It is suggested that subtyping of complex inflammatory diseases can be based on genetic susceptibility and relevant environmental exposure (G+E). We propose that using matched cellular phenotypes in human subjects and corresponding pre-clinical models with the same G+E combinations are useful to this end. As an example, defective Paneth cells can subtype Crohn's disease (CD) subjects; Paneth cell defects have been linked to multiple CD susceptibility genes and are associated with poor outcome. We hypothesized that CD susceptibility genes interact with cigarette smoking, a major CD environmental risk factor, to trigger Paneth cell defects. We found that both CD subjects and mice with ATG16L1T300A (T300A; a prevalent CD susceptibility allele) developed Paneth cell defects triggered by tobacco smoke. Transcriptional analysis of full thickness ileum and Paneth cell-enriched crypt base cells showed the T300A-smoking combination altered distinct pathways, including pro-apoptosis, metabolic dysregulation, and selective down-regulation of the PPARγ pathway. Pharmacologic intervention by either apoptosis inhibitor or PPARγ agonist rosiglitazone prevented smoking-induced crypt apoptosis and Paneth cell defects in T300A mice and mice with conditional Paneth cell-specific knockout of Atg16l1. This study demonstrates how explicit G+E can drive disease relevant phenotype, and provides rational strategies to identify actionable targets.
Interaction Between Smoking and \textit{ATG16L1}^{T300A} Triggers Paneth Cell Defects in Crohn's Disease

\textbf{Authors:} Ta-Chiang Liu \textsuperscript{1}, Justin T. Kern \textsuperscript{1}, Kelli L. VanDussen \textsuperscript{1}, Shanshan Xiong \textsuperscript{1}, Gerard E. Kaiko \textsuperscript{1}, Craig B. Wilen \textsuperscript{1}, Michael W. Rajala \textsuperscript{2}, Roberta Caruso \textsuperscript{2}, Michael J. Holtzman \textsuperscript{3}, Feng Gao \textsuperscript{4}, Dermot P.B. McGovern \textsuperscript{5}, Gabriel Nunez \textsuperscript{2}, Richard D. Head \textsuperscript{1}, Thaddeus S. Stappenbeck \textsuperscript{1*}

\textbf{Affiliations:}

\textsuperscript{1} Department of Pathology and Immunology, Washington University School of Medicine, Saint Louis, MO 63110.

\textsuperscript{2} Department of Pathology, University of Michigan School of Medicine, Ann Arbor, MI 48109.

\textsuperscript{3} Department of Medicine, Washington University School of Medicine, Saint Louis, MO 63110.

\textsuperscript{4} Department of Surgery, Washington University School of Medicine, Saint Louis, MO 63110.

\textsuperscript{5} The F.Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048.

*To whom correspondence should be addressed:

Thaddeus S. Stappenbeck

Department of Pathology and Immunology, Washington University School of Medicine,
Abbreviations:

CD: Crohn's disease

G+E: genetic susceptibility and environmental factors

HD5: human defensin 5

IBD: inflammatory bowel disease

LCM: laser capture microdissection

SNP: single nucleotide polymorphism
Abstract:

It is suggested that subtyping of complex inflammatory diseases can be based on genetic susceptibility and relevant environmental exposure (G+E). We propose that using matched cellular phenotypes in human subjects and corresponding pre-clinical models with the same G+E combinations are useful to this end. As an example, defective Paneth cells can subtype Crohn's disease (CD) subjects; Paneth cell defects have been linked to multiple CD susceptibility genes and are associated with poor outcome. We hypothesized that CD susceptibility genes interact with cigarette smoking, a major CD environmental risk factor, to trigger Paneth cell defects. We found that both CD subjects and mice with ATG16L1T300A (T300A; a prevalent CD susceptibility allele) developed Paneth cell defects triggered by tobacco smoke. Transcriptional analysis of full thickness ileum and Paneth cell-enriched crypt base cells showed the T300A-smoking combination altered distinct pathways, including pro-apoptosis, metabolic dysregulation, and selective down-regulation of the PPARγ pathway. Pharmacologic intervention by either apoptosis inhibitor or PPARγ agonist rosiglitazone prevented smoking-induced crypt apoptosis and Paneth cell defects in T300A mice and mice with conditional Paneth cell-specific knockout of Atg16l1. This study demonstrates how explicit G+E can drive disease relevant phenotype, and provides rational strategies to identify actionable targets.
Introduction

Subtyping of complex immune diseases such as inflammatory bowel disease (IBD) has traditionally been done by categorizing clinical phenotypes and disease presentations (1, 2). Incorporating key pathogenic elements, namely genetic susceptibility and relevant environmental exposure (G+E) (3) into subclassification schemes may facilitate identifications of therapeutic targets in these subtypes. The identification of >200 susceptibility single nucleotide polymorphisms (SNPs) and recently identified prognosis-associated SNPs (4, 5) for Crohn's disease (CD, one major form of IBD) poses a challenge to this approach, as many of these SNPs have been associated with gene expression and functional changes in various cell types, such as immune (6-9) and epithelial cells (10, 11). While development of gene scores has shown promise in subtyping patients (12), such scores do not account for environmental exposures that are likely to trigger phenotypes and disease. Likewise, environmental factors identified by epidemiologic studies require relevant and functional testing in preclinical models where exposure to these environmental factors is controlled. Conversely, novel environmental factors identified in preclinical models need to be confirmed in patients.

Development of a surrogate phenotype/biomarker that can integrate the effects from both genetics and environmental factors will facilitate subtyping of IBD. In CD, morphologic patterns of small intestinal Paneth cells (Paneth cell phenotype) are a surrogate phenotype that stratifies CD into prognostically distinct subtypes (13-15). We and others have shown that in mouse models, knockout of CD-associated genes (Atg16l1, Xbp1, Irgm, Lrrk2) resulted in Paneth cell defects manifested as secretory granule abnormalities (11, 16-18) that are similar to those observed in CD subjects (11, 14, 15), with potential additive effects between genes (14). We previously showed that administration of a chronic strain of murine norovirus (MNV) could induce Paneth cell defects
in $Atg16l1^{HM}$ (hypomorph) mice, which express low levels of Atg16l1 protein (19). In human subjects, Paneth cell defects in CD are associated with microbiota changes (20) and poor clinical outcome (14, 15). Thus, Paneth cell phenotypes are a biologically and clinically relevant surrogate phenotype ideally suited for mechanistic studies and identification of potential therapeutics in CD.

