Severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) infects epithelial cells of the human gastrointestinal (GI) tract and causes related symptoms. HIV infection impairs gut homeostasis and is associated with an increased risk of COVID-19 fatality. To investigate the potential link between these observations, we analysed single cell transcriptional profiles and SARS-CoV-2 entry receptor expression across lymphoid and mucosal human tissue from chronically HIV infected individuals and uninfected controls. Absorptive gut enterocytes displayed the highest co-expression of SARS-CoV-2 receptors ACE2, TMPRSS2 and TMPRSS4, of which ACE2 expression was associated with canonical interferon response and antiviral genes. Chronic treated HIV infection was associated with a clear antiviral response in gut enterocytes and, unexpectedly, with a significant reduction of ACE2 and TMPRSS2 target cells. Gut tissue from SARS-CoV-2 infected individuals, however, showed abundant SARS-CoV-2 nucleocapsid protein in both the large and small intestine, including an HIV co-infected individual. Thus, upregulation of antiviral response genes and downregulation of ACE2 and TMPRSS2 in the GI tract of HIV infected individuals, does not prevent SARS-CoV-2 infection in this compartment. The impact of these HIV-associated intestinal mucosal changes on SARS-CoV-2 infection dynamics, disease severity and vaccine responses remains unclear and require further investigation.
HIV infection drives interferon signalling within intestinal SARS-CoV-2 target cells

Rabiah Fardoos1,3*, Osaretin E. Asowata1,2*, Nicholas Herbert1,2, Sarah K. Nyquist7,8, Yenzekile Zungu1,2, Alveera Singh1, Abigail Ngoepe1, Ian Mbano1,2, Ntombifuthi Mthabela1, Dirhona Ramjit1, Farina Karim1, Warren Kuhn4, Fusi G. Madela5, Vukani T. Manzini5, Frank Anderson5, Bonnie Berger9, Tune H. Pers6, Alex K. Shalek7, Alasdair Leslie1,2,10 and Henrik N. Kløverpris1,2,3,10#

1Africa Health Research Institute (AHRI), Durban 4001, South Africa
2School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban 4001, South Africa
3Department of Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark
4ENT Department General Justice Gizenga Mpanza Regional Hospital (Stanger Hospital), University of KwaZulu-Natal, Durban 4001, South Africa
5Discipline General Surgery, Inkosi Albert Luthuli Central Hospital, University of KwaZulu-Natal, Durban 4001, South Africa
6Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen 2200, Denmark
7Institute for Medical Engineering & Science, Department of Chemistry, and Koch Institute for integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; Ragon Institute of MGH, Harvard & MIT, Cambridge, MA 02139
8Program in Computational & Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Computer Science & Artificial Intelligence Lab and Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

University College London, Division of Infection and Immunity, London, UK

*Equal contribution.

#Corresponding author:

Henrik N. Kløverpris, Africa Health Research Institute

K-RITH Tower Building

719 Umbilo Road

Durban 4001, South Africa

Cell: +27 74 546 6625

henrik.kloverpris@ahri.org
Severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) infects epithelial cells of the human gastrointestinal (GI) tract and causes related symptoms. HIV infection impairs gut homeostasis and is associated with an increased risk of COVID-19 fatality. To investigate the potential link between these observations, we analysed single cell transcriptional profiles and SARS-CoV-2 entry receptor expression across lymphoid and mucosal human tissue from chronically HIV infected individuals and uninfected controls. Absorptive gut enterocytes displayed the highest co-expression of SARS-CoV-2 receptors ACE2, TMPRSS2 and TMPRSS4, of which ACE2 expression was associated with canonical interferon response and antiviral genes. Chronic treated HIV infection was associated with a clear antiviral response in gut enterocytes and, unexpectedly, with a significant reduction of ACE2 and TMPRSS2 target cells. Gut tissue from SARS-CoV-2 infected individuals, however, showed abundant SARS-CoV-2 nucleocapsid protein in both the large and small intestine, including an HIV co-infected individual. Thus, upregulation of antiviral response genes and downregulation of ACE2 and TMPRSS2 in the GI tract of HIV infected individuals, does not prevent SARS-CoV-2 infection in this compartment. The impact of these HIV-associated intestinal mucosal changes on SARS-CoV-2 infection dynamics, disease severity and vaccine responses remains unclear and require further investigation.
**Introduction**

Gastrointestinal (GI) tract symptoms are observed in up to 60% of COVID-19 patients and precede respiratory symptoms (1). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can be detected within intestinal tissues (2, 3), and about 30% of COVID-19 patients harbour detectable viral RNA in their stool (4), which is associated with more severe gastrointestinal symptoms (1). The observed GI disorders include vomiting, nausea and diarrhoea, can manifest local and systemic disease, and may lead to faecal-oral transmission of virus (2, 5, 6). High expression of ACE2, the primary receptor for SARS-CoV-2, is found on the luminal surface of differentiated small intestinal epithelial cells, whereas crypt based cells express lower levels (7, 8). Studies of human small intestinal organoids show that the mature enterocytes are the major source for SARS-CoV-2 replication (2, 6, 9). These cells co-express the serine proteases TMPRSS2 and TMPRSS4, which promote SARS-CoV-2 spike protein fusion and viral entry into enterocytes (6).

