Gastroenterology 2014;146:200–209

Genetic Variants Synthesize to Produce Paneth Cell Phenotypes That Define Subtypes of Crohn's Disease

Kelli L. VanDussen,[1,*] Ta-Chiang Liu,[1,*] Dalin Li,[2,3] Fadi Towfic,[4] Nir Modiano,[2] Rachel Winter,[2] Talin Haritunians,[2,3] Kent D. Taylor,[3] Deepti Dhall,[5] Stephan R. Targan,[2] Ramnik J. Xavier,[4] Dermot P. B. McGovern,[2,3] and Thaddeus S. Stappenbeck[1]

[1]*Department of Pathology and Immunology, Washington University School of Medicine, St Louis, Missouri;* [2]*The F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute,* [3]*Medical Genetics Research Institute, and* [5]*Department of Pathology, Cedars-Sinai Medical Center, Los Angeles, California; and* [4]*Center for Computational and Integrative Biology, Massachusetts General Hospital, Harvard Medical School and Broad Institute, Boston, Massachusetts*

BACKGROUND & AIMS: Genetic susceptibility loci for Crohn's disease (CD) are numerous, complex, and likely interact with undefined components of the environment. It has been a challenge to link the effects of particular loci to phenotypes of cells associated with pathogenesis of CD, such as Paneth cells. We investigated whether specific phenotypes of Paneth cells associated with particular genetic susceptibility loci can be used to define specific subtypes of CD. **METHODS:** We performed a retrospective analysis of 119 resection specimens collected from patients with CD at 2 separate medical centers. Paneth cell phenotypes were classified as normal or abnormal (with disordered, diminished, diffuse, or excluded granule phenotypes) based on lysozyme-positive secretory granule morphology. To uncover the molecular basis of the Paneth cell phenotypes, we developed methods to determine transcriptional profiles from whole-thickness and laser-capture microdissected, formalin-fixed, paraffin-embedded tissue sections. **RESULTS:** The proportion of abnormal Paneth cells was associated with the number of CD-associated *NOD2* risk alleles. The cumulative number of *NOD2* and *ATG16L1* risk alleles had an additive effect on the proportion of abnormal Paneth cells. Unsupervised clustering analysis of demographic and Paneth cell data divided patients into 2 principal subgroups, defined by high and low proportions of abnormal Paneth cells. The disordered and diffuse abnormal Paneth cell phenotypes were associated with an altered transcriptional signature of immune system activation. We observed an inverse correlation between abnormal Paneth cells and presence of granuloma. In addition, high proportions of abnormal Paneth cells were associated with shorter time to disease recurrence after surgery. **CONCLUSIONS:** Histologic analysis of Paneth cell phenotypes can be used to divide patients with CD into subgroups with distinct pathognomonic and clinical features.

Keywords: Pathogenesis; Prognostic Factor; Diagnosis; Inflammatory Bowel Disease.

The inflammatory bowel diseases (IBD), Crohn's disease (CD) and ulcerative colitis (UC), currently affect approximately 1.5 million people in the United States and are a significant cause of morbidity, particular among young people.[1,2] A major hurdle in understanding the pathogenesis of complex chronic diseases such as IBD is incorporating the effects of both numerous genetic susceptibility loci and poorly defined environmental factors. This challenge often precludes precise genotype−phenotype correlations and further identification of disease mechanism-based surrogate markers.[3] One potential mitigating factor is the accessibility of clinical tissue samples from both involved and noninvolved areas for analyses, as this might allow the incorporation of histopathological and other tissue characteristics into genetic analyses to better understand the etiopathogenesis of disease. Among the major complex genetic diseases, IBD is unique in that tissue samples are routinely obtained as part of clinical practice. Therefore, IBD serves as an ideal platform to test the hypothesis that histological changes are a more homogenous phenotype than standard clinical manifestations for testing genotype−phenotype correlations.

One CD-relevant cell type is the Paneth cell, which is a specialized secretory cell type located at the bases of the crypts of Lieberkühn in the small intestine.[4] These cells produce a wide repertoire of antimicrobial peptides, such as lysozyme and α-defensins, to modulate the intestinal microbiome,[5,6] and are important mediators of the host innate immune response.[4,7] We previously demonstrated that the packaging of antimicrobial peptides into granules and their secretion were impaired in the Paneth cells of mice with hypomorphic expression of the CD susceptibility gene *Atg16l1*.[8,9] Importantly, we observed similar Paneth cell abnormalities in a small number of CD patients homozygous for the *ATG16L1 T300A* CD risk allele,[8] demonstrating that this is a valid approach to link genetics and phenotypes of a disease-relevant cell type. In addition to *ATG16L1*, *NOD2* has been identified as a CD susceptibility locus and has been predicted to disrupt Paneth cell function.[10–14]

CD is remarkable for both its heterogeneous clinical course and its varied histopathological findings.[15–17] The

*Authors share co-first authorship.

Abbreviations used in this paper: CD, Crohn's disease; IBD, inflammatory bowel disease; UC, ulcerative colitis.

