nature portfolio

Peer Review File

Identification and multimodal characterization of a specialized epithelial cell type associated with Crohn's Disease

REVIEWER COMMENTS

Reviewer #1 (expert in IBD, single-cell transcriptomics):

The manuscript by Li et al. employs a combination of single-cell RNA sequencing, spatial transcriptomics and single-molecule fluorescence in situ hybridisation to uncover a population of Crohn's disease (CD)-specific epithelial cells, termed 'LND cells' by the authors, that provide immune recruiting signals and correlate with response to anti-TNF treatment. While there have been a number of prior single-cell publications in the IBD space, as the authors argue, this manuscript uniquely takes an anatomically resolved look at the epithelial lineage during inactive and active disease. The authors provide analyses of the developmental origin, signalling partners, colocalising partners and IBD-susceptibility gene expression of LND cells. I also appreciate the inclusion of cohort details in the main text. Together, this is a convincing story that would be of interest to clinical and academic gastrointestinal fields.

Major:
1. Line 567-: The authors should control for the possibility that response to anti-TNF therapy can be predicted more generally by mature epithelial cells by also testing the transcriptional signature of mature enterocytes/colonocytes. If other mature epithelial cells cannot predict the outcome, then this would strengthen the case for late LND cells having a particular utility in anti-TNF response prediction.

Minor:
1. Given the low cell number per sample in the single-cell RNA sequencing data, it would be helpful to include a comment in the text about whether LND cells are present across all CD samples (it is quite hard to tell this from Supplementary Table S2).
2. Line 99: "CD often involves the ileum where Paneth cells are abundant, while the colon contains no or few Paneth cells" suggests that Paneth cells are a key cell in CD pathology. While this might be the case, the current manuscript does not focus on Paneth cells so perhaps change this line so that the reader isn't confused or distracted.
3. Were IgG+ plasma cells identified in the single-cell RNA seq data of CD samples?
4. Line 210: It may be interesting to list genes that change as a function of disease activity in the absorptive epithelium.
5. Line 237-240: Details of marker genes for immune cell types are repetitive of what was included earlier (~line 153).
6. Line 323: Citations for the six independent studies and that they are microarray and bulk mRNA data should be included.
7. Figure 3 panel I: Please include the number of repeats for the HCR-FISH staining in the figure legend. The same applies to Figure 4E.

Reviewer #2 (expert in IBD, gastroenterology):

The manuscript showed important insights into IBD pathogenesis, particularly in CD, a heterogeneous disease whose etiology and pathophysiology are still far from fully understood. In this scenario, the authors made much effort in compiling a broad single-cell expression

profiling of CD mucosa, in ascending colon and terminal ileum.
I found the level of data analysis impressive, very well-detailed, and comprehensive of all biological and pathophysiological questions that may be raised by such investigation. My enthusiasm was mainly raised by the elegant analysis of localization, characterization, and immune cell interactions.

The authors interestingly identify a specific epithelial subpopulation named LND based on the expression of specific genes (LCN2, NOS2, and DUOX2). This is very interesting since these cells correlate with inflammation and are specifically associated with CD.
However, despite the enormous amount of computational analysis and data produced, I believe that the work is still too descriptive.
One of the major concerns is the difficult reading for people not familiar with this kind of research and, given the broad readership of the journal, I suggest that the authors rewrite it more fluidly, so that the main messages are caught up easily and quickly.
Also, biological validations are required, at least for the LND population. Indeed, it would be helpful for the scientific audience to see whether these populations are acting in a specific way in the tissue, explaining the pathogenesis, at least in part, ultimately strengthening the biological insights provided.
Another concern is about the clinical relevance. May the author correlate these subpopulations with clinical activity or characteristics of the patients, further expanding the scientific value of their findings? Are there other correlations with these cells?
Moreover, they suggest this subpopulation as a potential biomarker. Is there a straightforward way to detect these cells and use them in clinics? A proof of concept may be sufficient at this stage.

Overall, although the work and the analysis provided are very well addressed, discussed, and valuable, I believe that the manuscript needs to be shortened, simplified, and carry more mechanistic proofs to endow it with importance in the field of IBD. As it is, it remains a descriptive work, unfortunately far from the journal requirements. Indeed, considering that the main output of the manuscript is descriptive, I would refer to the work as an atlas, rather than a scientific contribution expanding the mechanistic knowledge of the disease pathogenesis.

Reviewer #3 (expert in IBD, gastroenterology, intestinal epithelium):

Description:
The authors have submitted a paper that describes the results of transcriptomic analysis of the terminal ileum and ascending colon of 170 tissue samples obtained 65 patients with Crohn's disease (CD) and 18 subjects with no history of inflammatory bowel disease. The most interesting aspect of the work is the discovery of a purported novel cell type, termed "LNG" for lipocalin, NOS2, and DUOX2, which the authors state co-exist in a single cell type. Almost all of the rest of the data consist of a cell type atlas for the two studied intestinal segments in subjects with CD and disease controls, with a particular emphasis on advanced and sophisticated computational and statistical data analysis, including cell type deconvolution and spatial colocalization, cell-cell interaction, and developmental trajectory analysis.

General comments:
1) This appears to be a highly technically sophisticated study from an experienced group that

utilizes advanced methodology.

2) The study's novelty appears to lie almost exclusively with the finding of the "LND" cell type since many prior publications using similar technologies described the transcriptomics of human and rodent colon mucosa in control and IBD models (Zheng; PMID: 37576705; Serigado PMID: 35530046), although this study may boast the largest human sample size that includes both ileum and ascending colon.

3) The phenotype of the CD population is not described in detail; in particular, there is no quantification of disease activity using the CDAI or equivalent validated disease activity scale, which, although most useful in therapeutic trials, could provide a basis for correlations and associations in this study.

4) Also missing is an inflammation control, such as ischemic, infectious, or diverticular colitis, that could support the authors' contention that the observed changes in subjects with CD were "driven by inflammation" (lines 211-212). By that matter, ulcerative colitis samples would add specificity as well.