HIV infection in the GI tract results in rapid and massive CD4 T-cell depletion, with associated changes in the microbiome and elevated translocation of microbial products across the epithelial barrier. These event precipitate a so-called ‘leaky gut syndrome’, which is thought to be central to HIV pathology (10–13). The elevated systemic immune activation, inflammatory responses and gut dysbiosis associated with this syndrome (14) may influence the overall type I interferon responses, and therefore compromise both local and systemic responses to SARS-CoV2 infection, including responses within the lung mucosa (15). Multiple studies prior to the SARS-CoV-2 pandemic have demonstrated that HIV induced systemic immune activation renders HIV infected individuals vulnerable to airborne infection, such as tuberculosis (16, 17). Recent data suggest that SARS-CoV-2 and HIV co-infected individuals
overall have a 2.1 fold increased risk of dying from COVID-19 (16) with a risk factor reaching >3.5 in viremic individuals with CD4 count <200 cells/mm³ (18). This is consistent with low nadir CD4 T-cell counts associated with increased mortality in COVID-19 patients (19). Therefore, understanding the dynamics of SARS-CoV-2 infection in the gut of HIV infected individuals is a pressing area of research.

Single-cell transcriptomic analysis from different tissues reported increased co-expression of ACE2 and TMPRSS2 transcripts in the ileum of SHIV infected non-human primates and within the lung of HIV infected humans compared to uninfected controls (20), suggesting a potential higher risk of SARS-CoV-2 infection in HIV infected individuals. One of the isoforms of ACE2 is encoded by Exon 1c (dACE2), has recently been confirmed as an interferon-stimulating gene (ISG), which responds directly to type I interferons and SARS-CoV-2 infection in the small intestine (21). Therefore, type 1 interferon stimulation and altered regulation of ACE2 could be exploited by SARS-CoV-2 to enhance infection (20). However, no human studies have determined the impact of HIV infection on SARS-CoV-2 entry receptors in the gut, or on the transcriptional landscape of the gut enterocytes expressing them.

To explore the potential impact of HIV infection on gut epithelial cells susceptible to SARS-CoV-2 infection, we investigated the gene expression profile of ACE2, TMPRSS2 and TMPRSS4 using single-cell RNA-seq (scRNA-seq) datasets from human SARS-CoV-2 uninfected tonsil, liver, lymph-node, duodenum and blood, as well as a published human lung scRNA-seq dataset (20). All study participants were from clinics in KwaZulu-Natal, South Africa, recruited within extremely high HIV-1 endemic areas. Our data show that ACE2 is associated with interferon response genes in HIV uninfected individuals, supporting the hypothesis that ACE2 expression is linked with interferon signalling within gut enterocytes. In addition, and as expected,
we show that HIV infection itself drives a strong interferon signalling response in these cells. Surprisingly, however, HIV infection was associated with a down regulation of ACE2 expression in all cell types studied. These data suggest that, although ACE2 is associated with interferon signalling genes, the isoforms detected here are differentially regulated from canonical ISGs. The reduction in potential target cells may impact infectability, propagation of infection of SARS-CoV2 in HIV infected individuals, and local and systemic immunity beyond the gut, such as the lung mucosa (15). Nevertheless, using gut biopsies from co-infected individuals we observed the presence of SARS-CoV2 and HIV proteins in gut enterocytes, indicating this compartment remains vulnerable to infection.
Results

**SARS-CoV-2 host entry receptors are enriched in the human small intestine**

SARS-CoV-2 infection of human cells requires surface expression of the primary receptor **ACE2** and one of the co-receptors **TMPRSS2** (22) or **TMPRSS4** (3, 6). scRNA-seq analysis showed that **ACE2** expression is primarily restricted to type II pneumocytes in the lung, gut absorptive enterocytes and goblet secretory cells of the nasal mucosa (20). In this study, we applied high throughput scRNA-seq (23, 24) to profile distinct human tissue samples (Figure 1A) and analysed 32,381 cells from blood, tonsil, lung, duodenum, mesenteric lymph node, (lung dataset has been previously described here (20))(Figure 1B), for expression of genes encoding **ACE2**, **TMPRSS2** and **TMPRSS4**. We identified 18 cellular distinct sub-clusters in this dataset (Figure S1A) with the majority of **ACE2**, **TMPRSS2** and **TMPRSS4** expressing cells located within duodenum and to a lesser extent in lung tissue, and no expression in the other tissues analysed (Figure 1C). We next used fluorescent immunohistochemistry to co-stain epithelial cells (EpCAM) for **ACE2** and **TMPRSS2** and found distinct *in situ* expression patterns between the small and large intestine (Figure 1D and Figure S1B). The duodenum tissue expressed high levels of **ACE2** facing outward towards the lumen side, whereas the colon tissue was characterised by **ACE2** expression at the bottom half of the crypt. **TMPRSS2** was expressed in most of the epithelial cells and was co-expressed with **ACE2** in both duodenum and colon tissue. We found no quantitative differences in protein expression levels between both compartments for **ACE2** and **TMPRSS2** (Table S1 and Figure 1E). High SARS-CoV-2 entry receptor expression in the small intestine is consistent with recent reports (2, 5, 6, 20) and with detection of SARS-CoV-2 nucleocapsid in the small intestine of
infected individuals (3). Locational differences in ACE2 expression between the duodenum and colon may be linked to the physiology of ACE2 in these compartment such as regulation of intestinal inflammation and amino acid homeostasis (7, 8) and may influence SARS-CoV-2 infection dynamics in each compartment.