One G+E trigger for Paneth cell defects in mouse models, MNV (19), as yet has no correlate in human subjects (21, 22). Therefore, our goal was to identify an environmental trigger for Paneth cells defects that occurs in both CD subjects and analogous mouse models. Among the known CD environmental risk factors (1, 23), cigarette smoking is amongst the most reproducible (23, 24). It is also associated with an aggressive disease course in patients with established CD (25). A recent study suggested potential interactions between genetics and cigarette smoking (26). Based on these findings, we hypothesized that smoking would induce Paneth cell defects in genetically susceptible CD patients. As a proof of concept, we investigated the correlation of smoking exposure, Paneth cell defects, and post-operative recurrence after ileal/ileocolonic resections in CD subjects with $ATG16L1^{T300A}$, the most prevalent CD susceptibility SNP in Caucasians (4). We then performed functional studies using the $Atg16l1^{T300A}$ mouse model to identify host factors that mediated smoking-induced Paneth cell defects. Finally, we validated rationally designed therapeutic strategies targeting these factors that resulted in Paneth cell defects.
Results

CD subjects with ATG16L1{T300A} were susceptible to smoking-associated Paneth cell defects

We found that in CD subjects (demographics described in table S1) who received ileocolonic anastomosis and post-operative immunomodulatory and/or biologics prophylactic therapy (a known confounder for outcome; n=128), smoking status and Paneth cell phenotype were prognosticators of recurrence (fig. S1), and the combination of these factors further stratified patients into prognostically distinct subgroups (Fig. 1A). In addition, CD subjects who were of the ATG16L1{T300A} genotype and who were also smokers (T300A-smoker group) showed significantly shorter time to recurrence after surgery (fig. S2). We therefore hypothesized that cigarette smoking was a trigger for Paneth cell defects in CD subjects. Given that the most common risk allele for CD susceptibility known to be associated with Paneth cell defects was ATG16L1{T300A} (11), we further hypothesized that smoking triggers Paneth cell defects preferentially in CD subjects that harbored the ATG16L1{T300A} risk allele(s). In support of this hypothesis, ATG16L1{T300A} genotype in CD subjects who were smokers was associated with a lower percentage of normal Paneth cells, whereas subjects with no risk allele (NR) were not (Fig. 1B, 1C, and Table S2). We have previously described several distinct classes of abnormal Paneth cell morphology (14, 27). We determined the distribution of each subclass of abnormal Paneth cells and found that the majority of the abnormal Paneth cells were of the D2 subclass (decreased granules) (Fig. S3), similar to previous findings in adult CD (14, 15, 27). None of the individual abnormal morphology subclasses showed a significantly different distribution across the groups; rather, the sum percentage of these abnormal classes (or, conversely, the percentage of normal Paneth cells) provided the most robust association in the T300A-smoker group (Fig. 1C).
Given that NOD2 is the other CD susceptibility gene known to be associated with Paneth cell defects in North American CD cohorts (14), we also examined the correlation between common NOD2 variants (R702W, G908R, and L1007fs) carrier status, smoking status, and Paneth cell phenotype. There were no significant changes in the percentage of normal Paneth cells in subjects carrying NOD2 variants that were smokers (fig. S4A). We further correlated the total numbers of ATG16L1T300A and NOD2 risk alleles, smoking status, and Paneth cell phenotype. There was no significant difference in the genetic burden regarding Paneth cell phenotype and smoking status (fig. S4B). Therefore, smoking-induced Paneth cell defect correlated specifically with ATG16L1T300A alleles in this cohort.

\textit{Atg16l1T300A} mice were susceptible to smoking-induced Paneth cell defects

We modeled the four patient populations above in a mouse model representative of the gene-environment (G+E) interactions by exposing \textit{Atg16l1T300A} mice (28) and wild type (WT) littermates to cigarette smoke for four weeks (Fig. 2A). Paneth cell defects were triggered only in the \textit{Atg16l1T300A} mice (Fig. 2B and table S3), recapitulating the findings in CD subjects. We also examined the distribution of each class of abnormal Paneth cells. Similar to the observations in the human cohort, the abnormal Paneth cells were predominantly of the D2 subclass, with a small percentage of D3 subclass (diminished) (fig. S5A). Increased percentages of these two subclasses of abnormal Paneth cells were largely responsible for the decreased percentage of normal Paneth cells in the T300A-smoking group (fig. S5B-D).

We also performed transmission electron microscopy to investigate potential ultrastructural changes in the Paneth cells. We found that Paneth cells of the T300A mice contained...
cytoplasmic vesicles and degenerative mitochondria (fig. S6A-C), similar to our previous observations in Atg16l1HM mice (11). Importantly, these changes were more frequent in T300A mice exposed to smoking compared to those that were not (fig. S6D). Paneth cells from CD patients that were of the ATG16L1 T300A genotype and smokers also possessed similar features (fig. S6E, F). To exclude the possibility that the Paneth cell defects in the T300A-smoking mice were the result of stress associated with the physical presence in the smoking chamber rather than exposure to cigarette smoke, T300A mice were placed in the smoking chamber and exposed to normal air pumped through the machine (i.e. no exposure to cigarette smoke). Physical presence in the smoking chamber alone with exposure to normal air did not result in Paneth cell defects in these mice (fig. S7).

We next altered the duration of smoking to determine the impact on Paneth cell phenotypes. We found that a two-week exposure was sufficient to trigger Paneth cell defects in Atg16l1T300A mice, but that exposure beyond two weeks did not increase the percentage of defective Paneth cells (Fig. 2C). We also tested the durability of smoking-induced Paneth cell defects. Paneth cell phenotype was examined in Atg16l1T300A mice at the end of a 4-week smoking period (baseline), followed by a 2- or 4-week washout period where smoking was discontinued. Four weeks of washout (but not two weeks) were required to restore normal Paneth cell morphology (Fig. 2D). Therefore, Paneth cell defects in Atg16l1T300A mice after short-term smoking exposure were reversible upon smoking cessation.