5) There is only a superficial description of the function of proteins encoded by the three identified transcripts, which is unfortunate in that these proteins all generate products that can damage tissue, thus providing new directions for IBD research:

a) For example, LCN2 is closely linked with ulcerative colitis and upregulated in human colitis and in colitis models. Importantly, the LCN2 product lipocalin2 is implicated in a form of programmed cell death
termed ferroptosis that has been hypothesized to contribute to colitis-associated mucosal damage (Deng PMID: 37000707)

b) The nitric oxide synthases (NOS) produce NO that has diverse biological functions. NOS2, also termed inducible NOS or iNOS, is induced in human IBD (Kimura; PMID: 9149061; Rachmilewitz; PMID: 7541008).
Furthermore, the contribution of other NOS isoforms, in particular eNOS (NOS3) and nNOS (NOS1), to IBD pathogenesis has been studied in several experimental models, where NOS2 activity was predominantly
localized to submucosal inflammatory cells (Vallance; PMID: 15217783, Beck; PMID: 14665440). The finding of epithelial NOS2 gene expression, even in inflamed tissues, is unusual.

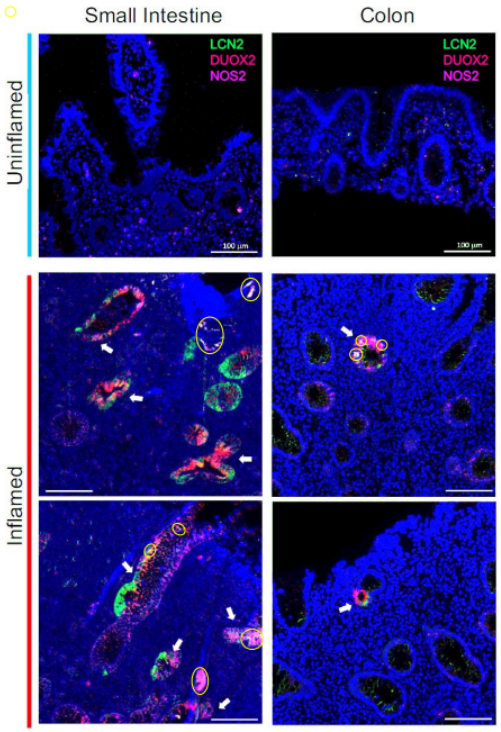
c) The NADPH oxidase DUOX2 generates H_2O_2, which, combined with the pseudohalide SCN^- or halide iodide in the presence of lactoperoxidase generates the potent antimicrobial reactive oxygen species (ROS) HOCl or
HOI. (Sarr; PMID: 29858825). ROS generated by the upregulation of gut NADPH oxidases has been implicated in IBD pathogenesis (Dang; PMID: 31968991).

6) The choice of gene markers for the different cell types (lines 148-156) was derived from prior studies of normal human tissue. Due to the documented "rewiring" (Smillie, 2019) of the inflamed mucosa, the authors should consider validating the markers when needed with immunohistochemistry and related means. Perhaps I missed it, but I was unclear on how the authors distinguished early, intermediate, and mature enterocytes and colonocytes solely

based on transcriptomic analysis. Furthermore, the BEST4/OTOP4 gene used to identify crypt cells does not identify a separate cluster termed "crypt top colonocytes" or "enterocytes" in other studies (cf. Serigado PMID: 35530046)

7) Although the authors provided a helpful comparison between the transcriptomic signatures between ileum and colon for goblet, EEC, and tuft cells, they should also consider comparing microfold (M-cells), that are implicated in bacterial translocation, since they apparently have differing function between the two organs that has not been studied in detail (cf. Serigado PMID: 35530046).

8) The most controversial aspect of the submission is the authors' contention that LNG cells are a novel cluster of absorptive epithelial cells. Data supporting their hypothesis other than the associational analysis of the transcriptomic data appears in Fig 3I, which is HCR-FISH used to demonstrate co-expression of the three RNAs. The figure, as a small panel of the entire figure is difficult to interpret, requiring the printing of a color-enhanced enlargement of the figure, which, when examined, revealed that the co-expression of the three transcripts (assumed to be white fluorescence identified by yellow ellipses) appeared to occur in a small fraction of the labeled cells that appeared to be located in the epithelial cells lining the crypts of the TI and AC in the inflamed samples (see attached image),



that the authors describe as occurring in a "subset, of, but not all epithelial cells" (line 371). Given this sparse co-localization of LNG gene products, all independently confirmed to be upregulated in inflammation of the colon and and likely the TI, the authors, rather than by isolating and functionally characterizing a unique cell fraction, resort to increasingly complex statistical associations and inferences to support their lead hypothesis. As an example, Parikh et al. (2019), in their survey of the human colonic epithelium in controls and in patients with IBD, used transcriptomics for hypothesis generation, confirming their findings with a combination of immunohistochemistry, electron microscopy, FACS, and in vitro cell culture to confirm the function and identity of the novel pH-sensing epithelial BEST4/OTOP2 cell cluster.

9) The cell interaction and development trajectory analyses, though technically sophisticated, cannot be considered definitive until they are confirmed by more conventional analysis of cell-cell interactions and development.

10) The authors should consider including even more of the inferential, speculative, confirmatory, and background data as Appendices, focusing on the primary novelty of their data, the identity of the LNG cells and the comparisons between ileum and AC.

Minor:
1) Please provide a space between the references
2) Line 367 "Inflammed"
3) Line 603 "significiant"
4) Please ensure that all colors used in figures are identified in teh figure itself and in the caption

NCOMMS-23-40308-T
A Specialized Epithelial Cell Type Regulating Mucosal Immunity and Driving Human Crohn's Disease

REVIEWER COMMENTS

Reviewer #1 (expert in IBD, single-cell transcriptomics):

The manuscript by Li et al. employs a combination of single-cell RNA sequencing, spatial transcriptomics and single-molecule fluorescence in situ hybridisation to uncover a population of Crohn's disease (CD)-specific epithelial cells, termed 'LND cells' by the authors, that provide immune recruiting signals and correlate with response to anti-TNF treatment. While there have been a number of prior single-cell publications in the IBD space, as the authors argue, this manuscript uniquely takes an anatomically resolved look at the epithelial lineage during inactive and active disease. The authors provide analyses of the developmental origin, signalling partners, colocalising partners and IBD-susceptibility gene expression of LND cells. I also appreciate the inclusion of cohort details in the main text. Together, this is a convincing story that would be of interest to clinical and academic gastrointestinal fields.

