Published in final edited form as:

Gastroenterology. 2014 January ; 146(1): 200–209. doi:10.1053/j.gastro.2013.09.048.

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^{2,4}, Fadi Towfic³, Nir Modiano², Rachel Winter², Talin Haritunians^{2,4}, Kent D. Taylor⁴, Deepti Dhall⁵, Stephan R. Targan², Ramnik J. Xavier³, Dermot P. B. McGovern^{2,4,*}, and Thaddeus S. Stappenbeck^{1,*}

¹Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

²The F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

³Center for Computational and Integrative Biology, Massachusetts General Hospital, Harvard Medical School and Broad Institute, Boston, MA 02115, USA

⁴Medical Genetics Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

⁵Department of Pathology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

Abstract

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

© 2013 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

*Corresponding Authors: Thaddeus S. Stappenbeck, 660 S. Euclid Ave, Box 8118, St. Louis, MO 63110, Phone: 314-362-4214, stappenn@pathology.wustl.edu. Dermot P. B. McGovern, 8797 Beverly Blvd, Suite 300, Los Angeles, CA, 90048, Dermot.McGovern@cshs.org.

[†]These authors contributed equally to this work.

Author Disclosures: None

Transcript Profiling: Microarray data are deposited at ArrayExpress <http://www.ebi.ac.uk/arrayexpress/> (accession number E-MTAB-1281).

Author Contributions:

K.L.V., T.C.L., D.P.B.M. and T.S.S. contributed to study concept and design. K.L.V. and T.C.L. contributed to data acquisition. R.W., D.P.B.M. and S.R.T. contributed to collection of the patient cohort. T.C.L., N.M., D.D. and T.S.S. performed pathological analysis. K.L.V., T.C.L., F.T., T.H., K.D.T. and R.J.X. contributed to data analysis and interpretation. K.L.V., T.C.L., D.P.B.M. and T.S.S. drafted the manuscript. K.L.V., T.C.L., D.L. and F.T. performed statistical analysis. D.P.B.M. and T.S.S. obtained funding and supervised the study.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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 the presence of granuloma. Moreover, 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

INTRODUCTION

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 markers³. One potential mitigating factor is the accessibility of clinical tissue samples from both involved and non-involved areas for analyses as this may allow the incorporation of histopathological and other tissue characteristics into genetic analyses in order 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 specialized secretory cell type located at the bases of the crypts of Lieberkühn in the small intestine⁴. These cells produce a wide repertoire of antimicrobial peptides, such as lysozyme and α -defensins, to modulate the intestinal microbiome^{5, 6}, and thus are important mediators of the host innate immune response^{4, 7}. We previously demonstrated that the packaging of anti-microbial peptides into granules and their secretion was 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 allele⁸, 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 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 endpoints as well as sufficient numbers of genotyped patients. Recent genome-wide association studies (GWAS) have extended the number of known IBD susceptibility loci to more than 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 IL-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 (i.e., Paneth cells) may 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 (i.e., lack of substantial pathologic hallmarks) will provide an objective endpoint that will more accurately reflect disease pathogenesis, as these areas may 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, the development of new therapies.

MATERIALS AND METHODS

Description and genotyping of patient cohort

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 (DDRCC) using matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry and by the Genome Technology Access Center (GTAC) 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, a hematoxylin and eosin-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 (Figure S1). Lysozyme distribution was quantified as previously described⁸. 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: 206–2,702) 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 previously described⁸. 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-E²², which were then used as the distance measure for unsupervised, hierarchical clustering of the patients. A marker selection strategy based on signal-to-noise ratios²³ was used to identify clinical variables associated with patient subtypes. A Chi-Square test and a log-rank test were performed for the analysis of granuloma incidence and time to disease recurrence, respectively (Prism GraphPad software). $P < 0.05$ was considered to be 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 may 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 protein⁸, a highly-expressed antimicrobial protein that is normally packaged into Paneth cell secretory granules⁷. 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 1A; Figure S2). 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 1B). 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 phenotype⁸, we initially excluded patients with this variant. The *ATG16L1 T300A* CD risk allele is common, with a risk allele carriage rate greater than 80% in people of European descent²⁰ (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2241880). Therefore, we included cases from both the Washington University School of Medicine and the Cedars-Sinai Medical Center patient cohorts. Confounding effects due to specimen source were not apparent. In this combined cohort *ATG16L1* 'safe' cases (cases 1–59, Table S1), we first demonstrated that the group of cases without *NOD2* risk variants ($n = 29$) and the group of cases with one or more of the common and/or rare²¹ CD-associated *NOD2* variants ($n = 30$) had similar demographic compositions (Table 1). In control cases without *NOD2* CD risk alleles, an average of 87% of the total number of Paneth cells counted per case had a normal phenotype based on lysozyme localization. In cases with one or more *NOD2* CD susceptibility alleles, the average percent of Paneth cells scored as normal was significantly lower than controls (76% with one variant allele and 69% with two variant alleles) (Figure 1C). Accordingly, cases with one or more *NOD2* CD susceptibility variants had significantly increased proportions of disordered, diminished and diffuse abnormal Paneth cell phenotypes compared to the control cases (Figure 1C).