We also determined if administration of nicotine, a major component in cigarette, would elicit similar effects on Paneth cells. We administered a daily dose of nicotine (0.7 mg/mouse/day) to mice instead of cigarette smoke. This dose is greater than the predicted absorbed nicotine dose (approximately 0.42 mg/mouse/day) achieved in the cigarette smoking experiments based on the
nicotine content of the cigarettes and known absorption kinetics (29, 30). Surprisingly, nicotine administration did not induce Paneth cell defects or crypt base apoptosis in Atg16l1T300A mice (fig. S8).

The gut microbiota did not alter smoking-induced Paneth cell defects

Paneth cell function is important for maintaining the homeostasis of the gut microbial community (31-33), and dysbiosis can develop as a downstream effect of Paneth cell defects or loss (20, 34, 35). Thus, we examined if microbiota changes occurred upstream of or as part of a feedback loop with Paneth cell defects in the context of G+E. We first compared the microbial compositions in Atg16l1T300A mice and littermates with and without smoking. There was no significant difference in microbial compositions between Atg16l1T300A mice and littermates without smoking (fig. S9A). Smoking did not result in significant changes in alpha (fig. S9B, 9C) or beta diversity (fig. S9D, 9E) in either Atg16l1T300A mice or littermates. Deeper examination of specific microbial taxa showed only limited differences between the groups of mice. For example, smoking induced relatively increased abundance of Lactobacillales and Turcibacterales and reduced the abundance of Alphaproteobacteria and Betaproteobacteria in the Atg16l1 T300A mice, whereas it only increased the abundance of Coriobacterales and Turcibacterales in the WT littermates (fig. S9F, 9G). Therefore, smoking only modestly altered the composition of the gut microbiota, regardless of genotype.

Because we did detect small differences in the microbiota composition that depended on smoking in Atg16l1T300A mice, we functionally tested the microbiota for its ability to induce the Paneth cell defects. We co-housed Atg16l1T300A mice and WT littermates exposed to smoking
(microbiota donors) with non-smoked mice of the same genotypes (microbiota recipients) (Fig. 3A). Recipients were pre-treated with antibiotics (36) to allow successful colonization of donor microbiota. The recipients showed indistinguishable microbiota compositions as their respective donors after 4 weeks (fig. S10). Co-housing of donors and recipients did not induce Paneth cell defects in $Atg16l1^{T300A}$ recipients (Fig. 3B). Therefore, the limited differences in microbial composition observed with smoking in the $Atg16l1^{T300A}$ mice did not contribute to Paneth cell defects.

*Smoking-induced Paneth cell phenotype did not correlate with lung or systemic inflammation*

To determine if Paneth cell defects could be due to secondary changes of lung and/or systemic inflammation, we further examined the lungs for histopathology and serum for inflammatory markers. No overt inflammation was seen in the lungs in any of the mice (fig. S11), consistent with a previous report that longer smoking exposure may be required to elicit lung inflammation (37). Likewise, none of the samples showed detectable TNFα in serum (fig. S12A). In addition, there was no significant difference in serum myeloperoxidase, RAGE, CXCL1, CXCL2, IL6, or IL1β levels between the groups (fig. S12B-G). Therefore, there was no correlation between lung or systemic inflammation and G+E-associated Paneth cell defects.

*Natural MNV infection was an unlikely cause for smoking-induced Paneth cell defects*

We previously showed that administration of a chronic strain of MNV could induce Paneth cell defects in $Atg16l1^{HM}$ mice (19). To exclude that natural MNV infection, not infrequently encountered in animal facilities (19), could result in Paneth cell defects in $Atg16l1^{T300A}$ mice
exposed to smoking, we also determined the MNV titers in fecal samples. We found that 16% of mice were indeed infected with MNV. However, among the $Atg16l1^{T300A}$ mice exposed to smoking, there was no significant difference in the percentages of normal Paneth cells between MNV-uninfected and MNV-infected mice (fig. S13). Therefore, natural MNV infection was an unlikely cause for the Paneth cell defects observed in this study.

**Smoking and $Atg16l1^{T300A}$ genotype interaction led to unique host transcriptomic changes**

The lack of a causative link between the microbiota, systemic inflammatory markers, and Paneth cell defects indicate that the underlying mechanisms most likely stemmed from the host intestine *per se*. To comprehensively analyze the effect of G+E in all ileal cell types, we performed global RNA-seq using mRNAs isolated from full thickness ileal sections from $Atg16l1^{T300A}$ mice and WT littermates ± smoking. Identified transcriptomic differences were categorized as associated with either genetics ("G" patterns), smoking ("E" patterns), or combinatorial effects of genetics and smoking ("G+E" patterns) (fig. S14A).

The "G+E" patterns could be further classified into "T300A-smoking" and "WT-smoking" patterns (Fig. 4A, 4B, and fig. S13B-E). The "T300A-smoking" pattern was striking, as it included signatures associated with the promotion of apoptosis, and the down-modulation of insulin signaling, predominantly through $Ppara/g$ regulation (Fig. 4A, fig. S14B-E and table S4). Interestingly, $Pparg$ activity has been shown to modulate Paneth cell function during high fat diet exposure (38).
The "WT-smoking" pattern included induction of genes associated with \textit{Hnf4\textalpha} and \textit{Ppar} regulation in addition to lipid and amino acid metabolism (Fig. 4B, fig. S14F, and table S5). These molecular responses to smoking were either attenuated or not activated in the \textit{Atg16l1}^{T300A} mice.

The "G" patterns specific to the \textit{Atg16l1} \textit{T300A} ileum included genes that function in metabolism and complement activation (Fig. 4C, fig. S14G, S14H, and table S6). We had previously found that deletion of \textit{Atg16l1} in mouse intestinal epithelial cells targets metabolism and acute phase reactants like complement (11). The "E" patterns enriched with smoking included genes that function in cell death, interferon signaling, cyclooxygenase pathway, and gluconeogenesis (Fig. 4D, fig. S14I, 14J, and tables S7, S8). Of note, smoking can induce apoptosis in a mouse model of emphysema (39).