***Response**: We sincerely appreciate the reviewer for recognizing the contributions of our work and offering valuable suggestions to improve the manuscript and strengthen the conclusion significantly.*

Major:
1. Line 567-: The authors should control for the possibility that response to anti-TNF therapy can be predicted more generally by mature epithelial cells by also testing the transcriptional signature of mature enterocytes/colonocytes. If other mature epithelial cells cannot predict the outcome, then this would strengthen the case for late LND cells having a particular utility in anti-TNF response prediction.
***Response**: Thank you for the great suggestion. Following the reviewer's suggestion, we assessed the correlation between each epithelial cell type and response to anti-TNF therapy. Our analysis revealed that none of the epithelial cell types were related to anti-TNF therapy response except for late LND cells (Supplementary Fig. S20). This strengthens the conclusion that late LND cells have particular utility in anti-TNF response prediction. We have added this result in the manuscript (lines 369-370, page 15 and Supplementary Fig. S20).*

Minor:
1. Given the low cell number per sample in the single-cell RNA sequencing data, it would be helpful to include a comment in the text about whether LND cells are present across all CD samples (it is quite hard to tell this from Supplementary Table S2).
***Response**: Thank you for the suggestion. LND cells were present in 73.6% and 87.2% of CD samples in endoscopic TI and AC, respectively. We have added the description in the manuscript (lines 254-255, page 10, and Table S3).*

2. Line 99: “CD often involves the ileum where Paneth cells are abundant, while the colon contains no or few Paneth cells” suggests that Paneth cells are a key cell in CD pathology. While this might be the case, the current manuscript does not focus on Paneth cells so perhaps change this line so that the reader isn’t confused or distracted.

***Response**: Thank you for pointing this out. We have revised this sentence to “the ileum and colon exhibit distinct cellular compositions and have specialized preferences for nutrient absorption, microbe defenses, and endocrine function[1]” (lines 95-97, page 4).*

3. Were IgG+ plasma cells identified in the single-cell RNA seq data of CD samples?

***Response**: Thank you for the question. We observed a minimal presence of IgG+ plasma cells, constituting only 1% (Supplementary Fig. S2), as determined by the criteria outlined in a prior study[2]. In brief, we computed the mean normalized expression of IgA heavy chain genes (IGHA1 and IGHA2) and IgG heavy chain genes (IGHG1, IGHG2, IGHG3, and IGHG4). A predefined cutoff of 6 was employed to differentiate between the two types of plasma cells. We have incorporated these results in Section 1 of the Supplementary Information (Supplementary Fig. S2).*

4. Line 210: It may be interesting to list genes that change as a function of disease activity in the absorptive epithelium.

***Response**: Thank you for this great suggestion. We have described the genes that change as a function of disease activity in mature enterocytes/colonocytes in Section 3 of the Supplementary Information (Supplementary Fig. S5). Other cell-type specific changes as a function of disease activity have been added in Table S4.*

5. Line 237-240: Details of marker genes for immune cell types are repetitive of what was included earlier (~line 153).

***Response**: Thank you for pointing this out. We have removed the redundancies.*

6. Line 323: Citations for the six independent studies and that they are microarray and bulk mRNA data should be included.

***Response**: Thank you for pointing this out. We have included citations for the six independent studies (lines 270-271, 274-277, page 11). We have specified whether they are microarray or bulk RNA-seq data in the Results section (lines 270-271, page 11) and in the Data availability section (page 37).*

7. Figure 3 panel I: Please include the number of repeats for the HCR-FISH staining in the figure legend. The same applies to Figure 4E.

***Response**: Thank you for the comment. We have included the number of repeats in the legends of those figures.*

Reviewer #2 (expert in IBD, gastroenterology):

The manuscript showed important insights into IBD pathogenesis, particularly in CD, a heterogeneous disease whose etiology and pathophysiology are still far from fully understood. In this scenario, the authors made much effort in compiling a broad single-cell expression profiling of CD mucosa, in ascending colon and terminal ileum.
I found the level of data analysis impressive, very well-detailed, and comprehensive of all biological and pathophysiological questions that may be raised by such investigation. My enthusiasm was mainly raised by the elegant analysis of localization, characterization, and immune cell interactions.

Response*: We sincerely thank the reviewer for recognizing our efforts and providing valuable suggestions and comments, which have significantly contributed to the improvement of the manuscript.*

The authors interestingly identify a specific epithelial subpopulation named LND based on the expression of specific genes (LCN2, NOS2, and DUOX2). This is very interesting since these cells correlate with inflammation and are specifically associated with CD.

However, despite the enormous amount of computational analysis and data produced, I believe that the work is still too descriptive.

One of the major concerns is the difficult reading for people not familiar with this kind of research and, given the broad readership of the journal, I suggest that the authors rewrite it more fluidly, so that the main messages are caught up easily and quickly.
Response*: Thank you for the valuable suggestion. We have rewritten and reorganized the manuscript to enhance its flow and clarity. Four minor sections, which are not closely related to LND cells, have been moved to the Supplementary Information. Additionally, we have merged the section titled "Colocalization between LND and immune cells" into "LND cells actively interact with immune cells". This adjustment aims to emphasize the key findings on LND-mediated immunoregulation and explain the potential mechanism of LND involvement in CD pathogenesis.*

Also, biological validations are required, at least for the LND population. Indeed, it would be helpful for the scientific audience to see whether these populations are acting in a specific way in the tissue, explaining the pathogenesis, at least in part, ultimately strengthening the biological insights provided.
Response*: Thank you for the insightful comment. Following the reviewer's suggestion, we conducted further biological validations of the LND subpopulation using immunohistochemistry and multiplex immunofluorescence (lines 321-325, page 13 and Supplementary Figs. S17 and S18). These experiments confirmed increased expression and colocalization of the three proteins (*LCN2, NOS2, and DUOX2*) in active CD with high LND. Moreover, we investigated the impact of LND cells on immune cell migration using transwell migration assays. Our results demonstrated that conditioned media derived from active CD with high LND significantly increased monocyte migration compared to those from active CD with low LND (p=0.001) (lines 397-410, pages 16-17, Figs. 5F and 5G). These findings support the role of LND in recruiting monocytes and provide insights into the potential mechanism of LND involvement in CD pathogenesis.*

Another concern is about the clinical relevance. May the author correlate these subpopulations with clinical activity or characteristics of the patients, further expanding the scientific value of their findings? Are there other correlations with these cells?