The *NOD2* locus contains three common CD susceptibility alleles, namely *R702W*, *G908R* and *L1007fsXinsC*. Several studies have suggested that each of these variants attenuate *NOD2* function based on an impaired cellular response following exposure to the *NOD2* ligand, bacterial peptidoglycan^{24–27}. We investigated whether these variants were similarly associated with abnormal Paneth cell phenotype in the *ATG16L1* 'safe' cohort (Figure 1D). We found that the *R702W* and *L1007fsXinsC* variants were associated with increased proportions of abnormal Paneth cells. Of note, two groupings of cases were observed for

each of these alleles, one with < 20% abnormal Paneth cells and the other with 20% abnormal Paneth cells. This observation suggests that the *NOD2* variant alone is not sufficient to drive the abnormal Paneth cell phenotypes, but rather additional genetic variants or host/environmental factors also contribute to this phenotype. We did not observe an association between *NOD2* G908R and abnormal Paneth cell phenotypes in this study. However, we were able to obtain only a small number of patients with this allele and thus cannot exclude the possibility of such an association. Similarly, potential associations of abnormal Paneth cell phenotypes and compound *NOD2* variants (e.g., R702W + L1007fsXinsC), rare CD-associated *NOD2* variants or homozygous variants were not interpretable due to the rarity of cases with these genotypes (Supplemental Figure S3).

Paneth cell phenotypes define CD patient subtypes

As our previous studies showed that the CD risk variant *ATG16L1* T300A was associated with increased proportions of abnormal Paneth cell phenotypes, we next wanted to test for a potential interaction between this variant and the CD-associated *NOD2* variants. For this analysis, we used a larger CD cohort that included both *ATG16L1* T300A 'safe' and 'risk' cases (cases 1–119, Tables S1 and S2). Similar to our previous study⁸, we observed an association between *ATG16L1* T300A and abnormal Paneth cells in this expanded cohort (Figure S4). We observed a significant, positive correlation between the cumulative number of *ATG16L1* T300A or CD-associated *NOD2* risk alleles and the proportion of abnormal Paneth cells, demonstrating an additive effect of these alleles (Figure 2A).

As a second experiment to test the association between the CD-associated *NOD2* and *ATG16L1* T300A variants and Paneth cell phenotype, we performed a patient-patient comparison using the same expanded cohort. Unsupervised analysis of Paneth cell phenotypes, genetics, documented environmental exposures and demographic information was used to cluster the CD cases according to similarity (Figure 2B). Interestingly, the resulting heat map showed two principal subtypes of cases. A minor number of cases did not fit either subtype. Using a marker selection strategy²³, we identified the factors that defined the two CD patient subtypes: the abnormal Paneth cell phenotype categories of diminished, diffuse and excluded granule Paneth cells and the total percent of abnormal Paneth cells (Figure 2C). The cumulative number of *NOD2* risk variants and *NOD2* R702W genotype were also defining factors for the two CD patient subtypes, supporting our initial finding that CD-associated *NOD2* variants are associated with abnormal Paneth cell phenotypes. We also performed the inverse experiment and looked for potential correlations between Paneth cell phenotypes and the other experimental factors (genetics, documented environmental exposures and demographic information). Significant correlations were again observed between the abnormal Paneth cell phenotype categories and *NOD2* genotype, but we did not observe significant correlations between the abnormal Paneth cell phenotypes and the environmental/demographic parameters (Figure S5). Taken together, these data demonstrate that Paneth cell phenotype is strongly linked to particular genetic alleles and can more clearly define subtypes of CD patients than other demographic parameters (treatment, disease behavior, disease location, etc.).