**ACE2 expression is dominated by absorptive enterocytes**

Next, focusing on the small intestine only, for which most samples were available, we evaluated the expression of ACE2, TMPRSS2, TMPRSS4 and other transmembrane serine proteases in duodenal cellular subsets using scRNA-seq data from 4 HIV uninfected female donors (Table 1 and Table S1). To assign cellular identity, we performed variable gene selection, dimensionally reduction by uniform manifold approximation and projection (UMAP) with graph-based clustering (Figure 2A), created a cells-by-genes expression matrix (Figure 2B) and identified 15 distinct major cell clusters from 13,056 duodenum cells. Cell cluster identities were assigned based on most highly differentially expressed genes (DEGs) for each cluster (Figure 2B, Table S2A). This approach identified absorptive enterocytes as the predominant ACE2 expressing cell type, defined by expression of APOA1 and APOA4 (41.3% ACE2+, false discovery rate [FDR]-adjusted p = 1.01E-227) (Figure 2B). These cells also more frequently expressed TMPRSS2 and TMPRSS4 than any of the other subsets (Figure 2C and Figure S2A). The majority of absorptive enterocytes also expressed ST14 and TMPRSS15, and to a much lesser extent TMPRSS3 and 6 (Figure S2B). There is no known role for these other serine proteases in SARS-CoV-2 infection, but it does imply the importance of this family of molecules to the biology of absorptive enterocytes (25).

Focusing on epithelial cells, we consistently identified absorptive enterocytes as the main source of ACE2, TMPRSS2 and TMPRSS4 gene expression (Figure S3A-D) and
SARS-CoV-2 putative target cells due to co-expression of these markers; 

\(ACE2^+TMPRRSS2^+\) (412 cells, 12% of all epithelial cells), and \(ACE2^+TMPRSS4^+\) subset 

(392 cells, 11% of all epithelial cells), and a smaller subset co-expressing all 3 genes \(ACE2^+TMPRSS2^+TMPRSS4^+\) 154 cells, 5%; Figure S3E). Using a monocle single cell trajectory analysis, we uncovered a clear potential differentiation trajectory from ‘intestinal stem cell like’ cells (LGR5, CD44, EPHB2), through transit amplifying cells (OLFM4) and paneth cells (DEFA5, DEFA6) to the dominant absorptive enterocyte clusters (APOA1, APOA4) enriched for putative SARS-CoV-2 susceptible cells (Figure S4), suggesting that terminal differentiated cells are likely to be the most susceptible to infection. Thus, cellular differentiation was linked to SARS-CoV-2 entry receptor expression. Overall, these observations are in line with previous reports indicating that \(TMPRSS2\) and \(TMPRSS4\) are important co-receptors for SARS-CoV-2 fusion and entry into \(ACE2^+\) differentiated enterocytes in the human duodenum and small intestine (6, 26).

**ACE2 expression is associated with interferon stimulating genes**

Next, we compared \(ACE2^+\) expressing (\(ACE2^+\)) with non-expressing (\(ACE2^-\)) cells and identified 381 DEGs \((P<0.05)\) (Table S2B). The majority of these genes were upregulated together with APOB and APOA4, confirming that \(ACE2\) expression is associated with absorptive enterocytes (Figure 3A). In addition, these DEGs included a range of known interferon response genes, such as \(ISG20\), which interferes with viral replication (27); \(STAT6\), which is important for immune signals emanating from interleukin-4 receptors (28); \(TNFSF10\), a proapoptotic cytokine (29, 30); \(TNFRSF1A\) involved in TNF alpha signalizing, and the proinflammatory cytokine \(IL-32\), which is involved in the pathogenesis and progression of a number of inflammatory disorders.
In support of the cell sub-setting approach, pathway analysis of ACE2+ cells was consistent with functional absorptive cells, including pathways involved in ‘transport of small molecules’, ‘mineral absorption’, ‘intestinal absorption’ and ‘bile secretion’ etc (Figure 3C). Further sub-setting on ACE2+TMPRSS2+ cells as putative SARS-CoV-2 targets and comparison to the remaining epithelial cells (Table S2C), showed significant upregulation of IFNGR1 within ACE2+TMPRSS2+ cells (Table S2C) further supporting the expression of ISG signatures in susceptible target cells. Analysis of potential upstream drivers of the ACE2 associated DEGs, identifies canonical ISG genes such as STAT3, IRF2 and IRF4 (Figure 3D and Table S2D). These data show that epithelial cells, and particularly the absorptive enterocytes, are enriched for ACE2, TMPRSS2 and TMPRSS4 expression, and that ACE2 expression in these cells is upregulated in conjunction with known ISGs.