\textit{Apoptosis was a central mechanism behind smoking-associated Paneth cell defects in CD subjects and mice with ATG16L1^{T300A}}

The analysis of the transcriptomics data suggested that apoptosis-associated cell death might be critical mediator of G+E-induced Paneth cell defects. We found that in CD subjects, the highest level of crypt base apoptosis was present in \textit{Atg16l1}^{T300A} smokers as compared to all other groups (Fig. 5A, fig. S15A, and table S2). Further analysis showed that Paneth cells themselves were sensitive to apoptosis in \textit{ATG16L1}^{T300A} smokers through defensin 5 (HD5)/TUNEL co-localization (Fig. 5B, table S2). In addition, \textit{ATG16L1}^{T300A} smokers but not NR smokers also had lower Paneth cell numbers/crypt (Fig. 5C and table S2). The effects of smoking and genotype showed no detectable effect on crypt proliferation, the other major function of epithelial cells in the crypt base (fig. S16A, S16B, and table S2).
Atg16l1T300A mice exposed to smoking also showed increased crypt apoptosis (Fig. 5D, fig. S15B, and table S3), and specifically increased Paneth cell apoptosis (Fig. 5E and table S3), confirming that apoptosis-associated cell death in the crypt base compartment was directly linked to smoking-induced Paneth cell defects. Of note, the Paneth cells that co-expressed cleaved caspase-3 and lysozyme were exclusively of the abnormal morphology (human: 92% D2, 8% D3; mice: 97% D2, 3% D3). Atg16l1T300A mice exposed to smoking tended to have fewer Paneth cells as compared to smoked littermate controls but this was not significant (Fig. 5F and table S3). We also found that smoking did not induce apoptosis in villus epithelial cells (Fig. 5G, fig. S15C), further demonstrating that the specificity of smoking and genotype on Paneth cells. Finally, crypt proliferation was not altered by smoking of mice with either genotype (fig. S16C, S16D, and table S3). Therefore, crypt base apoptosis was a specific response to G+E, and the process did not elicit compensatory alterations in proliferation. To determine if apoptosis mediated the smoking-induced Paneth cell defects in Atg16l1T300A mice, we administered pan-caspase inhibitor Z-VAD-FMK to the Atg16l1T300A mice. Z-VAD-FMK prevented the Paneth cell defects (Fig. 5H) and crypt apoptosis (Fig. 5I) induced by smoking, confirming apoptosis is upstream of Paneth cell defects. In addition, we also determined the potential role of necroptosis in mediating Paneth cell defects (40), as a recent report has suggested a link between Atg16l1 and necroptosis (41). Administration of the necroptosis inhibitor nec-1 did not prevent the Paneth cell defect (fig. S17A) nor crypt base apoptosis (fig. S17B) phenotypes in T300A-smoked mice, confirming the lack of association with necroptosis in this experimental design.
Repressed Pparg activation resulted in smoking-induced crypt apoptosis and Paneth cell defects

The unique "G+E" patterns in the full thickness ileal transcriptomic analysis demonstrated an attenuation or repression of Ppara/g activation in the Atg16l1T300A mice exposed to smoking as compared to WT animals (fig. S14C-E, Table S9). These data suggested the possibility that Ppara/g activation in WT-smoked mice may be protective of the normal Paneth cell phenotype. In a subsequent analysis of the “G+E” patterns using Enrichr to probe the GEO drug perturbations datasets (42, 43) (fig. S18A), the WT-smoked pattern genes uniquely matched compounds with multiple highly-significant adjusted P-value entries (adjusted \( P < 0.01 \)) (fig. S18B). The analysis revealed that selective PPARγ agonists, including rosiglitazone, pioglitazone, and troglitazone (especially among the drug-like molecules) regulate the genes in this potentially protective pattern, whereas fibrates (PPARα agonists) were not detected in this analysis. This suggested the general Ppara/g signature observed in the pathway-level analysis may be more specific to Pparg.

To further justify analysis of this pathway in Paneth cells, we performed global transcriptional analysis of crypt base material (enriched for Paneth cells) from these mice collected by laser capture microdisseciton (LCM) (Fig. 6A). We found that Atg16l1T300A mice exposed to smoking showed significantly diminished expression of many Pparg-associated genes as compared to the other groups of mice in this experiment (Fig. 6B and table S10). We also found a similar enrichment of downregulated PPARγ pathway genes in two specific G+E groups from our previous LCM-procured Paneth cell datasets: 1) Atg16l1HM mice infected with MNV (19) (Fig. 6B and table S11), and importantly, 2) CD subjects who were smokers and ATG16L1T300A genotype (14) (Fig. 6B and table S12). These data collectively suggest that PPARγ pathway is a central mechanism closely linked to Paneth cell defects in CD subjects and relevant mouse models as a result of G+E interaction.
We next functionally tested the role of PPAR\(\gamma\) pathway in smoking-induced Paneth cell defects. Administration of the PPAR\(\gamma\) agonist rosiglitazone rescued the smoking-induced Paneth cell defects (Fig. 6C) and crypt base apoptosis (Fig. 6D) in the \textit{Atg16l1}\textsuperscript{T300A} mice. In parallel, we also generated mice with intestinal epithelium-specific \textit{Pparg} deletion (\textit{Pparg/Villin-Cre} mice). These mice showed reduced percentage of normal Paneth cells, reduced Paneth cell numbers/crypt, and increased crypt base apoptosis compared to the \textit{Pparg}\textsuperscript{0/0} littermate controls (fig. S19). Therefore, the PPAR\(\gamma\) pathway is a critical mediator of crypt apoptosis and Paneth cell defects.

\textit{G+E interactions directly affected Paneth cells and precursors}

The crypt base transcriptomic data also suggested that the G+E interaction induced effects directly on Paneth cells themselves. To test this hypothesis \textit{in vivo}, we utilized a conditional knock-out model where \textit{Atg16l1} was deleted exclusively in Paneth cells (\textit{\(\alpha\)-defensin-4-IRES-Cre Atg16l1}\textsuperscript{\textsuperscript{fl/fl}} mice; herein termed PC-Cre+ mice) (44). As seen in fig. S20A, the PC-Cre+ mice elicited Paneth cell defects when exposed to cigarette smoking, whereas \textit{Atg16l1}\textsuperscript{\textsuperscript{fl/fl}} mice (herein termed PC-Cre- mice) did not, as expected. The PC-Cre+ mice also showed increased crypt base apoptosis (fig. S20B), and increased Paneth cell apoptosis (fig. S20C). Finally, rosiglitazone administration also prevented smoking-induced Paneth cell defect and crypt apoptosis in PC-Cre+ mice (Fig. 6E, 6F). These data further support the notion that the G+E effect can act directly on Paneth cells and their \textit{Defa4}-expressing precursors, although we cannot exclude Paneth cell-independent mechanisms with certainty.