An activated immune response gene signature is associated with diffuse Paneth cell phenotype

To investigate the molecular basis of the Paneth cell phenotypes, we developed a method whereby we could obtain RNA from the same archival formalin-fixed, paraffin-embedded tissue samples used to morphologically phenotype Paneth cells. Transcriptional profiles of the CD patient material were generated by performing microarray analysis with one set of RNAs procured by laser capture microdissection of ileal crypt bases (enriched for Paneth cells, n = 15 samples) and a second set of RNAs procured from whole, unstained ileal tissue

sections scraped from glass microscope slides (n = 40). To analyze these data (Figure S6), we first compared the two sets of transcriptional profiles to identify transcripts enriched in the Paneth cells of CD patients. Next, we performed a correlation analysis between the expression of the Paneth cell-enriched genes and the quantitative Paneth cell phenotype data for the same 40 patients using the whole ileal tissue data set to identify distinct sets of Paneth cell-enriched transcripts with expression values that were significantly correlated with each Paneth cell phenotype (Table S4).

We performed gene ontology (GO) term analysis for each transcript set that was correlated with a particular Paneth cell phenotype (Figure 3 and Figures S7, S8 and S9). This analysis showed that the transcript sets that correlated with the disordered and diffuse Paneth cell phenotypes were both enriched for immune system-related biological processes. We previously observed a gene signature of cytokine stimulation in the Paneth cells of mice with hypomorphic expression of *Atg16l1* (which have increased proportions of Paneth cells with the disordered and diffuse phenotypes)⁸. These data suggest that altered immune activation may be occurring in human and mouse intestinal tissue that exhibits abnormal Paneth cell phenotypes, although the underlying causes of these transcriptional profiles remains unclear. Thus, the disordered and diffuse Paneth cell phenotypes appear to be histological readouts of specific types of inflammatory responses occurring in a subset of CD patients.

Abnormal Paneth cell phenotype inversely correlates with granuloma incidence

We next examined whether other, more classic morphological features of CD, such as the presence of granuloma, were correlated with the Paneth cell phenotypes. Granulomas are a distinguishing feature of CD that is not present in all patients²⁸. For this analysis, pathologists (T.C.L. and D.D.) reviewed the entire case for each patient, including both involved and non-involved intestinal regions (n = 107, cohort from patients 1–119 with available material, Tables S1 and S2). We found that cases with low proportions of abnormal Paneth cell phenotypes (arbitrarily defined as < 20% based on this study and our previous *in vivo* studies⁸) had a higher incidence of granuloma than cases with high proportions of abnormal Paneth cell phenotypes (43.8% vs. 19.4%, respectively; $P = 0.0160$) (Figure 4A).

We then examined whether the presence of granuloma was associated with a particular abnormal Paneth cell phenotype. We found that the excluded granule phenotype exhibited the strongest inverse correlation with granuloma incidence, with 39.4% of cases with low proportions of excluded granule Paneth cells (< 1%) having granuloma compared to 17.9% of cases with high proportions of excluded granules (> 1%) (Figure 4B; $P = 0.0344$). A similar finding was observed with the diminished granule Paneth cell phenotype (46.5% vs. 26.8% incidence in cases with low [$<10\%$] and high [$>10\%$] proportions of diminished granule Paneth cells, respectively; $P = 0.0347$) (Figure 4C). Finally, cases with low proportions of diffuse granule Paneth cells (<5%) had 37.0% granuloma incidence compared to 14.3% granuloma incidence observed in cases with high proportions of diffuse granules (> 5%), although this correlation was not significant ($P = 0.4159$; Figure 4D). In accordance with previous reports, we did not find a significant association between *NOD2* status (including individual risk alleles) and granuloma incidence in our cohort (data not shown)^{28–30}. In summary, we identified that abnormal Paneth cell phenotypes are inversely associated with the presence of granuloma.