**HIV infection drives interferon response in gut absorptive enterocytes**

HIV infection depletes CD4 T-cells in the gut and has a major impact on the barrier integrity (10, 12), and the intestinal epithelium is dominated by absorptive enterocytes with high expression levels of SARS-CoV-2 entry receptors compared to other epithelial subsets (see Figure S3 and Table S3A). Therefore, we next determined the influence of HIV infection on potential SARS-CoV-2 target cells, by comparing the gene expression of cells from duodenal tissues collected from individuals with chronic HIV infection on long term ART and HIV uninfected participants. We found 160 differentially expressed genes (134 upregulated and 26 downregulated; Figure 4A, Table S3B). Consistent with the gut mucosal inflammation associated with HIV (32), these included strong upregulation of numerous canonical ISG and antiviral genes, including ISG15, IFI6, LY6E, IFITM3, IFI27, MX1 and IRF7 (Figure 4B, Table S3B). In
contrast to earlier studies (20), and despite the observed interferon signalling, ACE2 expression was significantly reduced in enterocytes isolated from subjects with treated chronic HIV infection (Figure 4B) suggesting that ACE2 expression itself may not directly act as an ISG (21). Pathway analysis of DEGs in absorptive enterocytes, confirmed that HIV infection was associated with a profound upregulation of interferon signalling (Figure 4C). Analysis of potential upstream drivers of these DEGs, indicated type I IFN and multiple interferon response factors (IRFs), consistent with a constitutive antiviral response program within absorptive enterocytes induced by HIV infection despite ART (Figure 4D and Table S3C). Finally, sub-setting on the remaining ACE2 expressing epithelial cells only showed a consistent strong upregulation of these canonical ISGs in HIV infected subjects (Figure S5A-B, Table S3D). Taken together, these data show that, despite suppression of plasma viremia, HIV infection induces a strong interferon antiviral response in gut enterocytes in general and also within ACE2 expressing putative SARS-CoV-2 target cells. However, this does not drive ACE2 expression itself, which is significantly reduced in HIV infected individuals.

**Reduced numbers of ACE2 expressing absorptive enterocytes and SARS-CoV-2 target cells in HIV infected individuals**

Having observed an unexpected down regulation of ACE2 in HIV infected subjects, we quantified the number of epithelial cells susceptible for SARS-CoV-2 infection and compared the relative frequencies of absorptive enterocytes, goblet cells and transit amplifying cells (see Figure 2) between HIV infected and uninfected individuals. Overall, the number of ACE2 and TMPRSS2 expressing absorptive enterocytes were significantly reduced in HIV infection, while TMPRSS4 was unaffected (Figure 5A). The same trend was observed for goblet cells but not for transit amplifying cells (Figure
When we analysed absorptive enterocytes expressing two or more entry receptors, we found a similar significant reduction of all combinations of SARS-CoV-2 putative target cells with the same trend for goblet cells, but not for transit amplifying cells (Figure S6). Thus, chronic HIV infection appears to reduce the total frequency of SARS-CoV-2 putative target cells within the small intestine.

**Abundant SARS-CoV-2 detection in small and large intestine irrespective of HIV co-infection**

Finally, to determine if HIV associated loss of SARS-CoV-2 target cells and upregulation of antiviral genes prevented infection of the GI tract, we collected gut tissue from confirmed SARS-CoV-2 antigen PCR positive participants with and without HIV infection. From a duodenum biopsy obtained from a SARS-CoV-2 and HIV co-infected individual on antiretroviral therapy and with undetectable plasma HIV viremia (Table S1), we identified abundant expression of SARS-CoV-2 nucleocapsid protein (NP) within the epithelial layer and colocalized with that of ACE2 entry receptor (Figure 6A). We found small, but detectable levels of HIV-p24 protein in the same area, but in cells that did not express ACE2, consistent with distinct viral entry receptor usage between SARS-CoV-2 and HIV infection. We repeated this staining in a pre-pandemic control sample and observed no SARS-CoV-2 NP staining (Figure 6B). Histology of colon samples collected from an HIV uninfected SARS-CoV-2 PCR positive donor also revealed abundant SARS-CoV-2 NP expression within the epithelial tissue overlapping with ACE2 (Figure 6C). These data confirm the presence of high levels of SARS-CoV-2 virus production in the GI tract (3) and show it is likely to occur in HIV infected individuals despite upregulation of antiviral immunity and a loss of putative target cells in the small intestine.
Discussion

In this study we performed single cell transcriptomic profiling across different human tissue sites and identified high expression of SARS-CoV-2 entry receptors within absorptive enterocytes from the small intestine that we confirmed by in situ protein staining. We detected overlapping expression of SARS-CoV-2 viral NP and ACE2 within both the small and large intestine of SARS-CoV-2 infected individuals, confirming the infectability of these cells in vivo. ACE2, TMPRSS2 and TMPRSS4 expression were highest in the duodenum followed by the lung, with little or no expression detected in the tonsil, liver, lymph-node and blood consistent with published studies (20, 25, 33, 34). We found that ACE2 protein expression was restricted to the luminal region of the enterocytes in the duodenum, whilst in the colon, ACE2 was located closer to the crypt base (7, 8). This distinct location may be explained by differences in the physiological processes within these compartments where luminal duodenal ACE2 is reported to be important in amino acid transport and protein synthesis (35, 36) and consistent with our pathway analysis of ACE2 expressing duodenal epithelial subsets (see Figure 2F).