***Response**: Thank you for the insightful comment. In response to the reviewer's comment, we investigated the correlation between each cell type and the Clinical Disease Activity Index (CDAI), a validated non-invasive disease activity assessment. Unfortunately, we did not observe any significant correlations between them. This outcome is likely due to inherent limitations in the CDAI, including subjective questions based on patient reported symptoms/overall well-being and variability when incorporating ideal versus actual body weight. We have included these findings in the Discussion (lines 507-513, page 21, and Supplementary Fig. S24), and the CDAI scores have been provided in Table S1. Additionally, no correlation of the LND subpopulation with medication exposure was observed (lines 260-261, page 11, and Supplementary Fig. S11).*

Moreover, we assessed cytokine levels in the serum from 37 CD patients. Our findings reveal that individuals with high LND (n=18) exhibit significantly elevated levels of CXCL1 (p=0.04), CXCL9 (p=2.9e-5), and IL6 (p=0.03) compared to those with low LND (n=19). This observation suggests that LND status may predict differences in systemic inflammation levels. These results have been integrated into the Discussion section (lines 525-529, page 22 and Supplementary Fig. S26).

Moreover, they suggest this subpopulation as a potential biomarker. Is there a straightforward way to detect these cells and use them in clinics? A proof of concept may be sufficient at this stage.

***Response**: This is a great question. We attempted to use the average expression level of LCN2, NOS2, and DUOX2 as a surrogate for both the LND population and disease activity. Our analysis revealed a very significant correlation between the average expression of these three genes and LND proportions ($r>0.8$ & $p<2.2e-16$), as well as disease activity ($p<0.001$), in both TI and AC (Supplementary Fig. S25). This finding implies that the average expression level of these three genes may serve as a reliable surrogate for detecting LND cells and as a supplementary indicator for disease activity beyond histology. We have included this result in the Discussion (lines 517-525, pages 21-22 and Supplementary Fig. S25).*

Overall, although the work and the analysis provided are very well addressed, discussed, and valuable, I believe that the manuscript needs to be shortened, simplified, and carry more mechanistic proofs to endow it with importance in the field of IBD. As it is, it remains a descriptive work, unfortunately far from the journal requirements. Indeed, considering that the main output of the manuscript is descriptive, I would refer to the work as an atlas, rather than a scientific contribution expanding the mechanistic knowledge of the disease pathogenesis.

***Response**: Thank you for the comment. We have made every effort to condense and streamline the manuscript, while also enhancing it with additional biological characterization and functional studies on the LND subpopulation. Hopefully, our work not only serves as an atlas, but also broadens the mechanistic understanding of disease pathogenesis, particularly regarding the involvement of LND cells in CD susceptibility through immune regulation.*

Reviewer #3 (expert in IBD, gastroenterology, intestinal epithelium):

Description:
The authors have submitted a paper that describes the results of transcriptomic analysis of the terminal ileum and ascending colon of 170 tissue samples obtained 65 patients with Crohn's disease (CD) and 18 subjects with no history of inflammatory bowel disease. The most interesting aspect of the work is the discovery of a purported novel cell type, termed "LNG" for lipocalin, NOS2, and DUOX2, which the authors state co-exist in a single cell type. Almost all of the rest of the data consist of a cell type atlas for the two studied intestinal segments in subjects with CD and disease controls, with a particular emphasis on advanced and sophisticated computational and statistical data analysis, including cell type deconvolution and spatial colocalization, cell-cell interaction, and developmental trajectory analysis.

Response*: We express our sincere gratitude to the reviewer for acknowledging the strengths of our work and providing great suggestions and comments to improve the manuscript significantly.*

General comments:

1) This appears to be a highly technically sophisticated study from an experienced group that utilizes advanced methodology.
Response*: Thank you for acknowledging the strengths of our work.*

2) The study's novelty appears to lie almost exclusively with the finding of the "LND" cell type since many prior publications using similar technologies described the transcriptomics of human and rodent colon mucosa in control and IBD models (Zheng; PMID: 37576705; Serigado PMID: 35530046), although this study may boast the largest human sample size that includes both ileum and ascending colon.
Response*: Thank you for acknowledging the novelty of our work. Indeed, the groundbreaking aspect of this study is the discovery of the LND cell type, revealing its potential involvement in CD susceptibility through immunoregulation.*

3) The phenotype of the CD population is not described in detail; in particular, there is no quantification of disease activity using the CDAI or equivalent validated disease activity scale, which, although most useful in therapeutic trials, could provide a basis for correlations and associations in this study.
Response*: Thank you for this great suggestion. In response to the reviewer's suggestion, we investigated the correlation between each cell type and the Clinical Disease Activity Index (CDAI). Unfortunately, we did not observe any significant correlations between proportions of any cell types and CDAI scores. This outcome is likely attributable to inherent limitations in the CDAI including subjective questions based on patient reported symptoms/overall well-being and variability when incorporating ideal versus actual body weight. We have incorporated the findings in the Discussion (lines 507-513, page 21, and Supplementary Fig. S24). We have provided the CDAI scores in Table S1.*

4) Also missing is an inflammation control, such as ischemic, infectious, or diverticular colitis, that could support the authors' contention that the observed changes in subjects with CD were "driven by inflammation" (lines 211-212). By that matter, ulcerative colitis samples would add specificity as well.

