genetic factors associated very early onset and adult-onset IBD. TRIM22 is also associated with antiviral and mycobacterial

***Authors share co-first authorship; §Authors share co-senior authorship.**

Abbreviations used in this paper: **eQTL, expression quantitative trait loci; GWAS, genome-wide association study; MDP, muramyl dipeptide; NF-κB, nuclear factor–κB; NOD2, nucleotide binding oligomerization domain containing 2; RSV, to respiratory syncytial virus; TNF, tumor necrosis factor; TRIM22, tripartite motif containing 22 gene; VEOIBD, Very-early-onset inflammatory bowel diseases; WES, whole exome sequencing.**

Most current article

0016-5085

http://dx.doi.org/10.1053/j.gastro.2016.01.031

effectors and markers of inflammation, such as fecal calprotectin, C-reactive protein, and Crohn's disease activity index scores. **CONCLUSIONS:** In WES and targeted exome sequence analyses of an infant with severe IBD characterized by granulomatous colitis and severe perianal disease, we identified a homozygous variant of *TRIM22* that affects the ability of its product to regulate NOD2. Combined computational and functional studies showed that the *TRIM22-NOD2* network regulates antiviral and antibacterial signaling pathways that contribute to inflammation. Further study of this network could lead to new disease markers and therapeutic targets for patients with very early and adult-onset IBD.

Keywords: VEOIBD; NF-kB; Antiviral and Antibacterial Networks.

Very-early-onset inflammatory bowel diseases (VEOIBD) often present with severe multisystemic disease that is difficult to treat with conventional therapies. These young patients frequently have novel causal[1–3] and risk variants[4–8] and common networks are now being established (reviewed in Uhlig et al[9]). For example, mutations in *IL10RA/B* genes cause a Mendelian form of VEOIBD with severe colitis and perianal disease[10] and mutations in *TTC7A* cause a severe form of apoptotic enterocolitis.[1]

Tripartite motif-containing 22 (TRIM22; also known as STAF50*)* is a RING finger E3 ubiquitin ligase[11] that is expressed in the intestine[12] and in macrophages,[13] and has a role in lineage-specific differentiation of lymphocytes.[14] TRIM22 was originally identified as an interferon inducible protein that possesses antiviral activity[15–17] and activates nuclear factor−κB (NF-κB) signaling.[13] Here we identify *TRIM22* functional variants associated with a distinct VEOIBD phenotype characterized by granulomatous colitis and severe perianal disease and show the *TRIM22-NOD2* network as a key antiviral and mycobacterial regulator.

Materials and Methods

See Supplementary Material for full methods information.

Subjects

All experiments were carried out with the approval of the Research Ethics Board at the Hospital for Sick Children. Informed consent to participate in research was obtained. A copy of the consent is available on the NEOPICS (International Early Onset Pediatric IBD Cohort Study) website (http://www.neopics.org/NEOPICS_Documents.html).

Whole Exome Sequencing

For patient 1 and her parents (trio), whole exome sequencing (WES) was performed using the Agilent SureSelect Human All Exon 50-Mb kit with high-throughput sequencing conducted using the SOLiD 4 System at The Center for Applied Genomics through the Hospital for Sick Children (Toronto). Sanger sequencing was used to verify variant genotypes in family 1 and infantile patients from the collaborating institutions were screened for *TRIM22* variants.

Validation

In order to validate these findings, we examined WES results from 150 infantile international VEOIBD patients without a genetic diagnosis who were previously sequenced in Toronto (NEOPICS), Oxford, and Munich (Care for Rare) and targeted exome sequencing of the *TRIM22* gene in 10 *IL10RA/B,* and *IL10*-negative patients without previous WES.

Computational Analysis

Datasets. Biopsy data were collected at baseline from anti−tumor necrosis factor (TNF)-resistant Crohn's patients enrolled in the ustekinumab trial described previously.[18] The ustekinumab trial expression data were used for construction of the adult IBD network. RNA-seq from the RISK cohort, as described previously,[19] was used for generation of the pediatric IBD Bayesian network.

Differential expression. To ascertain tissue-specific signatures, we examined inflamed vs noninflamed tissue from various anatomic regions of the small intestine, colon, and rectum. To identify differential expression signatures, we used an unbiased univariate filter to select top-varying genes and then applied Significance Analysis of Microarrays[20] (SAM) to whole-genome expression data collected from biopsy tissue samples and whole blood of individuals in the ustekinumab trial. To control for multiple hypothesis testing, we used the Benjamini & Hochberg adjustment on the raw P values to control the family-wise error rate and set a false discovery rate threshold of 1% or 5%. All analyses were performed using the R statistical package, version 2.15.24.1

Bayesian network. We employed Monte Carlo Markov Chain[21] simulation to identify potentially thousands of different plausible networks that are then combined to obtain a consensus network. eSNP data were used as priors as follows: genes with cis-eSNP[22] are allowed to be parent nodes of genes without cis-eSNPs, but genes without cis-eSNPs are not allowed to be parents of genes with cis-eSNPs.

Key driver analysis. Key driver analysis takes as input a set of genes (G) and a directed gene network (N) (eg, Bayesian network).[19,23–26] The objective is to identify the key regulators for the gene sets with respect to the given network. Key driver analysis first generates a subnetwork N_G, defined as the set of nodes in N that are no more than h-layers away from the nodes in G, and then searches the h-layer neighborhood (h = 1,..,H) for each gene in N_G ($HLN_{g,h}$) for the optimal h*, such that

$$ES\,h* \;=\; \max\,(\,ES_{\,h,\,g}\,)\,\forall\; g \in N_{\,g},\; h \in \{1\,..\,H\,\}$$

where $ES_{h,g}$ is the computed enrichment statistic for $HLN_{g,h}$. A node becomes a candidate driver if its HLN is significantly enriched for the nodes in G. Candidate drivers without any parent node (ie, root nodes in directed networks) are designated as global drivers, while the remaining are designated as local drivers.

Pathway enrichment. Expression quantitative trait loci (eQTL) analysis was performed on IBD, ulcerative colitis, and Crohn's disease genome-wide association study (GWAS) hits from the National Human Genome Research Institute catalog, plus all Wellcome Trust Case Control Consortium IBD GWAS single nucleotide polymorphisms with $P \leq .001$. Ten percent false discovery rate eGenes were thus extracted, and tested for enrichment of pathways from the Metacore database.