In HIV uninfected subjects, ACE2 expression in the small intestine was associated with genes involved in interferon signalling, in agreement with recent observation and experimental data demonstrating upregulation of ACE2 in response to interferon signalling (20). However, more recently investigators have established that it is the truncated isoform of ACE2, dACE2, that is most likely to act as a ISG and not ACE2. This study found that dACE2 was directly upregulated by both interferon stimulation and SARS-CoV-2 infection within human intestinal organoid cells, but the full length ACE2 was not (21). Importantly, dACE2 does not function as a SARS-CoV-2 entry
receptor. In future, it would be interesting to also sequence the gut virome of these individuals to determine how this may contribute to interferon signalling within the intestinal mucosa (37). The transcriptomic profile of enterocytes from chronic HIV infected individuals was characterised by a strong interferon signalling pathway that included upregulation of canonical ISGs such as \( ISG15 \) and \( IFI27 \) predicted to be driven by type I and II interferons. These data highlight the impact from HIV infection on the small intestine and contribute to mechanisms underlying the functional consequences in gut barrier integrity and overall pathology (10–13). The reduced frequency of \( ACE2 \) expressing cells in the intestinal mucosa was therefore unexpected. Both \( ACE2 \) and canonical ISGs (\( STAT1 \) and \( IFI6 \)) levels remained elevated in absorptive enterocytes from NHP with treated SIV infection (20). Indeed, \( IFI6 \) was also highly upregulated in HIV infected subjects in our cohort (Figure 3B). Why \( ACE2 \) expressing cells were reduced is not clear from this study. Changes in the microbiome, however, have recently shown to alter \( ACE2 \) expression levels and disrupt the ACE/ACE2 axis and HIV is known to cause gut dysbiosis (38, 39). It will therefore be interesting to investigate the link between HIV infection, altered microbiomes and \( ACE2 \) expression levels in the gut of HIV infected individuals. The monocle lineage analysis conducted here suggests that \( ACE2 \) expression may be upregulated as enterocytes progress towards terminal differentiation. Therefore, a reduction of \( ACE2 \) expressing cells could result from interference in this process, or in increased cell death of terminally differentiated \( ACE2 \) expressing cells. Individuals in this study were all female on long-term fully suppressive ART, which may distinguish them somewhat from the experimentally infected NHP, which were treated for 6 months. In addition, the impact of the gut microbiome on \( ACE2 \) expression in the small intestine, discussed above, may affect comparisons between experimental NHP
studies and human cohorts (38, 39). We only used females for the transcriptional data in this study to avoid gender biased gene expression, and therefore extending these gene expression profiles beyond females will require further validation. However, the persistent interferon signature in these subjects clearly implies that upstream drivers, such as type I and II interferons, have not diminished.

Although the implications for infectability of the gut mucosa for the SARS-CoV-2 virus remains unclear, our data from HIV and SARS-CoV-2 co-infected participants from which we obtained ex vivo duodenum biopsies showed that SARS-CoV-2 infection certainly can occur in the gut of HIV infected individuals. Whether HIV infected individuals have longer SARS-CoV-2 sequela from the gut (3) would be interesting to study in a large co-infected gut biopsy cohort, and particularly the impact of variant viruses, which may have different affinity for cell entry receptors (40, 41).

Although the literature is still emerging, in general, studies have observed COVID-19 patients with controlled HIV co-infection and preserved CD4 T cell counts have similar clinical trajectories to those without HIV infection (18, 42, 43). By contrast, immune compromised HIV infected individuals with CD4 T cell counts below 200 cells/mm³ are associated with increased COVID-19 disease severity and mortality (16, 18). Whether reduction of potential SARS-CoV-2 target cells in gut mucosa of HIV infected subjects limits the additional effects of SARS-CoV-2 infection in this compartment warrant further studies. Further intestinal sampling from co-infected individuals are currently being sought to address this urgent question. Indeed, although the actual number of SARS-CoV-2 target cells are reduced, interferon signalling and pathways of metabolic and absorptive changes observed here in HIV infected subjects prior to SARS-CoV-2 infection, may still exacerbate disease or limit immunity in some individuals. Whether increased transmissibility of new SARS-CoV-2 variants, including beta and the recent
dominat delta variants identified in high HIV prevalence populations in South Africa (44, 45), alter the GI related symptoms and overall COVID-19 disease severity is also currently unknown. Finally, potentially altered induction of type I interferons by evolving SARS-CoV-2 variants could potentially also be linked to signalling within intestinal enterocytes and should be explored further.
Methods

Study participants

Patients presenting to the gastrointestinal (GI) surgical unit of Inkosi Albert Luthuli Central Hospital (IALCH) were recruited into this study after obtaining written informed consent. Tonsil, liver, gut lymph node, duodenum and colon biopsies with participants-matched blood samples were obtained during surgical procedures. Clinical information, including HIV status and demographic details of these participants, was collected using a structured questionnaire. HIV status was confirmed using the Determine HIV 1/2 Set (Abbott Laboratories) and COBAS® TaqMan® HIV-1 Test (Roche).

Sample processing

Mononuclear cells were isolated from blood, tonsil, liver, lymph node and pooled duodenum pinches to average over individual pinch variation. Blood was collected in BD vacutainers with sodium heparin (Becton Dickinson). Peripheral blood mononuclear cells (PBMCs) were isolated using the Ficoll-Histopaque 1077 (Sigma-Aldrich) density gradient centrifugation. Duodenum and colon pinch biopsies (2 to 4 pinches) were removed by the operating GI surgeon and transported to the laboratory in cold Phosphate-Buffered Saline (PBS) (pH 7.2). The PBS was decanted from the tubes containing the gut biopsies, which are about 5-8 mm in size, and they were incubated in epithelial strip buffer (PBS; 0.5M EDTA; 1M DTT; FBS and Penicillin/Streptomycin) in a 37°C water bath for 10 minutes, with occasional agitation. Thereafter, the epithelial strip buffer was removed, and the tissues were digested in a buffer containing Collagenase-D (0.5 mg/ml; Roche) and DNase-I (20 µg/ml; Sigma-
Aldrich) for 30 minutes in a 37°C water bath with occasional agitation. Digested tissue was passed through a 70µM cell strainer to isolate the cells, and these cells were washed with PBS.