***Response**: Thank you for pointing this out. Unfortunately, including ulcerative colitis patients in the study is not feasible as we have reached the Gut Cell Atlas grant's end date and concluded the patient data collection phase. We agree that it is inappropriate to assert "driven by inflammation" without an inflammation control. Therefore, we have revised the sentence to state "driven by disease activity" (line 173-174, page 7).*

5) There is only a superficial description of the function of proteins encoded by the three identified transcripts, which is unfortunate in that these proteins all generate products that can damage tissue, thus providing new directions for IBD research:

***Response**: We highly appreciate the reviewer for providing numerous literature references and biological findings related to these three genes. We have integrated these insights into the manuscript, striving to offer a more comprehensive description of the functions attributed to these genes (page 12).*

a) For example, LCN2 is closely linked with ulcerative colitis and upregulated in human colitis and in colitis models. Importantly, the LCN2 product lipocalin2 is implicated in a form of programmed cell death termed ferroptosis that has been hypothesized to contribute to colitis-associated mucosal damage (Deng PMID: 37000707)

***Response**: The reference and the description have been incorporated into the manuscript.*

b) The nitric oxide synthases (NOS) produce NO that has diverse biological functions. NOS2, also termed inducible NOS or iNOS, is induced in human IBD (Kimura; PMID: 9149061; Rachmilewitz; PMID: 7541008).

Furthermore, the contribution of other NOS isoforms, in particular eNOS (NOS3) and nNOS (NOS1), to IBD pathogenesis has been studied in several experimental models, where NOS2 activity was predominantly localized to submucosal inflammatory cells (Vallance; PMID: 15217783, Beck; PMID: 14665440). The finding of epithelial NOS2 gene expression, even in inflamed tissues, is unusual.

***Response**: The reference and the description have been incorporated into the manuscript.*
*It is very likely that NOS2 expression patterns differ between human and mice. In human tissues, this work involving scRNAseq, RNA-FISH, IHC and MxIF, along with previous studies[3,4](Fig. 3 and Fig. 6, respectively) demonstrated that NOS2 is induced in the epithelium of CD patients. Consistently, NOS2 has been reported to be produced in response to microbes and other activating stimuli in human epithelial cells in multiple studies[5-8]. In the two mouse studies mentioned by the reviewer[9,10], although NOS2 was predominantly localized to inflammatory cells in lamina propria and submucosa, it was also reported to be present in epithelial cells. One study stated that "iNOS was rare in the colon of normal wild-type mice (Fig. 2A). However, expression was evident on **epithelial cells** (arrowhead) and lamina propria mononuclear cells (arrow) within the same mouse strain during colitis (Fig. 2B)"[9]. The other study claimed that "following DSS exposure there also appeared to be a small increase in **intestinal epithelial cell** iNOS expression (Fig. 2d-2d)[10]". Moreover, our group has shown that murine colonic epithelial cells express increased iNOS (NOS2) in response to Citrobacter rodentium infection both in vivo[11] and in vitro[12],*

and that isolated colonic epithelial cells from dextran sulfate sodium-treated mice express increased iNOS[13].

c) The NADPH oxidase DUOX2 generates H2O2, which, combined with the pseudohalide SCN- or halide iodide in the presence of lactoperoxidase generates the potent antimicrobial reactive oxygen species (ROS) HOCl or HOI. (Sarr; PMID: 29858825). ROS generated by the upregulation of gut NADPH oxidases has been implicated in IBD pathogenesis (Dang; PMID: 31968991).
***Response**: The reference and the description have been incorporated into the manuscript.*

6) The choice of gene markers for the different cell types (lines 148-156) was derived from prior studies of normal human tissue. Due to the documented "rewiring" (Smillie, 2019) of the inflamed mucosa, the authors should consider validating the markers when needed with immunohistochemistry and related means. Perhaps I missed it, but I was unclear on how the authors distinguished early, intermediate, and mature enterocytes and colonocytes solely based on transcriptomic analysis. Furthermore, the BEST4/OTOP4 gene used to identify crypt cells does not identify a separate cluster termed "crypt top colonocytes" or "enterocytes" in other studies (cf. Serigado PMID: 35530046)
***Response**: Thank you for the great suggestion, and we apologize for not explaining clearly.*

Following the reviewer's suggestion, we further validated the LND markers using immunohistochemistry and multiplex immunofluorescence (lines 321-325, page 13 and Supplementary Figs. S17 and S18), which demonstrated increased expression and colocalization of the three proteins in active CD with high LND.

We distinguished early enterocytes/colonocytes, intermediate enterocytes/colonocytes, and mature enterocytes/colonocytes based on three perspectives: 1) markers identified in previous studies[14,15]; 2) cell differentiation state; 3) trajectory analysis. We have added a detailed description for distinguishing these cell types in the Methods section (lines 754-767, pages 31-32 and Supplementary Fig. S10) and below.

***Markers identified in previous studies:** Burclaff et al.[14] defined early enterocytes/colonocytes, intermediate enterocytes/colonocytes, and mature enterocytes/colonocytes in TI and AC, respectively. They also identified markers distinguishing these cell types. We employed these markers to delineate early/intermediate/mature enterocytes/colonocytes in our study, such as high expression of FREM1, PCCA, and DMBT1 in early enterocytes, RBP2 in intermediate enterocytes, and APOA1, APOC3, APOA4, and GUCA2B in mature enterocytes. Similarly, early colonocytes were defined by high expression of B3GNT7, ABR, and ADH1C, intermediate colonocytes by ATP5G1 and PCNP, and mature colonocytes by AQP8, GUCA2A, CA4, and CEACAM1 (Supplementary Fig. S10A). Furthermore, Parikh et al.[15] introduced a crypt-axis score, utilizing 15 genes to predict the location of cells within the colonic crypt. In our study, the classification of early/intermediate/mature colonocytes aligns with our in-silico maturity/crypt gradient predictions (Supplementary Fig. S10B). Here, mature colonocytes are positioned at the crypt-top with the highest scores, while early clonotypes are situated at the crypt-bottom with the lowest scores. It is noteworthy that the crypt-axis score is specifically defined for the colon; therefore, we did not conduct this analysis in the TI.*

***Cell differentiation state and trajectory analysis:** Our classification of early/intermediate/mature enterocytes/colonocytes aligns with the predicted cell differentiation state (Fig. 4D in the manuscript) and trajectory analysis (Fig. 4A in the manuscript). As anticipated,*

stem cells exhibit the highest score, indicating undifferentiated status, followed by early and intermediate enterocytes, while mature enterocytes register the lowest score, signifying full differentiation (Fig. 4D in the manuscript). In the trajectory analysis, a pronounced directional flow is observed starting from stem cells, progressing through early enterocytes, intermediate enterocytes, and culminating in mature enterocytes (Fig. 4A in the manuscript).