Paneth cell phenotype is associated with disease prognosis

We next investigated if Paneth cell phenotypes were associated with the disease prognosis post-resection. For this analysis, we used our expanded CD cohort (cases 1–119, Tables S1

and S2) as well as an additional 59 CD cases that have not yet been genotyped but had Paneth cell scoring and disease recurrence data (cases 120–178, Table S3). In patients who received prophylactic therapy post-resection (i.e., suggestive of patients with a more aggressive clinical course prior to resection; $n = 102$), those with high proportions of abnormal Paneth cells ($\geq 20\%$) had a significantly shorter time to disease recurrence compared to those with low proportions of abnormal Paneth cells ($< 20\%$) ($P = 0.0200$; Figure 5). A significant difference was not observed in patients who did not receive prophylaxis post-resection (i.e., patients with a less aggressive clinical course prior to resection) (Figure S10). Thus, we conclude that Paneth cell phenotypes can identify clinically relevant subtypes of CD patients.

DISCUSSION

Here, we have provided the first evidence that a defined cellular phenotype (in Paneth cells) is linked to multiple CD genetic susceptibility loci and subdivides patients into two groups. In addition, we defined the molecular consequences of this phenotype in human Paneth cells and found an association with immune activation. We demonstrated that Paneth cell phenotypes are associated with the presence of granuloma, a classic histological finding associated with CD. We also showed that the Paneth cell phenotypes are associated with a specific clinical outcome in the current cohort, i.e., time to disease recurrence. While there clearly is a genetic basis for this phenotype, because of the potential complexity of the genetics and environmental interactions, we propose that the Paneth cell phenotypes would be a readily accessible, integrative readout of this information (rather than genetics alone) and should be further evaluated for their ability to stratify CD patients in a clinically meaningful way.

Here, we demonstrated that CD cases with one or more *NOD2* susceptibility alleles had increased proportions of abnormal Paneth cells compared to those without *NOD2* susceptibility alleles. In this study, we also found an additive effect of *NOD2* susceptibility alleles and *ATG16L1 T300A* on Paneth cell abnormalities. This finding is of interest because Paneth cell phenotypes have now been demonstrated to link two IBD susceptibility genes that have been suspected to act in a shared biochemical pathway^{31–33}. It has been proposed that IBD genes act in pathways¹¹, and our findings provide additional evidence that supports this theory.

There is a pressing need to classify subtypes of CD based on underlying pathological mechanisms including genetics and not solely by clinical parameters. This need is highlighted by the fact that treatment with anti-TNF- α monoclonal antibodies, the most advanced biologic currently employed, does not induce remission in the majority of CD cases, and the majority of initial responders do not maintain long-term remission^{34, 35}. Furthermore, many potential therapeutics for CD have failed to induce or maintain remission of active disease³⁵. Although some of these “failed” therapeutics truly do lack efficacy, inappropriate endpoint selection and patient heterogeneity have been cited as factors contributing to treatment inefficacy in clinical trials³⁴. Thus, it is becoming increasingly apparent that novel strategies to define and stratify CD patients that are based on genotype and other molecular parameters will be needed to progress towards improved diagnostics, prognostics and therapeutics. Because the ability to test potential CD therapeutics in a particular subtype of patients may yield better outcomes, the abnormal Paneth cell phenotypes, as an objective biomarker that is both related to genetic susceptibility loci and disease pathogenesis mechanisms, should be explored as a method to identify a subtype of CD patients that has a similar disease pathogenesis mechanism.

In this study, we observed a novel and important histological correlate to the Paneth cell phenotype, an inverse relationship between the presence of granuloma and the proportion of abnormal Paneth cells. This is, to our knowledge, the first report linking defects in a CD-relevant cell type and other, more classical histologic changes of CD. The standard practice for pathologic examination for CD includes extensive sampling of the specimens (at a minimum of 1 section/10 cm), and as there might be unsampled granulomas in patients where no granulomas were found, it is perhaps best to classify the two subsets of CD patients as “granuloma-rich” and “granuloma-poor”. We showed that the total proportion of abnormal Paneth cells, and in particular, excluded and diminished granules, were strong indicators of the granuloma-poor subtype. This is important as while we showed that *NOD2* status was associated with Paneth cell phenotype, we and others also showed that there was no significant correlation between *NOD2* status and granuloma^{28–30}, and that Paneth cell phenotype is the best predictor of granuloma. Our work supports the model that cell-specific readouts that can integrate the effects of both genetic and environmental factors are more informative and clinically relevant (Figure 6). There has been controversy over the biologic and prognostic implications of granulomas in CD^{36, 37}, with some recent studies indicating that granulomas might be linked to more aggressive clinical behavior²⁸ and others that it is not^{37, 38}. An immediate clinical implication of our finding is that stratification of CD based on Paneth cell phenotypes may be a more reliable approach for clinical management and clinical trial design, as features such as granuloma could be sparse and more likely to be undersampled, especially in biopsy specimens, whereas Paneth cell phenotypes are more easily analyzed within limited samples. Prospective studies using Paneth cell phenotypes as stratification for novel CD therapeutics may yield new insights.