**Single-cell RNA-seq using Seq-Well S³**

After obtaining single-cell suspension from fresh biopsies, we used the Seq-Well S³ platform. Full methods on implementation of this platform is described (23, 24). Briefly, 15,000 cells in 200 mL RPMI + 10% FBS were loaded onto one PDMS array preloaded with barcoded mRNA capture beads (ChemGenes) and settled by gravity into distinct wells. The loaded arrays were washed with PBS and sealed using a polycarbonate membrane with a pore size of 0.01 µm, which allows for exchange of buffers but retains biological molecules within each nanowell. Arrays were sealed in a dry 37°C oven for 40 min and submerged in a lysis buffer containing guanidium thiocyanate (Sigma), EDTA, 1% beta-mercaptoethanol and sarkosyl (Sigma) for 20 min at room temperature. Arrays were transferred to hybridization buffer containing NaCl (Fisher Scientific) and supplemented with 8% (v/v) polyethylene glycol (PEG, sigma) and agitated for 40 min at room temperature, mRNA capture beads with mRNA hybridized were collected from each Seq-Well array, and beads were resuspended in a master mix for reverse transcription containing Maxima H Minus Reverse Transcriptase (ThermoFisher EP0753) and buffer, dNTPs, RNase inhibitor, a 50 template switch oligonucleotide, and PEG for 30 min at room temperature, and overnight at 52°C with end-over-end rotation. Exonuclease I treatment (New England Biolabs M0293L) to remove excess primers. After exonuclease digestion, bead-associated cDNA denatured for 5 min in 0.2 mM NaOH with end over end rotation. Next, beads were washed with TE + 0.01% tween-20, and second strand synthesis was carried out by
resuspending beads in a master mix containing Klenow Fragment (NEB), dNTPs, PEG, and the dN-SMRT oligonucleotide to enable random priming off of the beads. PCR amplification were carried out using KAPA HiFi PCR Mastermix (Kapa Biosystems KK2602) with 2.00 beads per 50 µL reaction volume. Post—whole transcriptome amplification, libraries were then pooled in sets of six (12,000 beads) and purified using Agencourt AMPure XP SPRI beads (Beckman Coulter, A63881) by a 0.6x volume ratio, followed by a 0.8x. Libraries size was analysed using an Agilent Tapestation hsD5000 kit (Agilent Genomics) with an expected peak at 1000 bp and absence of smaller primer peaks. Libraries were quantified sing Qubit High-Sensitivity DNA kit and preparation kit and libraries were constructed using Nextera XT DNA tagmentation (Illumina FC-131-1096) using 800 pg of pooled cDNA library as input using index primers with format as in Gierahn et al. Amplified final libraries were purified twice with AMPure XP SPRI beads as before, with a volume ratio of 0.6x followed by 0.8x yielding library sizes with an average distribution of 650-750 pb. Libraries from 16 Seq-Well arrays were pooled and sequenced together using a Illumina NovaSeq 6000 S2 Reagent kit v1.5 (100 cycles) using a paired end read structure with custom read 1 primer: read 1: 20 bases with a 12 bases cell barcode and 8 bases unique molecular identifier (UMI). Read 2: 82 bases of transcript information, index 1 and index 2: 8 bases.

**Single-cell RNA-seq Computational Pipeline and Analysis**

Raw sequencing data was converted to demultiplexed FASTQ files (online access XXX) using bcl2fastq2 based on Nextera N700 indices corresponding to individual arrays. Reads were then aligned to hg19 genome assembly and aligned using the Dropseq-tools pipeline on Terra (app.terra.bio). Data was normalized and scaled using
Seurat R package v.3.1.0 ([https://satijalab.org/seurat/](https://satijalab.org/seurat/)), any cell with fewer than 750 UMI count or greater than 2500 UMI was excluded for further analyses. This cell-by-genes matrix was then used to create a Seurat object. Cells with any gene expressed in fewer than 5 cells were discarded from downstream analysis and any cell with at least 300 unique genes was retained. Cells with >20% of UMI mapping to mitochondrial genes were then removed. These objects were then merged into one object for pre-processing and cell-type identification. The combined Seurat object was log-normalized to UMI+1 using a scale factor of 10,000. We examined highly variable genes across all cells, yielding 2000 variable genes. Principal component analysis was applied to the cells to generate 100 principal components (PCs). Using the JackStraw function within Seurat, we identified significant PCs to be used for subsequent clustering and further dimensionality reduction. For 2D visualization and cell type clustering, we used a Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction technique and with “min_dist” set to 0.5 and “n_neighbors” set to 30. To identify clusters of transcriptionally similar cells, we employed unsupervised clustering as described above using the FindClusters tool within the Seurat R package with default parameters and k.param set to 10 and resolution set to 0.5. We applied the default parameters with a shared nearest neighbour parameter optimized for each combined dataset inside Monocle3 package (V3.2.0) to construct single cell pseudo-time trajectory to discover differential transitions. We used highly variable genes identified by Seurat to sort cells in pseudo-time order. The actual precursor determined the beginning of pseudo-time in the first round of “orderCells”. UMAP was applied to reduce dimensional space and the minimum spanning tree on cells was plotted by the visualization functions “plot_cells” for Monocle 3. To further characterize substructure within cell types (for example, epithelial cells), we performed
dimensionality reduction (PCA) and clustering over those cells alone. Differential
expression analysis between the negative and positive groups of the same cell type
was performed using the Seurat package FindAllMarkers in Seurat v3 (setting
“test.use” to bimod). For each cluster, differentially expressed (DEGs) were generated
relative to all of the other cells. Gene ontology, gene-set enrichment analysis and
KEGG pathway analyses from DEGs were performed using Metacape webtool
(www.metascape.org), which supports statistical analysis and visualization profiles for
genes and gene clusters.