\textit{TNF antagonism rescued smoking-induced Paneth cell defects}
We previously suggested that tumor necrosis factor alpha (TNFα), a major therapeutic target in CD, may be an important mediator of Paneth cell defects in CD subjects (15). We have also shown that anti-TNFα treatment can ameliorate the intestinal pathology in Atg16l1HM mice infected with MNV and treated with DSS (19). Interestingly, it has also been shown that PPARγ antagonism in preadipocytes conferred increased sensitivity to TNFα-induced apoptosis (45), and that treatment with PPARγ agonist blocked TNFα-induced apoptosis in vitro (46), suggesting defective PPARγ pathway could prime the host tissue to TNFα-induced apoptosis. In Atg16l1T300A mice, anti-TNFα treatment prevented Paneth cell defects (Fig. 7A) and crypt apoptosis (Fig. 7B). Anti-TNFα administration did not alter the expression of genes involved in the PPARγ pathway (fig. S21), suggesting that TNFα acts downstream of PPARγ and mediates apoptosis induction. We further tested the role of TNFα by crossing Atg16l1T300A mice to TNF receptor 1 (Tnfr1)–deficient mice. Atg16l1T300A/Tnfr1−/− mice were resistant to smoking-induced Paneth cell defects (Fig. 7C) and crypt base apoptosis (Fig. 7D). Therefore, the TNFα signaling pathway is a key mediator and therapeutic target for smoking-induced crypt apoptosis and Paneth cell defects in Atg16l1T300A mice.

Discussion

We previously showed that Paneth cell phenotypes are associated with CD genotypes, microbiota composition, a pathologic hallmark, unique transcriptomic profiles, and clinical outcome (11, 14, 15, 19, 20). However, clinically relevant environmental trigger(s) and the mechanism(s) driving Paneth cell defects were unclear (Fig. 8A). In the current study, we show
that relevant environmental stimuli can trigger Paneth cell defects in genetically susceptible hosts, confirming this phenotype as a unique readout to functionally test potential genetic and environmental interactions. We also show that the G+E interactions resulted in previously unpredicted intestinal metabolic dysregulation, leading to crypt base apoptosis and Paneth cell defects mediated by PPARγ that could additionally be blocked by anti-TNFα (Fig. 8B). Finally, we show that the G+E interactions directly affect Paneth cells and their Defa4-expressing precursors, although other cell types may also be affected. The Paneth cells that underwent apoptosis were exclusively of the abnormal morphology patterns. Therefore, the abnormal Paneth cells may undergo apoptosis; however, they may also revert to normal morphology once smoking is discontinued.

We focused on Atg16l1T300A, as a knock-in model exists which possesses the same polymorphism as CD patients, providing a mechanistic advantage over the whole gene knockout models in select cell types. Similar mouse models for Nod2 polymorphisms exist, but this allele is much less common in CD cohorts of European ancestry (4). Other CD susceptibility genes associated with abnormal Paneth cells do not yet have mouse models of their respective genetic polymorphisms. In addition, while we have previously shown that Paneth cell defects are induced in hypomorphic Atg16l1 mice after murine norovirus infection, we are unable to yet identify such a link in CD patients (21). Even so, smoking is a clinically relevant environmental trigger (23, 47). We show that the combination of relevant host genetic and environmental factors can provide insight into disease pathogenesis and therapeutic targets, as has been recently demonstrated in studies such as nonalcoholic fatty liver disease (48, 49).

The majority of the microbiome studies on the effect of smoking have centered on the oral cavity microbiota (50). A recent population-based microbiome study showed that smoking status
and history showed modest effect on Bray-Curtis distance without significant associations for individual species or pathways (51). A small cross-sectional study of CD patients showed that smoking is associated with reduced microbiota diversity, with reduced abundance of limited taxa at genera level (52). Our in vivo study was consistent with these findings. Along with the failure of horizontal transmission of Paneth cell defects in co-housing experiments, our data strongly suggest that microbiota changes were not a cause of Paneth cell defects. Overall, this supports our working model (53) that Paneth cell defects promote dysbiosis only in the presence of active inflammation. In addition, recent studies have suggested that necroptosis modulates Paneth cell function (40, 41). In our study, smoking-induced Paneth cell defects in T300A mice was only rescued by apoptosis inhibition but not necroptosis inhibition, suggesting that different injuries may illicit different predominant cell death responses.

The value of the unbiased global transcriptomics approach is highlighted by the identification of Pparg as a central mediator in the T300A-smoking patterns, a finding that was not deducible from examining the effect of each single factor alone. PPARγ has been shown to be down-regulated in smoking-associated emphysema (54), suggesting that the modulation of this pathway by smoking is likely a general tissue response. In addition, Pparg is linked to reduced Paneth cell numbers in mice fed with a high fat diet (38). Given that agents targeting this pathway (e.g., rosiglitazone) are readily available for routine clinical use, further clinical studies using these agents to treat CD patients who are smokers with Paneth cell defects will validate the importance of this pathway in CD. Furthermore, metabolic dysregulation is tightly connected with the TNFα-associated apoptosis pathway. Therefore, previously recognized important genetic factors, an environmental factor, and inflammatory pathways converged to affect Paneth cell health and clinical prognosis. Of note, our complementary approaches (crypt base LCM, PC-Cre+ mice)
support the transcriptomics analysis from full thickness ileum that Paneth cells are the main target of the G+E effect.