In contrast to the success that has been achieved in treating single gene diseases (e.g., chronic myelogenous leukemia and imatinib (Gleevec)³⁹), the development of effective therapeutic options for complex genetic diseases, such as rheumatoid arthritis, systemic lupus erythematosus and IBD has been difficult^{34, 40, 41}. Of the complex genetic diseases, CD studies are uniquely positioned to lead the way in the development of novel therapies. First, there is a clearer understanding of the genetic susceptibility loci that are associated with CD susceptibility compared to rheumatoid arthritis or systemic lupus erythematosus. Second, relevant tissue and cell types (not restricted to Paneth cells) are relatively easy to access. In contrast, for diseases such as type I diabetes, whereas there is a good understanding of the associated genetics, there is poor tissue availability³. Thus, we believe that CD can be the prototype disease for moving forward the study of complex genetic diseases for both the development of new therapeutics and a more personalized approach to disease management.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant Support: K.L.V. was supported by an NIH training grant (T32 AI007163). The research was funded by a Washington University Institute of Clinical and Translational Sciences Pilot Award (CTSA308). The Washington University Digestive Disease Research Core Center is supported by a grant from the National Institute of Diabetes and Digestive and Kidney Disease (NIDDK) (P30DK052574). IBD Research at Cedars-Sinai is supported by USPHS grant PO1DK046763 and the Cedars-Sinai F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute Research Funds. Genotyping at CSMC is supported in part by the National Center for Research Resources (NCRR) grant M01-RR00425, UCLA/Cedars-Sinai/Harbor/Drew Clinical and Translational Science Institute (CTSI) Grant (UL1 TR000124-01), the Southern California Diabetes and Endocrinology Research Grant (DERC) (DK063491). Project investigators are supported by The Helmsley Charitable Trust (D.P.B.M.), The European Union (D.P.B.M.), The Crohn's and Colitis Foundation of America (CCFA) (D.P.B.M.), The Feintech Family Chair in IBD (S.R.T.), The Joshua L. and Lisa Z. Greer Chair in IBD

Genetics (D.P.B.M.), and grants DK043351, DK097485, DK092405 (R.J.X.), DK062413, DK046763-19, AI067068, HS021747 (D.P.B.M.) and AI08488702 (T.S.S.). T.S.S. and R.J.X. are supported by the CCFA Genetics initiative.

K.L.V. was supported by an NIH training grant (T32 AI007163). The research was funded by a Washington University Institute of Clinical and Translational Sciences Pilot Award (CTSA308). The Washington University Digestive Disease Research Core Center is supported by a grant from the National Institute of Diabetes and Digestive and Kidney Disease (NIDDK) (P30DK052574). R.J.X. is supported by grants DK043351, DK097485, DK092405. T.S.S. is supported by grant AI08488702. Cedars-Sinai is supported by an USPHS grant (PO1DK046763), The F.Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute Research Funds, a National Center for Research Resources (NCRR) grant (M01-RR00425), UCLA/Cedars-Sinai/Harbor/Drew Clinical and Translational Science Institute (CTSI) Grant (UL1 TR000124-01), Southern California Diabetes and Endocrinology Research Grant (DERC) (DK063491) and grants DK062413, DK046763-19, AI067068, HS021747 (D.P.B.M.). Research support was also received from The Helmsley Foundation (D.P.B.M. and T.S.S.) and the Crohn's and Colitis Foundation of America (R.J.X., D.P.B.M. and T.S.S.).

Abbreviations

GO	gene ontology
GWAS	genome-wide association studies
TNF-α	tumor necrosis factor α

References

Author names in bold designate shared co-first authorship.