**Histology and Multi-colour fluorescent immunohistochemistry**

Formalin (4%) fixed duodenum and colon tissue samples were embedded in paraffin,
and a 4 µm section was obtained on a glass slide. These sections were deparaffinized
and incubated with the following antibodies: anti-HIV p24 (clone: Kal-1, Dako), anti-
SARS-CoV-2 nucleoprotein (clone: 40143-T-62, Sino Biological) anti-ACE2 (clone:
ab15348; Abcam), anti-TMPRSS2 (clone: ab109131; Abcam) and anti-EpCAM (clone:
ab71916; Abcam) followed by a secondary antibody incubation using the Opal™ 4-
color manual IHC (Perkin Elmer, USA) as instructed by the manufacturer. Opal
fluorophores: FITC was used for EpCAM and p24, Texas Red for ACE2 and Cy5 for
TMPRSS2 and SARS-CoV-2 nucleoprotein signal generation. DAPI was used as the
nuclear counterstain. The sections were mounted with the Opal mounting oil, cover-
slipped and the edges were sealed with nail vanish. The slides were stored 2-8°C
temperature until images are acquired.

**Microscopy and Quantitative Image Analysis**
Images of the tissue sections were acquired using the TissueFAXS software (TissueGnostics) connected with a Zeiss Axio Observer Z1 inverted microscope (Olympus). The quantitative analysis of the cells of the different phenotypes within the images was done using the TissueQuest quantitation software (TissueGnostics).

**Statistical methods**

Graphs were plotted using Prism 8.4.3 (GraphPad Inc.). Difference between groups were analysed using the Seurat package FindAllMarkers in Seurat v3 (setting “test.use” to bimod). If any other specific test used, it has been stated in the figure legends.

**Study Approval**

This study was approved by the Biomedical Research Ethics Committee (BREC) of the University of KwaZulu-Natal (BE 021/13) and (BE061/13).

**Author contributions**

RF and OA performed experiments and analysed transcriptional data. SN supervised transcriptional analysis. NH, YZ, AS, AN and IM contributed to experimental work. NM consented participants and collected samples. DR and FK coordinated human tissue sample collection. WK, FM, VM and FA contributed surgical human tissues samples. BB, TP and AKS supervised data analysis. AKS, AL and HK provided intellectual input. RF, OA, AL and HK prepared the manuscript. HK conceptualized and supervised the work.

**Declaration of interests**

Authors declare no competing interests.
Acknowledgements

HNK is supported by The Wellcome Trust (#202485/Z/16/Z). AL is supported by the Wellcome Trust (210662/Z/18/Z). This work was supported through the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), a DELTAS Africa Initiative [grant # DEL-15-006]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)’s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa’s Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant # 107752/Z/15/Z] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government. HK and AS was supported by SANTHE. AKS was supported, in part, by the Searle Scholars Program, the Beckman Young Investigator Program, the NIH (5U24AI118672, 2R01HL095791, 2U19AI089992, 1R01HL134539, 1R01AI138546), a Sloan Fellowship in Chemistry, and the Bill and Melinda Gates Foundation. SWK was supported by the Hugh Hampton Young Memorial Fellowship.
References


2. Lamers MM et al. SARS-CoV-2 productively infects human gut enterocytes [Internet]. *Science (80-. ).* 2020;369(6499):50–54.


8. Hashimoto T et al. ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation [Internet]. *Nature* 2012;487(7408):477–481.


44. San JE et al. Transmission Dynamics of SARS-CoV-2 Within-Host Diversity in Two Major Hospital Outbreaks in South Africa [Internet]. *Virus Evol.* [published online ahead of print: April 21, 2021]; doi:10.1093/ve/veab041

Figure 1. SARS-CoV-2 putative target cells are enriched in the human duodenum.

