

Journal Pre-proof

Coronavirus disease (COVID-19) and fertility: viral host entry protein expression in male and female reproductive tissues

Kate E. Stanley, BA, Elizabeth Thomas, BA, Megan Leaver, MSc, Dagan Wells, PhD



PII: S0015-0282(20)30435-0

DOI: <https://doi.org/10.1016/j.fertnstert.2020.05.001>

Reference: FNS 32483

To appear in: *Fertility and Sterility*

Received Date: 24 April 2020

Revised Date: 29 April 2020

Accepted Date: 2 May 2020

Please cite this article as: Stanley KE, Thomas E, Leaver M, Wells D, Coronavirus disease (COVID-19) and fertility: viral host entry protein expression in male and female reproductive tissues, *Fertility and Sterility* (2020), doi: <https://doi.org/10.1016/j.fertnstert.2020.05.001>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Copyright ©2020 Published by Elsevier Inc. on behalf of the American Society for Reproductive Medicine

Running title: Coronavirus disease (COVID-19) and fertility

Title: Coronavirus disease (COVID-19) and fertility: viral host entry protein expression in male and female reproductive tissues.

Authors: Kate E. Stanley¹, BA, Elizabeth Thomas¹, BA, Megan Leaver¹, MSc, Dagan Wells^{1,2}, PhD

- 1) Nuffield Department of Women's and Reproductive Health, University of Oxford, John Radcliffe Hospital, Women's Centre, Oxford OX3 9DU, UK.
- 2) Juno Genetics, Winchester House, Oxford Science Park, Oxford OX3 4GE, UK.

Email Addresses for all authors:

Kate E. Stanley, BA: kate.stanley@some.ox.ac.uk

Elizabeth Thomas, BA: elizabeth.thomas@gtc.ox.ac.uk

Meagan Leaver, MSc: megan.leaver@bnc.ox.ac.uk

Dagan Wells, PhD: dagan.wells@wrh.ox.ac.uk

Corresponding Author:

Dagan Wells, PhD

Juno Genetics

Winchester House

Oxford Science Park

Oxford

OX4 4GE

United Kingdom

Capsule

The effects of COVID-19 on male and female fertility are uncertain. To improve understanding, the gene and protein expression patterns of SARS-Cov-2 host entry proteins were assessed in reproductive tissues.

Keywords (MeSH terms)

COVID-19, severe acute respiratory syndrome coronavirus 2, Ovary, Testis, Fertility

Abstract

Objective: To identify cell types in the male and female reproductive systems at risk of SARS-CoV-2 infection due to expression of host genes and proteins used by the virus for cell entry.

Design: Descriptive analysis of transcriptomic and proteomic data.

Setting: Academic research department and clinical diagnostic laboratory.

Patients/Animals: None. Focused on previously generated gene and protein expression data.

Intervention(s): None.

Main Outcome Measure(s): Identification of cell types co-expressing the key *ACE2* and *TMPRSS2* genes and proteins, as well as other candidates potentially involved in SARS-CoV-2 cell entry.

Results: Based on single cell RNA sequencing data, co-expression of *ACE2* and *TMPRSS2* was not detected in testicular cells, including sperm. A subpopulation of oocytes in non-human primate ovarian tissue were found to express *ACE2* and *TMPRSS2*, but co-expression was not observed in ovarian somatic cells. RNA expression of *TMPRSS2* in 18 samples of human cumulus cells was shown to be low or absent. There was general agreement between publicly available bulk RNA and protein datasets in terms of *ACE2* and *TMPRSS2* expression patterns in testis, ovary, endometrial and placental cells.

Conclusion: These analyses suggest that SARS-CoV-2 infection is unlikely to have long-term effects on male and female reproductive function. While the results cannot be considered definitive they imply that procedures in which oocytes are collected and fertilized *in vitro* are associated with very little risk of viral transmission from gametes to embryos and may indeed have the potential to minimize exposure of susceptible reproductive cell types to infection when compared to natural conception.

Introduction

As of April 26, 2020, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is known to have infected 2 804 796 persons globally and 193 710 have had coronavirus disease 2019 (COVID-19) at the time of death (1). Investigations into the molecular details of SARS-CoV-2 infection have been rapidly initiated and several key facts are already known. Viral entry requires the binding of SARS-CoV-2 spike (S) glycoprotein to the host receptor angiotensin-converting enzyme 2 (*ACE2*; 2–4). Host proteases such as transmembrane serine protease 2 (*TMPRSS2*) are then needed to cleave the viral S protein to induce a conformational change to S that allows for permanent fusion of the viral and host cell membranes (2,4). The importance of *TMPRSS2* has been confirmed and is evident in studies showing that its inhibition blocks SARS-CoV-2 entry and spread in targeted lung cells (4).

TMPRSS2 is more broadly expressed in human tissues than *ACE2* indicating that *ACE2* may be one of the principal determinants of whether a given cell type is susceptible to viral infection (5). Single cell RNA sequencing (scRNAseq) in human and non-human primate respiratory tissues have shown co-expression of *ACE2* and *TMPRSS2* in pneumocytes in the lungs and goblet secretory cells in the nose (6), indicating that these cell types may serve as foci for infection and potentially explaining the range of respiratory symptoms associated with COVID-19. Co-expression has also been reported in other tissue types, including the ileum, heart and kidney for which there are also established COVID-19 symptoms (5,7–11). Detection of the virus in blood, feces and possibly urine suggests that the impact of SARS-CoV-2 on cardiac, enteric and renal function may be the direct result of infection of cells within these tissues rather than being secondary to acute respiratory distress syndrome (12–14).