- Burisch J, Munkholm P. Inflammatory bowel disease epidemiology. *Curr Opin Gastroenterol*. 2013; 29:357–62. [PubMed: 23695429]
- Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology*. 2012; 142:46–54. e42. quiz e30. [PubMed: 22001864]
- Virgin HW, Todd JA. Metagenomics and personalized medicine. *Cell*. 2011; 147:44–56. [PubMed: 21962506]
- Stappenbeck TS. Paneth cell development, differentiation, and function: new molecular cues. *Gastroenterology*. 2009; 137:30–3. [PubMed: 19497398]
- Salzman NH, Ghosh D, Huttner KM, et al. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature*. 2003; 422:522–6. [PubMed: 12660734]
- Salzman NH, Hung K, Haribhai D, et al. Enteric defensins are essential regulators of intestinal microbial ecology. *Nat Immunol*. 2010; 11:76–83. [PubMed: 19855381]
- Ouellette AJ. Paneth cells and innate mucosal immunity. *Curr Opin Gastroenterol*. 2010; 26:547–53. [PubMed: 20693892]
- Cadwell K, Liu JY, Brown SL, et al. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. *Nature*. 2008; 456:259–63. [PubMed: 18849966]
- Cadwell K, Patel KK, Maloney NS, et al. Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell*. 2010; 141:1135–45. [PubMed: 20602997]
- 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. [PubMed: 23128233]
- Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature*. 2011; 474:307–17. [PubMed: 21677747]
- Hugot JP, Chamaillard M, Zouali H, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*. 2001; 411:599–603. [PubMed: 11385576]
- Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*. 2001; 411:603–6. [PubMed: 11385577]

14. Wehkamp J, Harder J, Weichenthal M, et al. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. *Gut*. 2004; 53:1658–64. [PubMed: 15479689]
15. Abraham C, Cho JH. Inflammatory bowel disease. *N Engl J Med*. 2009; 361:2066–78. [PubMed: 19923578]
16. Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet*. 2012; 380:1590–605. [PubMed: 22914295]
17. Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet*. 2007; 369:1641–57. [PubMed: 17499606]
18. Wehkamp J, Fellermann K, Herrlinger KR, et al. Mechanisms of disease: defensins in gastrointestinal diseases. *Nat Clin Pract Gastroenterol Hepatol*. 2005; 2:406–15. [PubMed: 16265431]
19. Wehkamp J, Stange EF. Paneth's disease. *J Crohns Colitis*. 2010; 4:523–31. [PubMed: 21122555]
20. Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet*. 2010; 42:1118–25. [PubMed: 21102463]
21. Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet*. 2011; 43:1066–73. [PubMed: 21983784]
22. Gehlenborg N, Wong B. Heat maps. *Nature Methods*. 2012; 9:213–213.
23. Ramaswamy S, Tamayo P, Rifkin R, et al. Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci U S A*. 2001; 98:15149–54. [PubMed: 11742071]
24. Bonen DK, Ogura Y, Nicolae DL, et al. Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. *Gastroenterology*. 2003; 124:140–6. [PubMed: 12512038]
25. Chamaillard M, Philpott D, Girardin SE, et al. Gene-environment interaction modulated by allelic heterogeneity in inflammatory diseases. *Proc Natl Acad Sci U S A*. 2003; 100:3455–60. [PubMed: 12626759]
26. Inohara N, Ogura Y, Fontalba A, et al. Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J Biol Chem*. 2003; 278:5509–12. [PubMed: 12514169]
27. Li J, Moran T, Swanson E, et al. Regulation of IL-8 and IL-1beta expression in Crohn's disease associated NOD2/CARD15 mutations. *Hum Mol Genet*. 2004; 13:1715–25. [PubMed: 15198989]
28. Heresbach D, Alexandre JL, Branger B, et al. Frequency and significance of granulomas in a cohort of incident cases of Crohn's disease. *Gut*. 2005; 54:215–22. [PubMed: 15647184]
29. Pierik M, De Hertogh G, Vermeire S, et al. Epithelioid granulomas, pattern recognition receptors, and phenotypes of Crohn's disease. *Gut*. 2005; 54:223–7. [PubMed: 15647186]
30. Shaoul R, Karban A, Weiss B, et al. NOD2/CARD15 mutations and presence of granulomas in pediatric and adult Crohn's disease. *Inflamm Bowel Dis*. 2004; 10:709–14. [PubMed: 15626887]
31. Cooney R, Baker J, Brain O, et al. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med*. 2010; 16:90–7. [PubMed: 19966812]
32. Homer CR, Richmond AL, Rebert NA, et al. ATG16L1 and NOD2 interact in an autophagy-dependent antibacterial pathway implicated in Crohn's disease pathogenesis. *Gastroenterology*. 2010; 139:1630–41. 1641 e1–2. [PubMed: 20637199]
33. Travassos LH, Carneiro LA, Ramjeet M, et al. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol*. 2010; 11:55–62. [PubMed: 19898471]
34. Melmed GY, Targan SR. Future biologic targets for IBD: potentials and pitfalls. *Nat Rev Gastroenterol Hepatol*. 2010; 7:110–7. [PubMed: 20134493]
35. Danese S. New therapies for inflammatory bowel disease: from the bench to the bedside. *Gut*. 2012; 61:918–32. [PubMed: 22115827]
36. Chambers TJ, Morson BC. The granuloma in Crohn's disease. *Gut*. 1979; 20:269–74. [PubMed: 447106]