A. Schematic of protocol for isolation of different tissues for scRNA-seq using Seq-Well S3, to identify cell types. B. UMAP of 32381 cells colored by tissue source. C. Left: UMAP projection of epithelial cells showing expression of $ACE2$ (top), $TMPRSS2$ (middle) and $TMPRSS4$ (bottom) among all tissue source from human donors. Color coding is as follows: purple, RNA positive; grey, RNA negative. Right: Corresponding violin plots of expression values for $ACE2$ (top), $TMPRSS2$ (middle) and $TMPRSS4$ (bottom). D. Representative fluorescent immunohistochemistry image of gut tissue showing $ACE2$ (red), $TMPRSS2$ (orange), EpCAM (green) and DAPI (blue) of duodenum and colon. Bars, 20 µm for all images. E. Quantification of $ACE2$ and $TMPRSS2$ proportion of total cell stained with Epcam.
Expression Level

ACE2

TMPRSS2

TMPRSS4
**Figure 2. ACE2 expression is enriched in absorptive enterocytes.**

**A.** UMAP projection of 13056 cells from endoscopic pinch biopsies, colored by cell type. **B.** Dot plot of two defining genes for each cell type, with ACE2, TMPRSS2 and TMPRSS. Dot size represents fraction of cells within cell type expressing a given gene, and color intensity represents binned count-based expression amounts (log(scaled UMI+1)) among expressing cells. Red arrow indicated cell type with largest proportion of ACE2+TMPRSS2+TMPRSS4+ cells; full results can be found in Table S2A. **C.** Expression of ACE2 (left), TMPRSS2 (middle) and TMPRSS4 (right) among all subsets from duodenum.
Figure 3. ACE2 expressing absorptive enterocytes are linked to interferon signaling genes and functional absorptive pathways.

A. Volcano plot of DEGs (Table S2B) within epithelial cells from HIV uninfected individuals (n=4) highlighting genes with more than 0.5 fold change and adjusted P < 5.0E10-08. B. Genes differentially expressed among ACE2+ and ACE2− epithelial cells, FDR-adjusted P < 0.05; full results can be found in Table S2B. C. GO BP enrichment analysis of the DEGs of epithelial cells analysis upregulated in the ACE2+ compared to ACE2-. P-value was derived by a hypergeometric test. D. Selected upstream drivers of pathways shown in C from DEGs in Table S2D.
Figure 4. HIV infection downregulates *ACE2* expression and drive interferon signaling in absorptive enterocytes

**A.** Volcano plot of DEGs (Table S3B) within absorptive enterocytes in HIV infected and HIV uninfected cells highlighting genes with more than 0.5-fold change and adjusted $P < 5.0\text{E}10^{-08}$. **B.** Expression of *ACE2* and interferon-responsive genes among absorptive enterocytes from HIV negative (n= 4) and HIV+ART+ (n=5). **C.** GO BP enrichment analysis of the DEGs of absorptive analysis upregulated in the HIV negative (n= 4) and HIV+ART+ (n=5). *P*-value was derived by a hypergeometric test. **D.** Activation z-score of upstream drivers from DEGs shown in A and Table S3D color coded by their functional categories.
A Absorptive enterocytes

B Goblet cells

C Amplifying cells
Figure 5. HIV infection reduces the frequency of SARS-CoV-2 putative target cells within the small intestine.

A. Actual number of absorptive enterocytes (left) and % expression (right) ACE2, TMPRSS2, TMPRSS4 by HIV status. B. Number of goblet cells (left) and % (right) expressing ACE2, TMPRSS2, TMPRSS4 by HIV status. C. Number of transit amplifying cells (left) and % (right) expressing ACE2, TMPRSS2, TMPRSS4 by HIV status. ‘Rest’ refer to cells not expressing the indicated transcript. P-values by Fisher’s Exact Test.
A) Donor A - (HIV+ART+, SARS-CoV-2+) Duodenum

B) Donor B (pre-pandemic, HIV-, SARS-CoV-2-) Duodenum

C) Donor C - (HIV-, SARS-CoV-2+) Colon
Figure 6. SARS-CoV-2 nucleocapsid detection overlap with ACE2 expression in the small and large intestine. Representative fluorescent immunohistochemistry (F-IHC) images of duodenum and colon tissues showing HIV-p24 (green), ACE2 (red), SARS-CoV-2 nucleocapsid protein (orange) and DAPI (blue). A. F-IHC image of a duodenum tissue from an HIV⁺SARS-CoV-2⁺ PCR positive (nasopharyngal swab) including no antibody control (top). B. F-IHC image of a duodenum tissue from an HIV⁻SARS-CoV-2⁺ participant. C. F-IHC image of a colon tissue from HIV⁺SARS-CoV-2⁺ with two sections shown from the same biopsy tissue. Scale bars are shown at the bottom right of each image.
Table 1. Characteristics of study participants (n=24)

<table>
<thead>
<tr>
<th></th>
<th>HIV−</th>
<th>HIV−ART+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;Duo (n=7)</td>
<td>&quot;Colon (n=7)</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age years, median (IQR)</td>
<td>43 (33-53)</td>
<td>46.5 (35-53)</td>
</tr>
<tr>
<td>Gender, n (%) median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female, 19 (79)</td>
<td>43 (27-54)</td>
<td>46.5 (22-52)</td>
</tr>
<tr>
<td>Male, 5 (21)</td>
<td>39 (33-44)</td>
<td>50 (42-57)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, 14 (63)</td>
<td>3 (43)</td>
<td>3 (43)</td>
</tr>
<tr>
<td>Colored, 1 (4)</td>
<td>NA</td>
<td>1 (14)</td>
</tr>
<tr>
<td>Indian, 8 (33)</td>
<td>4 (57)</td>
<td>3 (43)</td>
</tr>
<tr>
<td>White, 1 (4)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CD4 count (cells/µl), median (IQR)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Viral load (copies/ml), median (IQR)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA=not applicable, ND=not done, "F-­IHC (n=14), "scRNA-­Seq (n=10)