

## Relative Abundance of SARS-CoV-2 Entry Genes in the Enterocytes of the Lower Gastrointestinal Tract

Jaewon J. Lee,<sup>1,2,3</sup> Scott Kopetz,<sup>4</sup> Eduardo Vilar,<sup>4,5</sup> John Paul Shen,<sup>4</sup> Ken Chen,<sup>6</sup> Anirban Maitra<sup>1,2</sup>

<sup>1</sup>Department of Translational Molecular Pathology

<sup>2</sup>Sheikh Ahmed Center for Pancreatic Cancer Research

<sup>3</sup>Department of Surgical Oncology

<sup>4</sup>Department of Gastrointestinal Medical Oncology

<sup>5</sup>Department of Clinical Cancer Prevention

<sup>6</sup>Department of Bioinformatics and Computational Biology

The University of Texas MD Anderson Cancer Center

### Correspondence:

Anirban Maitra, M.B.B.S.

Sheikh Ahmed Center for Pancreatic Cancer Research

Department of Translational Molecular Pathology

The University of Texas MD Anderson Cancer Center

6565 MD Anderson Blvd

Houston, TX 77030

Email: [amaitra@mdanderson.org](mailto:amaitra@mdanderson.org)

**Conflict of Interest:** A.M. receives royalties for a pancreatic cancer biomarker test from Cosmos Wisdom Biotechnology, and this financial relationship is managed and monitored by the UTMDACC Conflict of Interest Committee. A.M. is also listed as an inventor on a patent that has been licensed by Johns Hopkins University to Thrive Earlier Detection. E.V. has a consulting and advisory role with Janssen research and Development. The remaining authors disclose no conflicts.

COVID-19, the disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has rapidly spread throughout the world and was declared a pandemic by the World Health Organization, thus leading to a rapid surge in the efforts to understand the mechanisms of transmission, methods of prevention, and potential therapies. While COVID-19 frequently manifests as a respiratory infection,<sup>1</sup> there is evidence for infection of the gastrointestinal (GI) tract<sup>1-4</sup> with documented viral RNA shedding in the stool of infected patients.<sup>2,4</sup> In this study, we aimed to investigate the expression of *ACE2* and *TMPRSS2*, which are required for SARS-CoV-2 entry into mammalian cells,<sup>5</sup> from single-cell RNA sequencing (scRNA-seq) datasets of five different parts of the GI tract: esophagus, stomach, pancreas, small intestine, and colon/rectum.

## Methods

We utilized previously published scRNA-seq data, which are summarized in **Supplementary Table 1**. For in-house normal colon samples, a total of 7 patients were recruited at the University of Texas MD Anderson Cancer Center through written informed consent following Institutional Review Board approval. Single cell dissociation, library generation, and sequencing were performed following standard protocols described in **Supplementary Materials**.

## Results

Analysis of scRNA-seq was performed separately for each GI segment and available cell identities from the original studies were assigned to the new clusters after confirming the expression of relevant marker genes (**Figure 1**). Overall, there were 82,422 cells from the esophagus, 18,405 cells from the stomach, 13,407 cells from the pancreas,

9,858 cells from the small intestine, and 20,161 cells from the colon and rectum (**Supplementary Table 1**). The proportions of cells expressing *ACE2* were approximately 16-fold lower in the upper GI tract (esophagus, stomach; 1% or 1,004/100,827 cells) than in the lower GI tract (ileum, colon, rectum; 16% or 4,777/30,019 cells), and overall, higher proportions of cells expressed *TMPRSS2* than *ACE2* throughout the GI tract (**Supplementary Figures 1A-E**).