We apologize for any confusion. The BEST4/OTOP2 genes are specifically employed to identify a distinct cell type known as "BEST4/OTOP2" cells, rather than two separate cell types "crypt top" or "enterocytes." These BEST4/OTOP2 cells are specifically located at the crypt top, as indicated by Parikh et al.[15]. To eliminate potential confusion, we now refer to them simply as "BEST4/OTOP2 cells" instead of "BEST4/OTOP2 crypt-top cells.

7) Although the authors provided a helpful comparison between the transcriptomic signatures between ileum and colon for goblet, EEC, and tuft cells, they should also consider comparing microfold (M-cells), that are implicated in bacterial translocation, since they apparently have differing function between the two organs that has not been studied in detail (cf. Serigado PMID: 35530046).

***Response**: Thank you for the great suggestion. Previously, microfold cells were not detected in our single-cell data. In response to the reviewer's question, we reanalyzed each cluster using an alternative strategy. Regrouping BEST4/OTOP2 cells using an innovative clustering approach based on an adaptive k-nearest neighbor graph[16], we identified a small cluster only in the AC characterized by high expression of SPIB, SOX8, TNFRSF11B, CCL23, and CCL20 (33 out of 1172 cells, mostly from one patient) (Supplementary Fig. S1). These genes are recognized markers of colonic microfold-like cells[2,17]. Notably, SPIB is also highly expressed in BEST4/OTOP2 and Tuft cells, aligning with previous findings[2,17]. Since M-like cells were not identified in the TI and only detected in the AC of a few patients, we were unable to compare their transcriptional differences between TI and AC. This result has been added into Section 1 of the Supplementary Information (Supplementary Fig. S1).*

8) The most controversial aspect of the submission is the authors' contention that LNG cells are a novel cluster of absorptive epithelial cells. Data supporting their hypothesis other than the associational analysis of the transcriptomic data appears in Fig 3I, which is HCR-FISH used to demonstrate co-expression of the three RNAs. The figure, as a small panel of the entire figure is difficult to interpret, requiring the printing of a color-enhanced enlargement of the figure, which, when examined, revealed that the co-expression of the three transcripts (assumed to be white fluorescence identified by yellow ellipses) appeared to occur in a small fraction of the labeled cells that appeared to be located in the epithelial cells lining the crypts of the TI and AC in the inflamed samples (see attached image), that the authors describe as occurring in a "subset, of, but not all epithelial cells" (line 371). Given this sparse co-localization of LNG gene products, all independently confirmed to be upregulated in inflammation of the colon and and likely the TI, the authors, rather than by isolating and functionally characterizing a unique cell fraction, resort to increasingly complex statistical associations and inferences to support their lead hypothesis. As an example, Parikh et al. (2019), in their survey of the human colonic epithelium in controls and in patients with IBD, used transcriptomics for hypothesis generation, confirming their findings with a combination of immunohistochemistry, electron microscopy, FACS, and in vitro cell culture to confirm the function and identity of the novel pH-sensing epithelial BEST4/OTOP2 cell cluster.

***Response**: We highly appreciate the reviewer for this insightful comment. Following the*

reviewer's suggestion, we confirmed the identity and function of LND cells through a series of experiments including repeated RNA-FISH, immunohistochemistry, multiplex immunofluorescence, and in-vitro transwell migration assays.

First, we replicated RNA-FISH experiments in multiple TI and AC tissues, depending on their availability (Fig. S16 A-C). Addtionally, We found that transcripts appear as single molecule puncta rather than diffuse like proteins. This means that colocalization within a cell is not limited to overlapping signals but occurs within the borders of a segmented cell (Fig. S16D).

Next, we characterized LND cells at the protein level using immunohistochemistry and multiplex immunofluorescence, demonstrating increased expression and co-colocalization of the three proteins in active CD with high LND (lines 321-325, page 13 and Supplementary Figs. S17 and S18). Additionally,

Finally, we investigated the impact of LND cells on immune cell migration using transwell migration assays. Importantly, we observed a significantly higher number of migrated monocytes in conditioned media from high LND compared to low LND cells (p=0.001) (lines 397-410, pages 16-17, Figs. 5F and 5G), supporting the role of LND cells in recruiting immune cells.

9) The cell interaction and development trajectory analyses, though technically sophisticated, cannot be considered definitive until they are confirmed by more conventional analysis of cell-cell interactions and development.
***Response**: Thank you for pointing this out. We have revised those sections to convey a less definitive tone and relocated sections that are not closely related to the Supplementary Information (Sections 1-4 in Supplementary information).*

10) The authors should consider including even more of the inferential, speculative, confirmatory, and background data as Appendices, focusing on the primary novelty of their data, the identity of the LNG cells and the comparisons between ileum and AC.
***Response**: Thank you for the great suggestion. We have reorganized and simplified the manuscript, focusing on the identity and potential function of the LND cells. Four inferential subsections have been relocated to the Supplementary Information (Sections 1-4 in Supplementary information).*

Minor:
1) Please provide a space between the references
***Response**: Thank you for pointing this out. We have added a space between the references.*

2) Line 367 "Inflammed"
***Response**: corrected.*

3) Line 603 "significiant"
***Response**: corrected.*

4) Please ensure that all colors used in figures are identified in the figure itself and in the caption
***Response**: Thank you for pointing this out. We have added labels for all colors in the figure captions.*

References

1 Wang, Y. *et al.* Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. *J Exp Med* **217**, doi:10.1084/jem.20191130 (2020).

2 Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730 e722, doi:10.1016/j.cell.2019.06.029 (2019).

3 Leonard, N., Bishop, A. E., Polak, J. M. & Talbot, I. C. Expression of nitric oxide synthase in inflammatory bowel disease is not affected by corticosteroid treatment. *J Clin Pathol* **51**, 750-753, doi:10.1136/jcp.51.10.750 (1998).

4 Singer, II *et al.* Expression of inducible nitric oxide synthase and nitrotyrosine in colonic epithelium in inflammatory bowel disease. *Gastroenterology* **111**, 871-885, doi:10.1016/s0016-5085(96)70055-0 (1996).