37. Wolfson DM, Sachar DB, Cohen A, et al. Granulomas do not affect postoperative recurrence rates in Crohn's disease. *Gastroenterology*. 1982; 83:405–9. [PubMed: 7084618]
38. Rubio CA, Orrego A, Nesi G, et al. Frequency of epithelioid granulomas in colonoscopic biopsy specimens from paediatric and adult patients with Crohn's colitis. *J Clin Pathol*. 2007; 60:1268–72. [PubMed: 17293387]
39. Cervantes F, Mauro M. Practical management of patients with chronic myeloid leukemia. *Cancer*. 2011; 117:4343–54. [PubMed: 21413002]
40. Colmegna I, Ohata BR, Menard HA. Current understanding of rheumatoid arthritis therapy. *Clin Pharmacol Ther*. 2012; 91:607–20. [PubMed: 22357455]
41. Dall'era M, Chakravarty EF. Treatment of mild, moderate, and severe lupus erythematosus: focus on new therapies. *Curr Rheumatol Rep*. 2011; 13:308–16. [PubMed: 21584692]

Figure 1. CD-associated *NOD2* risk alleles are associated with abnormal Paneth cell morphology (A–C) Lysozyme immunostaining (red) was performed to visualize and score Paneth cell secretory granule morphology from cases with 0 ($n = 29$), 1 ($n = 25$) or 2 ($n = 5$) CD-associated *NOD2* risk variants. Paneth cells were scored as normal if they contained numerous small ($\sim 1 \mu\text{m}$), lysozyme-positive apically located granules. Disordered Paneth cells contained lysozyme-positive granules of normal size and quantity, but had some basally located granules. Diminished Paneth cells contained <10 granules, with the remaining granules frequently enlarged or fused. Diffuse Paneth cells did not contain any secretory granules and had diffuse lysozyme staining throughout their cytoplasm. (A) Representative images of lysozyme immunostaining. Nuclear counterstain, DAPI (blue). Well-oriented Paneth cells are outlined in white and the scored phenotype is indicated: normal (N), disordered (Dis), diminished (Dim), diffuse (Dif) or excluded granule (Exc). Bars = $10 \mu\text{m}$. (B) Representative images of Paneth cells with the excluded granule phenotype (arrowheads). (C) Quantification of Paneth cell phenotypes according to the number of CD-associated *NOD2* risk variants. (D) Quantification of the percent of abnormal Paneth cells in cases with 0 CD-associated *NOD2* susceptibility alleles ($n = 29$) or in cases that carried one allele of *R702W* ($n = 10$), *G908R* ($n = 4$) or *L1007fsXinsC* ($n = 9$). Data are presented as the mean $\pm$ s.e.m. of the percent of Paneth cells with the indicated phenotype out of the total number of Paneth cells counted for each case. Cases with the *ATG16L1 T300A* variant were excluded from these analyses. Permutation tests were performed to demonstrate the statistically significant association between the presence of CD-associated *NOD2* variants and the percentage of abnormal Paneth cells for each scoring category (*, P

< 0.05 ; **, $P < 0.01$). Mann-Whitney tests were used to demonstrate statistical difference between the presence of 1 or 2 *NOD2* risk variants and controls ($\dagger$, $P < 0.05$; $\dagger\dagger$, $P < 0.01$).