Our data also suggest that smoking cessation may be beneficial for ATG16L1T300A CD subjects with smoking-associated Paneth cell defects. Other potential intervention approaches include nicotine patch, PPARγ agonists, and anti-TNFα. Anti-TNFα is a major treatment modality for CD (1), and rosiglitazone has been shown to be efficacious in ulcerative colitis (another major form of IBD) (55). Our data indicates that Paneth cell phenotype may be used to stratify CD patients who may benefit from these therapies. One limitation of our study is that due to the physical restraint of the smoking chamber, the cigarette smoking experiments could not exceed 6 weeks. Therefore, the effect of long-term cigarette smoking on Paneth cell defects, in particular the reversibility of the approaches described above, is unclear. In addition, while neither nicotine, lung pathology, nor systemic inflammatory signals were shown to affect Paneth cells in this model, it is possible that the changes in gut transcriptomics and subsequent Paneth cell defects are the results of processes initiated external to the diseased/target organ (gut), similar to rheumatoid arthritis (56), where cigarette smoking has been shown to alter transcriptomic changes of the joints (57). One such possibility could be the lung-gut axis, such that cigarette smoking affects the lung on a molecular level (potentially through lung microbiome and/or metabolites) (58, 59), which may in turn affect gut transcriptomics. In addition, the PPARγ pathway has best been studied in liver, skeletal muscle, and adipocytes in the context of metabolism (60). For example, based on the known PPARγ upstream regulatory mechanisms, we also speculate that the combination of G+E could affect either fatty acid binding proteins or fatty acid transporters, which would potentially involve a liver-gut crosstalk (61).
In summary, we show that genetic and environmental factors synthesize to trigger unique biologic processes resulting in a clinically relevant phenotype. Our data also provide complementary mechanistic insights into the role of Paneth cells in mediating CD pathogenesis (16, 62) and identification of actionable therapeutic targets.
Materials and Methods

Study design

The overall objective of our study was to determine the associations between the \textit{ATG16L1T300A} genotype and exposure to cigarette smoking in triggering Paneth cell defects. For human subjects, based on the prevalence of adult CD subjects harboring Type I Paneth cell phenotype (14, 15) and the natural history of CD after resection (1), 90 subjects were required to achieve a power of 80%. For in vivo experiments, we used a previously described mouse strain (\textit{Atg16l1T300A}) that is known to possess Paneth cell defects (28). Cigarette smoking was performed following previously described protocol (63, 64), with the cigarette filters removed. Paneth cell analysis was performed using immunofluorescence (14, 15) on distal ileum. All the experiments were performed in several replicates over the course of 2 years. At least 3 to 6 biological replicates were used for each group/experiment. The mice were randomized, and the investigator performing the histologic analysis was blinded to the sample identity. The design for microbiome studies included proper littermate controls and co-housing (65), and microbiome composition was analyzed using 16S rRNA sequencing. Transcriptomic analysis was performed using RNA-seq. All data were included (no outliers were excluded). Additional details including the total numbers per study group are included in the respective figure legends.

CD Subjects

CD subjects who underwent ileocolectomy between 1999-2010 at Washington University, St. Louis, or Cedars-Sinai Medical Center, Los Angeles, were previously described (14). A second CD cohort from Washington University composed of consecutive CD patients who underwent
ileocolectomy between 2011-2013 were additionally included. De-identified tissue samples from ileal resection margins that were free of acute inflammation were used for Paneth cell phenotype analysis.

The following information was retrieved from the medical record: gender, age at operation, smoking history (never smoker vs. active/ex-smokers), medication history (including immunomodulators and biologics), endoscopic findings at the first visit 6-12 months after surgery. Recurrence was defined by endoscopy (Rutgeert’s score ≥ i2). The genotypes of the patients were obtained using immunochip (14) or through Taqman genotyping assay (Thermo Fisher) with genomic DNA extracted from formalin-fixed, paraffin-embedded tissue based on the vendor's instructions.

**Mice**

*Atg16l1T300A* mice have been described before (28) and were gifts from Dr. Ramnik Xavier (Harvard Medical School, Boston, USA), and heterozygotes were used to breed *Atg16l1T300A* and littermate controls. *Atg16l1T300A* mice were crossed with *Tnfrsf1atm1/imx* mice (*Tnfr1* knock out; The Jackson Laboratory #3244) to generate *Atg16l1 T300A/Tnfr1-/-* mice. *α-defensin-4-IRES-Cre* mice were generated by introducing Cre recombinase gene driven by the *α-defensin-4* promoter in the ES cells (44). The *α-defensin-4-IRES-Cre* (PC-Cre) mice was subsequently used to cross with *Atg16l1 β/β* mice to generate *α-defensin-4-IRES-Cre* (PC-Cre) *Atg16l1 β/β* mice. Mice with intestinal epithelium-specific knockout of *Pparg* were generated by crossing *Pparg β/β* mice (The Jackson Laboratory #004584) with *Villin-Cre* mice (The Jackson Laboratory #4586). All mice
were on a C57BL/6 genetic background. The distal ileum of the mice (distal 3 cm) was used for analysis in this study.

Mouse treatments

Four- to six-week-old mice were exposed to cigarette smoking 4 cigarettes/day for 5 days/week using Kentucky research cigarette 3R4F (with filters removed) and a previously described protocol (64, 66). Mice were exposed to 2 weeks of smoking unless otherwise indicated and then sacrificed for tissue collection. For co-housing experiments, \textit{Atg16l1}\textsuperscript{T300A} and WT littermates designated as microbial recipients were exposed to an antibiotics cocktail of vancomycin, neomycin, ampicillin, and metronidazole for 2 weeks (36), followed by co-housing with mice of the same genotype that were exposed to cigarette smoking (microbial donors). Co-housing lasted 4 weeks, and the microbial donors continued to be exposed to cigarette smoking during this period. For the nicotine experiment, nicotine was added into the drinking water at a concentration of 0.1 mg/ml for 4 weeks. In other experiments to study effects of pharmacological agents in preventing cigarette smoking-induced Paneth cell defect, mice with \textit{Atg16l1}\textsuperscript{T300A} genotype were exposed to smoking for two weeks. During this period, the mice were administered either pan-caspase inhibitor Z-VAD-FMK 10 mg/kg/day intraperitoneally (ApexBio; Houston, TX) or Ultra-LEAF\textsuperscript{TM} anti-mouse TNF-\(\alpha\) antibody 0.5 mg/mouse/injection, 2 injections/week (Biolegend; San Diego, CA). For nec-1 inhibitor administration, mice received intraperitoneal administration of necrostatin (Sigma-Aldrich) at 4 mg/kg/day for 2 weeks. For rosiglitazone administration, mice received daily oral gavage with either PBS or rosiglitazone (Sigma-Aldrich) at a dose of 20 mg/kg/day for 2 weeks.
Paneth cell phenotype analysis