5 Rumbo, M., Courjault-Gautier, F., Sierro, F., Sirard, J. C. & Felley-Bosco, E. Polarized distribution of inducible nitric oxide synthase regulates activity in intestinal epithelial cells. *FEBS J* **272**, 444-453, doi:10.1111/j.1742-4658.2004.04484.x (2005).

6 Eckmann, L. *et al.* Nitric oxide production by human intestinal epithelial cells and competition for arginine as potential determinants of host defense against the lumen-dwelling pathogen Giardia lamblia. *J Immunol* **164**, 1478-1487, doi:10.4049/jimmunol.164.3.1478 (2000).

7 Cross, R. K. & Wilson, K. T. Nitric oxide in inflammatory bowel disease. *Inflamm Bowel Dis* **9**, 179-189, doi:10.1097/00054725-200305000-00006 (2003).

8 Witthoft, T., Eckmann, L., Kim, J. M. & Kagnoff, M. F. Enteroinvasive bacteria directly activate expression of iNOS and NO production in human colon epithelial cells. *Am J Physiol* **275**, G564-571, doi:10.1152/ajpgi.1998.275.3.G564 (1998).

9 Vallance, B. A. *et al.* Relative contributions of NOS isoforms during experimental colitis: endothelial-derived NOS maintains mucosal integrity. *Am J Physiol Gastrointest Liver Physiol* **287**, G865-874, doi:10.1152/ajpgi.00187.2004 (2004).

10 Beck, P. L. *et al.* Paradoxical roles of different nitric oxide synthase isoforms in colonic injury. *Am J Physiol Gastrointest Liver Physiol* **286**, G137-147, doi:10.1152/ajpgi.00309.2003 (2004).

11 Gobert, A. P. *et al.* Protective role of arginase in a mouse model of colitis. *J Immunol* **173**, 2109-2117, doi:10.4049/jimmunol.173.3.2109 (2004).

12 Singh, K. *et al.* L-arginine uptake by cationic amino acid transporter 2 is essential for colonic epithelial cell restitution. *Am J Physiol Gastrointest Liver Physiol* **302**, G1061-1073, doi:10.1152/ajpgi.00544.2011 (2012).

13 Singh, K. *et al.* The apolipoprotein E-mimetic peptide COG112 inhibits NF-kappaB signaling, proinflammatory cytokine expression, and disease activity in murine models of colitis. *J Biol Chem* **286**, 3839-3850, doi:10.1074/jbc.M110.176719 (2011).

14 Burclaff, J. *et al.* A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. *Cell Mol Gastroenterol Hepatol* **13**, 1554-1589, doi:10.1016/j.jcmgh.2022.02.007 (2022).

15 Parikh, K. *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49-55, doi:10.1038/s41586-019-0992-y (2019).

16 Li, J., Shyr, Y. & Liu, Q. Single-cell and Spatial Transcriptomics Clustering with an Optimized Adaptive K-Nearest Neighbor Graph. *bioRxiv*, doi:10.1101/2023.10.13.562261 (2023).

17 Serigado, J. M. *et al.* Ulcerative Colitis: Novel Epithelial Insights Provided by Single Cell RNA Sequencing. *Front Med (Lausanne)* **9**, 868508, doi:10.3389/fmed.2022.868508 (2022).

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed my initial comments and their manuscript has been improved through the review process. The work represents a valuable resource and uncovers a previously unidentified LCN2+ NOS2+ DUOX2+ (LND) epithelial cell type with a pathogenic role in Crohn's disease and potential in TNF-treatment response prediction. I have a few further minor suggestions:

1. Include figure legend titles throughout the manuscript as already done for Figure 6.
2. Include version numbers for software packages in the methods section as documented in the Reporting Summary.
2. Include Code Availability regarding codes provided on GitHub.

Reviewer #1 (Remarks on code availability):

Codes for the bulk RNAseq analysis, ligand-receptor expression (cellchat) and cell proportions from scRNAseq data are included on GitHub. Although the codes are not extensively annotated, examples are provided for each script and they should be reproducible. Codes for trajectory analysis, spatial deconvolution and earlier preprocessing of single-cell RNA sequencing data are not provided.

Reviewer #2 (Remarks to the Author):

The authors replied to all comments.

Reviewer #3 (Remarks to the Author):

The authors have returned a revised manuscript in which a unique epithelial cell type was identified in the GI mucosa of patients with Crohn's disease.

In general, the authors have provided a detailed, constructive, positive, and helpful responses to the reviewers' comments with overall improvement of the comprehensibility and conciseness of the submission. Of particular note is the addition of functional studies such as those depicted in Figs. 5F-5G, additional methods and analyses, and more detailed characterization of the developmental stage of the epithelial cells in response to the reviewers' comments. As such the authors have enhanced the import of their data within the constraints of their original experimental design.

I just have a few remaining comments:

1) The title is at best aspirational in that it implies that the novel cell type "regulates" mucosal immunity and "drives human Crohn's disease". Recommend toning down the title to match conclusions that can be derived from the data shown, which are mostly descriptive and inferential.

2) Lines 293-6: Suggest stating that thiocyanate (SCN-) is the predominant pseudohalide converted to the ROS hypothiocyanite (OSCN-) by the action of lactoperoxidase in the presence of DUOX-generated H2O2. The antimicrobial effect of this system is suggested by alteration of the gut flora in selective DUOX knockout mice: (ref. 65; see attached figure).