Figure 2. Abnormal Paneth cell phenotypes define a distinct subtype of CD cases

(A) Total abnormal Paneth cell proportions were examined in a larger cohort ($n = 119$) with cumulative number of *ATG16L1 T300A* and CD-associated *NOD2* variants (including common and rare variants) 0 ($n = 29$), 1 ($n = 41$), 2 ($n = 33$), 3 ($n = 10$) or 4 ($n = 6$). Linear regression analysis demonstrated a significant, positive correlation between the percent of abnormal Paneth cells per case and cumulative risk alleles. (B) Unsupervised clustering was performed for the CD cases with Paneth cell phenotype, genetic, documented environmental exposure and demographic information available. Cases are arranged in the same order along the x and y axes. The Pearson correlation was determined for each patient-patient comparison, and then patients were clustered according to similarity with the results represented as a heat map. (C) Heat map displaying the factors that significantly contribute to the identification of Subtypes 1 and 2 in (B), which were defined by low and high proportions of abnormal Paneth cells, respectively. Heat map cells are colored red and blue to indicate positive and negative correlations, respectively. PCs, Paneth cells.

Figure 3. A specific gene signature associated with an activated immune response correlates with the disordered and diffuse Paneth cell phenotypes

The 114 and 218 transcripts with expression values that correlated with the disordered and diffuse Paneth cell phenotype, respectively, were analyzed using DAVID to identify biological process gene ontology (GO) terms significantly overrepresented in each gene set. (A, B) GO terms related to the immune response are indicated (fully labeled bar graphs are shown in Figures S8 and S9). Bars are colored according to the major GO term category: biological process (gray), cellular component (white), molecular function (black). PC, Paneth cell.

Figure 4. Abnormal Paneth cell phenotypes and presence of granuloma are inversely correlated

The incidence of granuloma was determined in our expanded CD cohort (n = 107).

Correlation of granuloma incidence and cases with low (white bars) vs. high proportions (black bars) of (A) abnormal Paneth cell phenotypes (cutoff = 20%; n = 63 [$< 20\%$], n = 36 [$\geq 20\%$]), (B) 'excluded granule' Paneth cell phenotypes (cutoff = 5%; n = 71 [$< 1\%$], n = 28 [$\geq 1\%$]), (C) diminished granules (cutoff = 10%; n = 42 [$< 10\%$], n = 57 [$\geq 10\%$]) and (D) diffuse granules (cutoff = 5%; n = 92 [$< 5\%$], n = 7 [$\geq 5\%$]). Statistical significance was determined by Chi-Square test (* $P < 0.05$).

Figure 5. Abnormal Paneth cell phenotypes correlate with shorter time to disease recurrence in CD patients that received prophylaxis post-surgery

The time to disease recurrence post-resection surgery was determined in the CD cases that received prophylaxis post-surgery (n = 102). Prophylaxis included use of immunomodulators (e.g., 6-mercaptopurine, azathioprine, methotrexate) or biologics (i.e., anti-TNF therapy). Cases were assigned to the “Abnormal Paneth cells: Low” or “Abnormal Paneth cells: High” groups using a cutoff of 20% abnormal Paneth cell phenotypes. Data are presented as a survival curve with *P = 0.0200 by Log-rank test.

Figure 6. Proposed model for the use of disease-relevant phenotypes as integrative readouts for the stratification of CD patients

Abnormal Paneth cell phenotypes can synthesize multiple genetic and environmental inputs. These phenotypes are associated with molecular and pathological features and can be used to subdivide CD patients.

Table 1Baseline characteristics of CD cohort¹.

	<i>NOD2</i> Variants Absent N = 29 (%)	<i>NOD2</i> Variant(s) Present N = 30 (%)	Fisher's Exact Test <i>P</i>
Sex (Male)	13 (45)	15 (50)	0.796
Family history of IBD	9 (31)	8 (27)	0.779
Presence of stricturing and/or penetrating disease	25 (86)	24 (80)	0.731
Presence of upper small bowel disease	7 (24)	6 (20)	0.761
Presence of perianal disease	5 (17)	11 (37)	0.143
ASA treatment	18 (67)	21 (70)	0.589
Steroids treatment	22 (81)	22 (73)	1.000
Antibiotics treatment	14 (52)	17 (57)	0.606
Anti-TNF alpha treatment	15 (56)	15 (50)	1.000
Non-smoker status	23 (79)	26 (87)	0.506

¹ Patients with the *ATG16LI T300A* allele were excluded from this analysis.

ASA, acetylsalicylic acid; TNFalpha, tumor necrosis factor.