Lysozyme and defensin 5 immunofluorescence was interpreted (by T.C.L.) as described previously (11, 14, 19, 20, 27). For both human and mouse samples, each Paneth cell was classified into normal or one of the 5 abnormal categories, including: disordered (abnormal distribution and size of the granules), diminished (≤10 granules), diffuse (smear of lysozyme or defensin within the cytoplasm with no recognizable granules), excluded (majority of the granules do not contain stainable material), and enlarged (rare, megagranules)(14, 27). The last two categories were only observed in human samples. The Paneth cell phenotype of CD subjects (used for outcome correlation) were then defined by the percentage of total abnormal Paneth cells in the sample. Type I Paneth cell phenotype was defined as ≥ 20% of total Paneth cells showing abnormal morphology patterns, whereas Type II Paneth cell phenotype was defined as < 20% of total Paneth cells showing morphologic defects (14).

Statistics

Clinical outcome correlation was performed using Log-rank test. For analysis between different genotype and smoking exposure combinations, Kruskal-Wallis tests followed by Dunn's tests between groups was performed. For Paneth cell phenotype and various cellular readout comparisons in mouse experiments, two-way ANOVA followed by Tukey's multiple comparison adjustment was used. For microbiome studies, principle coordinate analysis was performed using ANOSIM with 999 permutations. Relative OTU abundance data were input into LEfSe to determine biomarkers with significant linear discriminant analysis effect size (67).
determination of sample size and data analysis for animal studies followed the general guideline of Festing and Altman (68). Based on the law of diminishing returns, Mead recommended that a degree of freedom (DF) of 10-20 associated with error term in an ANOVA will be adequate to estimate preliminary information (69). All tests were two-tailed and a $P$ value of $<0.05$ was considered significant. Data were plotted and analyzed using GraphPad Prism (version 6.05) and SAS version 9.4 (SAS Institute, Cary, North Carolina). Data represent mean ± SEM.

**Study Approval**

The study was approved by the Institutional Review Boards of Washington University School of Medicine (St. Louis) and Cedars-Sinai Medical Center (Los Angeles). Subjects provided written informed consent. The animal studies were approved by the ethical committee at Washington University School of Medicine.
Author contributions: T-C.L, D.P.B.M., T.S.S designed the study. T-C.L, J.K acquired the data. D.P.B.M. recruited patients. M.H. provided experiment instruments. M.R., R.C. and G.N. provided mouse strain. T-C.L, K.L.V, F.G., R.H., and T.S.S. analyzed data. TCL and TSS drafted the manuscript. All the co-authors agreed on the content of the manuscript.

Acknowledgement

Funding: The study was supported by The Leona M and Harry B Helmsley Charitable Trust (D.P.B.M., T.S.S.), Doris Duke Charitable Foundation grant 2014103 (T-C.L.), Crohn's and Colitis Foundation grant 274415, NIH grants 1R56DK095820 (D.P.B.M., T.S.S.), P01 DK046763, and U01 DK062413 (D.P.B.M), K01DK109081 (K.L.V.), UL1 TR000448 (T-C.L.).

The authors thank the Genome Technology Access Center at Washington University School of Medicine for help with RNA-seq analysis. The Center is partially supported by NCI Grant #P30 CA91842 and by ICTS/CTSA Grant# UL1 TR000448 from the NIH. The authors also thank Dr. Wandy Beatty in the Molecular Microbiology Imaging Facility for assistance with TEM, and Dr. Robert Schmidt for interpretation of the TEM images.

Competing interests: The authors declare no financial interests.

Data and materials availability: The accession number for the Atg16l1T300A mouse study full thickness ileal RNA-seq data deposition at ArrayExpress is E-MTAB-5707. The accession number for the Atg16l1T300A mouse study laser capture microdissection-procured Paneth cell microarray data deposition at ArrayExpress is E-MTAB-6168. The accession number for the 16s rRNA sequencing data deposition for Atg16l1T300A mice at ArrayExpress is E-MTAB-5717. The
accession number for the 16s rRNA sequencing data deposition for cohousing experiments at ArrayExpress is E-MTAB-5720.
 References:


Fig. 1. CD subjects with ATG16L1T300A genotype (T300A) were more susceptible to cigarette smoking-associated Paneth cell defects. (A) In a cohort of CD subjects (n = 186) who underwent ileocolonectomy, 126 received postoperative prophylaxis. Within this prophylaxis subset, smokers with Type I Paneth cell phenotype (<80% Paneth cells with normal granule morphology) showed the shortest time to disease recurrence (P = 0.0183 by Log-rank test). (B) Representative defensin 5 immunofluorescence (scale bar: 10 μm). Asterisks indicate abnormal Paneth cells. NR: non-risk. Data represent mean ± SEM. (C) Cigarette smoking was associated with lower percentage of normal Paneth cells in patients with ATG16L1T300A allele(s), while no significant Paneth cell defects were seen between NR patients with or without smoking history (overall P =0.001). NR nonsmokers n=25, NR smokers n=14, T300A nonsmokers n=84, T300A smokers n=62. Data were analyzed by Kruskal-Wallis test followed by Dunn’s multiple comparison tests between groups. P values for comparisons between each group are shown in table S2. *: P < 0.05, **: P < 0.01.
**Fig. 2.** Atg16l1<sup>T300A</sup> mice were more susceptible to Paneth cell defects after exposure to cigarette smoking. (A) Schematic illustration of experimental design. Atg16l1<sup>T300A</sup> (T300A) mice and wild type (WT) littermates were treated with or without cigarette smoking for 4 weeks and Paneth cell morphology was assessed. (B) Smoking induced more Paneth cell defects specifically in T300A mice (overall $P<0.0001$). (A, B) WT-nonsmoked: $n=12$; WT-smoked: $n=21$; T300A-nonsmoked: $n=19$; T300A-smoked: $n=25$. Results of 6 independent experiments. Data were analyzed by 2-way ANOVA. (C) Two weeks of cigarette smoking was sufficient to induce Paneth cell defects in T300A mice ($P=0.0054$), while no additional Paneth cell defects were seen with longer exposure time up to six weeks ($P>0.9999$). Non-exposed: $n=17$, 2 weeks: $n=10$, 4 weeks: $n=10$. (D) After 4 weeks of cessation to cigarette exposure, the percentage of normal Paneth cells of the T300A mice returned to a level comparable to unexposed status ($P=0.0027$). Baseline: $n=6$, 2 weeks: $n=7$, 4 weeks: $n=7$. (C, D) Data were analyzed by one-way ANOVA followed by Mann-Whitney tests between groups. *: $P<0.05$, **: $P<0.01$, ***: $P<0.001$, ****: $P<0.0001$. (B-D) Data represent mean ± SEM.
Fig. 3. Smoking-associated Paneth cell defects in $Atg16l1^{T300A}$ mice was not horizontally transmissible by cohousing. (A) Schematic illustration of experimental design. Smoked $Atg16l1^{T300A}$ (T300A) mice and wild type (WT) littermates were used as microbiota donors, and cohoused with nonsmoked, antibiotics-pretreated mice of the same genotypes as recipients. Cohousing lasted 4 weeks, during which the donors continued to receive exposure to cigarette smoking. (B) The Paneth cell defects of the $Atg16l1^{T300A}$ microbiota donor mice did not transfer to recipient mice. WT donors: n=5, WT recipients: n=15, T300A donors: n=5, T300A recipients: n=17. Data analyzed by Kruskal-Wallis test followed by Dunn's multiple comparison tests between groups. *: $P < 0.05$, **: $P < 0.01$. Data represent mean ± SEM.
Fig. 4. Unique transcriptomic patterns associated with (A, B) combination from both genetics and smoking interactions (G+E patterns), (C) genetics alone (G patterns), or (D) smoking exposure alone (E patterns). Full thickness ileum from Atg16l1T300A mice and WT littermates with or without cigarette smoking exposure was analyzed by RNA-seq. X axis in all panels represent log P values. Significantly enriched pathways are shown.
Fig. 5. Paneth cell defects were mediated by apoptosis. (A) Smoking was associated with more crypt base apoptosis in CD subjects with the *Atg16l1*<sup>T300A</sup> genotype (*P*<0.0001), compared to NR subjects (*P*>0.9999). (B) Smoking was associated with more apoptotic Paneth cells in CD subjects with *ATG16L1*<sup>T300A</sup> (*P*=0.01) compared to NR subjects (*P*>0.9999). HD5: defensin 5. (C) *ATG16L1*<sup>T300A</sup> subjects who were smokers had reduced Paneth cell numbers/crypt (*P*=0.0103) compared to NR subjects (*P*>0.9999). (A-C) Sample sizes and data analysis as in Fig. 1. (D) In mice, smoking induced more profound crypt base apoptosis specifically in *Atg16l1*<sup>T300A</sup> mice (*P*<0.0001). (E) More Paneth cells were undergoing apoptosis in *Atg16l1*<sup>T300A</sup> mice exposed to cigarette smoking (*P*=0.0018). (F) Smoking did not induce significant alterations in Paneth cell numbers/crypt, irrespective of genotype (*P*=0.0948). (G) Smoking did not induce increased apoptosis in the villi, irrespective of genotype (*P*=0.5058). (D-G) Sample sizes and data analysis as in Fig. 2. Pan-caspase inhibitor Z-VAD-FMK administration prevented smoking-induced (H) Paneth cell defects and (I) crypt base apoptosis in *Atg16l1*<sup>T300A</sup> mice. (H, I) control: n=8, pancaspase inhibitor: n=10, nonsmoked: n=7. Data were analyzed by Kruskal-Wallis tests followed by Dunn's multiple comparison tests between groups. *: *P*<0.05; **: *P*<0.01, ***: *P*<0.001; ****: *P*<0.0001. Data represent mean ± SEM.
**Fig. 6.** *Pparg*-associated metabolism dysregulation resulted in crypt base apoptosis and Paneth cell defects. (A) Workflow for Paneth cell laser capture microdissection (LCM) of the mice in the T300A-smoking study. (B) Significantly downregulated *Pparg*-associated genes were found in all three datasets, including the *Atg16l1T300A* smoking mice (n=108 genes), *Atg16l1HM*-MNV infected mice (n=298 genes), and CD subjects that were *ATG16L1* T300A-smokers (n=166 genes). Adjusted *P* value for each group (by ChEA): \(^1\): 0.0044, \(^2\): 0.0004, \(^3\): 0.038. *Ppar-*γ agonist Rosiglitazone ("Rosi") treatment rescued (C) Paneth cell defects (*P* < 0.0001) and (D) crypt base apoptosis (*P* < 0.0001) in *Atg16l1T300A* mice exposed to cigarette smoking. (C, D) Total n: 9/group for nonsmoked groups, 10/group for smoked groups. Rosiglitazone administration prevented the (E) Paneth cell defects and (F) crypt apoptosis of PC-Cre+ smoked mice (n=4/group). (C-E) Statistical analysis was performed using 2-way ANOVA. \(*\): *P* < 0.05, \(*\*\*\*\*\): *P* < 0.0001. Data represent mean ± SEM.
Fig. 7. Anti-TNFα prevented Paneth cell defects. Administration of anti-TNFα prevented smoking-induced (A) Paneth cell defects ($P < 0.0001$) and (B) crypt base apoptosis ($P = 0.0011$) in Atg16l1T300A mice. (A, B) control: n=9, anti-TNF: n=10. Compared to the Atg16l1T300A mice, Atg16l1T300A/Tnfr−/− mice exposed to cigarette smoking showed significantly less (C) Paneth cell defects ($P = 0.0022$) and (D) crypt base apoptosis ($P = 0.0087$). (C, D) n = 6/group. Data were analyzed by Mann-Whitney test. **: $P < 0.01$, ****: $P < 0.0001$. Data represent mean ± SEM.
**Fig. 8.** Previous and current models of G+E resulting in Paneth cell defects. (A) Previous knowledge suggests G+E interaction may lead to Paneth cell defects, but the mechanisms were unclear. (B) Proposed mechanistic model based on the current study. Clinically relevant genetic and environmental factors of CD could interact and induce metabolism dysfunction (defective PPARγ signaling), leading to TNFα-mediated crypt base apoptosis and Paneth cell defects.