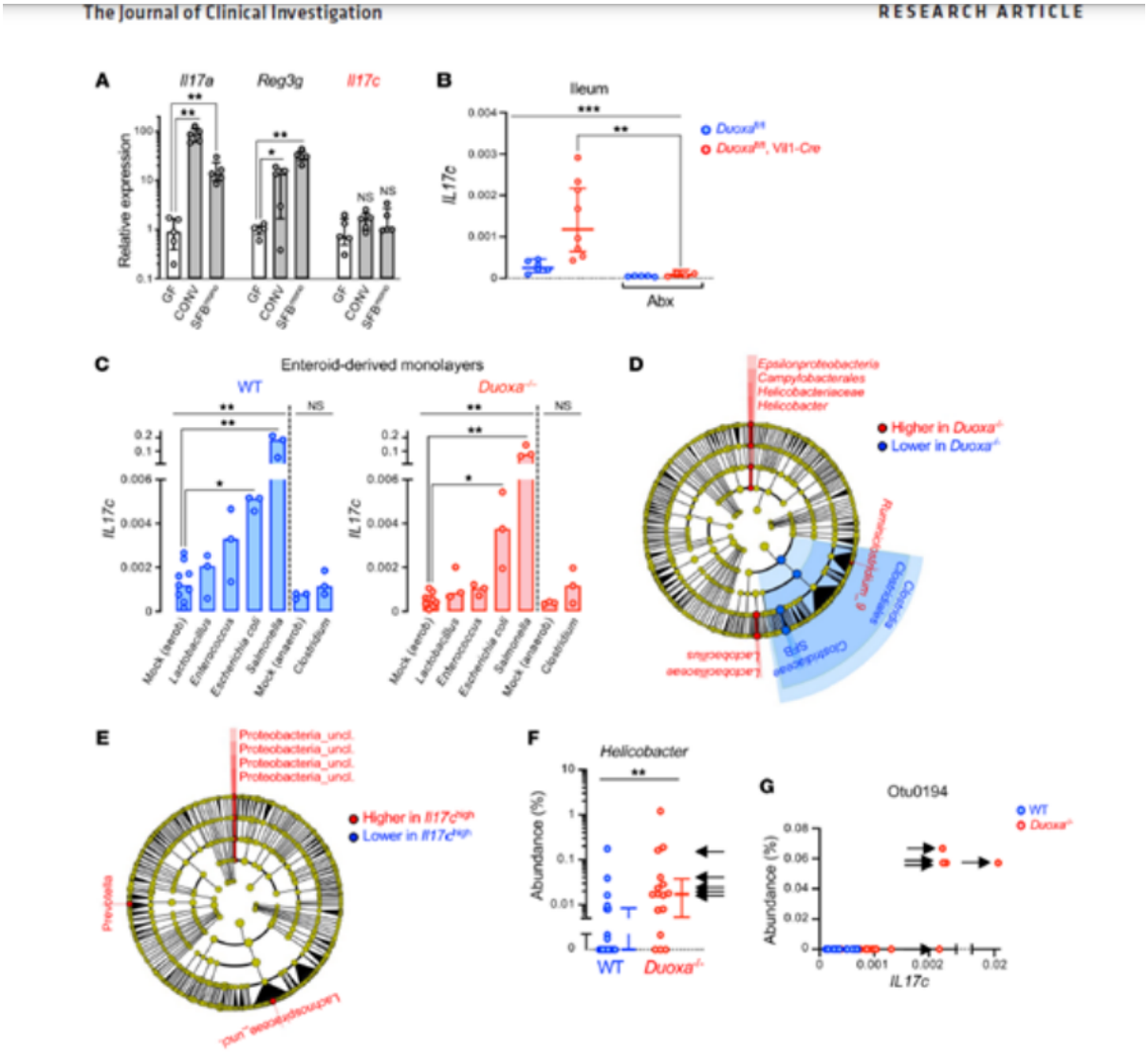


Figure 5. High *Il17c* expression in the intestinal epithelium of DUOX2-deficient mice is linked to the expansion of gram-negative pathobionts. (A)

NCOMMS-23-40308-A
A Specialized Epithelial Cell Type Regulating Mucosal Immunity and Driving Human Crohn's Disease

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed my initial comments and their manuscript has been improved through the review process. The work represents a valuable resource and uncovers a previously unidentified LCN2+ NOS2+ DUOX2+ (LND) epithelial cell type with a pathogenic role in Crohn's disease and potential in TNF-treatment response prediction. I have a few further minor suggestions:

***Response**: We thank the reviewer for constructive suggestions to further improve the work.*

1. Include figure legend titles throughout the manuscript as already done for Figure 6.

***Response**: We have included a brief title for each figure legend.*

2. Include version numbers for software packages in the methods section as documented in the Reporting Summary.

***Response**: We have included version numbers for software packages in the methods section.*

3. Include Code Availability regarding codes provided on GitHub.

***Response**: We have included Code Availability section.*

Reviewer #1 (Remarks on code availability):

Codes for the bulk RNAseq analysis, ligand-receptor expression (cellchat) and cell proportions from scRNAseq data are included on GitHub. Although the codes are not extensively annotated, examples are provided for each script and they should be reproducible. Codes for trajectory analysis, spatial deconvolution and earlier preprocessing of single-cell RNA sequencing data are not provided.

***Response**: We have added annotations for codes and added codes for trajectory analysis, spatial deconvolution and preprocessing of single-cell RNA sequencing and spatial transcriptomics data.*

Reviewer #2 (Remarks to the Author):

The authors replied to all comments.

Reviewer #3 (Remarks to the Author):

The authors have returned a revised manuscript in which a unique epithelial cell type was identified in the GI mucosa of patients with Crohn's disease.

In general, the authors have provided a detailed, constructive, positive, and helpful responses to the reviewers' comments with overall improvement of the comprehensibility and conciseness of the submission. Of particular note is the addition of functional studies such as those depicted in Figs. 5F-5G, additional methods and analyses, and more detailed characterization of the developmental stage of the epithelial cells in response to the reviewers' comments. As such the authors have enhanced the import of their data within the constraints of their original experimental design.

***Response**: We appreciate the reviewer for insightful comments that further improved the manuscript.*

I just have a few remaining comments:

1) The title is at best aspirational in that it implies that the novel cell type "regulates" mucosal immunity and "drives human Crohn's disease". Recommend toning down the title to match conclusions that can be derived from the data shown, which are mostly descriptive and inferential.

***Response**: Following the reviewer's recommendation, we have toned down the title to match our findings. The current title is "Identification and multimodal characterization of a specialized epithelial cell type associated with Crohn's Disease".*

2) Lines 293-6: Suggest stating that thiocyanate (SCN-) is the predominant pseudohalide converted to the ROS hypothiocyanite (OSCN-) by the action of lactoperoxidase in the presence of DUOX-generated H2O2. The antimicrobial effect of this system is suggested by alteration of the gut flora in selective DUOX knockout mice: (ref. 65; see attached figure).

***Response**: We thank the reviewer for the explanation. We have revised the sentence to state that Duox2-deficient mice had altered microbiota composition (page 20).*