Of note, higher proportions of esophageal columnar cells – which were mostly from Barrett’s esophagus samples – co-expressed *ACE2* and *TMPRSS2*, compared to the native squamous epithelium (**Figure 1A**). Pancreatic ductal and acinar cells co-expressed *ACE2* and *TMPRSS2* but endocrine cells did not show detectable co-expression (**Figure 1C**). Parenthetically, the expression of *TMPRSS2* in acinar cells underscores the rationale for using a *TMPRSS2* inhibitor (camostat mesylate) in acute pancreatitis. In fact, this agent is now undergoing early phase clinical trials in COVID-19 patients.<sup>5</sup>

Taken together, within the GI tract, the co-expression of *ACE2* and *TMPRSS2* transcripts was highest in the small intestine and colon. Approximately 20% of enterocytes from the small intestine and 5% of colonocytes (enterocytes from colon/rectum) co-expressed *ACE2* and *TMPRSS2* (**Figure 1D-E**). Analysis of 308 genes that were positively correlated (Pearson’s  $r > 0.1$ ) with *ACE2* in the small intestine revealed the enrichment of functional Gene Ontology (GO) terms such as metabolic and transport pathways in the small intestine (**Supplementary Figure 1F**), thus confirming the inherent digestive and absorptive functions of enterocytes. Interestingly, 251 genes positively correlated with *ACE2* in the colon and rectum showed enrichment of not only secretion-associated

pathways but also immune-related processes (**Supplementary Figure 1G**), which may be upregulated in functional colonocytes due to the microbiome.

## Discussion

To summarize, our results confirm the co-expression of SARS-CoV-2 entry genes *ACE2* and *TMPRSS2* in a subset of epithelial cells in the GI tract, especially functional enterocytes from the lower GI tract (small intestine and colon), and provide potential insights into why entero-colitic symptoms may arise in acute COVID-19 infection. These findings parallel the findings from recent studies that revealed the co-expression of *ACE2* and *TMPRSS2* in olfactory epithelium<sup>6</sup> and intrahepatic cholangiocytes,<sup>7</sup> which may explain atypical COVID-19 symptoms and laboratory findings such as anosmia/ageusia and transaminitis. Our results further support the feasibility of SARS-CoV-2 entry into the GI tract with implications for GI infection and transmission.

## Acknowledgements

We thank Charles Bowen for performing single cell dissociations from colonic samples, Kangyu Lin and Jinzhuang Dou for assistance in data acquisition and transfer, and the Single Cell Core at MD Anderson Cancer Center.

## Funding

This work was supported by the MD Anderson Pancreatic Cancer Moon Shot Program and the Khalifa Bin Zayed Foundation to A.M.; the Human Breast Cell Atlas Seed Network Grant HCA3-0000000147 from the Chan Zuckerberg Initiative to K.C.; the NIH K22 CA234406-01 and the Cancer Prevention and Research Institute of Texas (CPRIT)

Scholar in Cancer Research to J.P.S.; a gift from the Feinberg Family and MD Anderson Cancer Center Pre-Cancer Atlas Project to E.V.; the MD Anderson Colorectal Cancer Moon Shot Program to S.K. and E.V.; the NIH T32 CA009599 to J.J.L.; the NIH/NCI P30 CA016672 to the MD Anderson Cancer Center Core Support Grant; and the CPRIT SINGLE CORE Facilities Grant RP180684.

## References

1. Guan WJ, Ni ZY, Hu Y, et al. N Engl J Med 2020 [E-pub ahead of print].
2. Holshue ML, et al. N Engl J Med 2020;382:929-936.
3. Xiao F, Tang M, Zheng X, et al. Gastroenterology 2020 [E-pub ahead of print].
4. Wolfel R, Corman VM, Guggemos W, et al. Nature 2020 [E-pub ahead of print].
5. Hoffmann M, Kleine-Weber H, et al. Cell 2020 [E-pub ahead of print].
6. Brann DH, Tsukahara T, Weinreb C, et al. bioRxiv 2020.
7. Wen Seow JJ, et al. bioRxiv 2020.

## Figure Legends

**Figure 1.** Uniform manifold approximation and projection plots (left) of scRNA-seq data from esophagus (**A**), stomach (**B**), pancreas (**C**), small intestine (**D**), and colon and rectum (**E**); bubble plots showing expression levels of cell-type marker genes (middle); and bar plots depicting proportions of cell types co-expressing *ACE2* and *TMPRSS2* (right). EEC: enteroendocrine cell; GMC: antral basal gland mucous cell; PMC: pit mucous cell; TA: transit-amplifying cell.

## Figure 1