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2013.09.048

clinical variability in natural history and response to therapy is likely, in part, a consequence of the genetic heterogeneity that underlies these conditions. Major challenges to genotype−phenotype association studies are the lack of robust and reproducible criteria to define end points as well as sufficient numbers of genotyped patients. Recent genome-wide association studies have extended the number of known IBD susceptibility loci to >160. These studies and others have implicated multiple pathways in IBD pathogenesis, including epithelial barrier homeostasis, innate immune response, antigen presentation, autophagy, Paneth cell defects, and interleukin-23/T_H17 signaling.[10,11,18,19] Here, we hypothesized that linking genetics to disease-associated phenotypes in relevant cell types based on predicted disease mechanisms (ie, Paneth cells) can be a successful method for defining more homogenous subtypes of CD. We also propose that analyzing regions of intestine free of severe active or chronic inflammation (ie, lack of substantial pathologic hallmarks) will provide an objective end point that will more accurately reflect disease pathogenesis, as these areas can harbor early molecular and pathologic changes. A greater understanding of the causes of the observed clinical heterogeneity will lead to improved clinical management through a more individualized approach to disease management and, potentially, development of new therapies.

Materials and Methods

Description and Genotyping of Patient Cohort

Full methods are provided in the Supplementary Detailed Methods. Patients were recruited at Barnes-Jewish Hospital, St Louis between 2005 and 2013 or at Cedars-Sinai Medical Center, Los Angeles between 1999 and 2013. Patient DNA samples were genotyped for *ATG16L1 T300A* and the CD-associated *NOD2* variants.[10,20,21] Patients from the Barnes-Jewish Hospital cohort were genotyped by the Digestive Disease Research Core Center using matrix-assisted laser desorption ionization-time of flight mass spectrometry and by the Genome Technology Access Center using the Human OmniQuad SNP genotyping arrays (Illumina, San Diego, CA). Patients from the Cedars-Sinai cohort were genotyped using the Immunochip (Illumina). The study protocol was approved by the Institutional Review Boards of Washington University-St Louis and Cedars-Sinai Medical Center. Written informed consent was obtained from all study participants.

Morphological Analysis of Paneth Cells

For each resection case, an H&E-stained tissue section of the proximal margin (terminal ileum) was identified by pathologists (T.S.S. and T.C.L.). Cases were included for Paneth cell analysis if the section contained at least 100 well-oriented intestinal crypts and exhibited absent or minimal acute and/or chronic inflammation (Supplementary Figure 1). Lysozyme distribution was quantified as described previously.[8] For each case, a pathologist (T.C.L.) who was blinded to the characteristics of the cases scored a minimum of 200 Paneth cells (range, 206–2702) in well-oriented crypts. Paneth cells located within Peyer's patches were excluded.

Transcriptional Analysis

RNA was procured from the set of archived formalin-fixed paraffin-embedded surgical resection samples used for histological analysis. Microarrays were performed as described previously.[8] Data are deposited at ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) with accession number E-MTAB-1281.

Statistical Analyses

For analysis of lysozyme quantification, permutation tests were performed to determine the association between *NOD2* variants and the percentage of abnormal Paneth cells using R statistical software (version 2.13.1, R Foundation for Statistical Computing). Mann-Whitney tests were used to demonstrate statistical difference between cases with 1 or 2 *NOD2* risk variants and controls. Linear regression was used to analyze the cumulative number of risk variants. For correlation analyses, Pearson correlations were calculated using GENE-E,[22] which were then used as the distance measure for unsupervised, hierarchical clustering of the patients. A marker selection strategy based on signal-to-noise ratios[23] was used to identify clinical variables associated with patient subtypes. A χ^2 test and a log-rank test were performed for the analysis of granuloma incidence and time to disease recurrence, respectively (Prism GraphPad software). $P < .05$ was considered significant.

Results

Association of NOD2 CD Susceptibility Variants With Abnormal Paneth Cell Phenotype

We performed a retrospective analysis of Paneth cell phenotypes in genotyped CD patients (N = 119) using resection specimens. In order to study tissue that might exhibit early pathologic and molecular changes associated with disease pathogenesis, we examined ileal tissue samples that demonstrated no evidence of active/chronic disease. Paneth cell analysis was performed using our previously developed system for robust, quantitative scoring of Paneth cell phenotypes based on high-resolution localization of lysozyme protein,[8] a highly expressed antimicrobial protein that is normally packaged into Paneth cell secretory granules.[7] A pathologist (T.C.L.) who was blinded to the individuals' characteristics scored a minimum of 200 Paneth cells per case as normal or as abnormal (including Paneth cells scored as disordered, diminished, diffuse, or "excluded granule") (Figure 1*A* and Supplementary Figure 2). The excluded granule abnormal Paneth cell phenotype was identified in this study and is characterized by granule shapes that have low/absent lysozyme staining, but also contain diffuse cytoplasmic lysozyme staining and occasionally a few lysozyme-positive granules (Figure 1*B*). For most cases, Paneth cells with normal morphology were predominant. When present, Paneth cells with an abnormal phenotype were found interspersed among those with normal morphology.

In light of the established effect of the *ATG16L1 T300A* CD susceptibility variant on Paneth cell phenotype,[8] we initially excluded patients with this variant. The *ATG16L1 T300A* CD risk allele is common, with a risk allele carriage