Given that clinical features of COVID-19 appear to be largely determined by the tissues with co-expression of *ACE2* and *TMPRSS2* in their constituent cells, it is conceivable that viral infection could have an impact on reproductive function if cells of the male and/or female reproductive systems also express these genes. Expression of *ACE2* and *TMPRSS2* has been shown in testicular, endometrial and placental cells with varying interpretations (7,11). To our knowledge, expression of SARS-CoV-2 host entry proteins has not been assessed in human or non-human primate ovaries; specifically, the outer ovarian cortex which houses the germ cells (PubMed/MEDLINE). Reproductive indices are not typically assessed in the intensive care unit (ICU) setting and any effects of COVID-19 on fertility may not be readily apparent until epidemiological data becomes available. Therefore, we used publicly available scRNAseq data, our own unpublished transcriptomic data and publicly available bulk RNA and proteomics data in order to assess the expression patterns of viral host entry proteins in reproductive cell populations. We also considered expression of the receptor basigin (*BSG/CD147*), as limited data suggests it may have the capacity to mediate viral entry (15), and also the cysteine protease cathepsins L (CatL; gene symbol: *CTSL*), which potentially cleaves the viral S protein (2,4). Cathepsin L does not appear to be essential for viral infection, but it is possible that residual infection in cells treated with camostat mesylate, a *TMPRSS2* inhibitor, may reflect S protein priming by CatL (4,16).

Since the expression patterns of *ACE2* and *TMPRSS2* in tissues studied to date is cell-type specific (5), scRNAseq analyses, which reveal co-expression of these genes within individual cells, are expected to be particularly valuable to uncovering the etiology of disease. Cell types in the reproductive system have variable lifespans and cells that live for shorter periods may pose less of a threat to an individual's long-term reproductive potential if infected. For instance, somatic spermatogonial stem cells (SSC) in the testes are continuously self-renewing cells that propagate throughout male reproductive life, while differentiated spermatocytes are cleared from the reproductive system within approximately 60 days (17). In the ovary, a cohort of ~2,000,000 oocytes exists at birth, a diminishing subset of which will persist throughout female reproductive life, a cellular lifespan measured in decades. No new oocytes are produced. In contrast, cells of the decidua (outermost lining of the endometrium) differentiate and shed in successive menstrual cycles (regularly every 21-40 days). Cell-type specific expression patterns of genes that produce viral host entry proteins, and identification of potential loci of infection within the reproductive system, are therefore necessary in order to predict whether SARS-CoV-2 is likely to have any impact on fertility. Although not definitive, findings from studies such as this may also help to predict the likelihood of human embryos becoming infected with SARS-CoV-2. In the midst of the pandemic, fertility treatments in many countries have been curtailed, cancelled or even banned outright. Given the declining chances of treatment success with advancing female age, there is an urgent need to reestablish such treatments as soon as this can be safely accomplished. Understanding whether SARS-CoV-2 has the capacity to infect gametes and embryos produced *in vitro* is of great importance when considering the risks of natural and assisted reproduction during the COVID-19 pandemic.

Methods

Pre-processing and data download

Protocols for amplification, sequencing and alignment are previously described in the respective studies (18,19). In brief, Guo et al (2018) generated scRNAseq data in testicular tissue from three healthy males of peak reproductive age (17, 24 and 25 years old). Wang et al (2019) generated scRNAseq data in ovarian tissue from three young (4-5 years old) cynomolgus monkeys, a genus of *Macaca*. The combined expression matrices were downloaded from GSE112013 (testicular) and GSE130664 (ovarian).

Clustering analysis and cell type annotation using Seurat

The combined expression matrices from each study were converted to a sparse matrix using the Matrix package in R and written out using writeMM. The sparse matrix was then loaded into R using the Read10X function in the Seurat V3.0 package (<https://satijalab.org/seurat/9>) and an object was created using the CreateSeuratObject function. Cells from each experiment were kept only if >500 genes were expressed, and <20% of reads mapped to the mitochondrial genome. The top 5000 highly variable

genes were identified using the FindVariableGenes function. Data was then normalized and scaled using the SCTransform function by regressing out the percentage of mitochondrial gene content. Principal components analysis was then performed on the scaled data using the RunPCA function. To visualize “dimensionality” of the dataset, we plotted the principle components using an elbow plot. This indicated that PC 1-13 explained the majority of the variance in both datasets. Downstream tSNE and clustering analyses were therefore performed on PC 1-13. The FindNeighbors and FindClusters function were used for the clustering analyses. In brief, FindNeighbors constructs a K-nearest neighbor (KNN) graph and then determines the Jaccard similarity between any two cells. We manually assigned cell types to clusters with well-known markers from the original studies (18,19). Expression of an ‘average’ cell in each cluster was calculated based on the normalized, scaled and centered data. The average expression reflects the mean expression in each cluster compared to all other cells.

Cumulus cell RNA extraction and RNA sequencing

Cumulus cells associated with 18 independent oocytes from nine patients undergoing routine assisted reproductive treatment were collected (ethical approval: WIRB 20031397). Total RNA was extracted from each sample using the RNeasy Micro Kit (Qiagen) according to the manufacturer’s instructions and samples were eluted in 14µl of RNase-free water. cDNA was synthesized and amplified using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (TaKaRa Bio) according to the standard workflow. Sequencing library preparation was performed using the Nextera XT DNA Library Prep Kit (Illumina), according to manufacturer’s instructions. Briefly, purified double-stranded cDNA samples were fragmented, indexed, cleaned, normalized and pooled. The final denatured and diluted pooled libraries were sequenced on a NextSeq 550 (Illumina) instrument with 75 cycle paired-end synthesis, using a NextSeq 150 Cycle High Output Kit. An average of >50,000,000 sequencing reads were obtained per sample. Relative gene expression levels were given in Transcripts Per Million (TPM), calculated by dividing the number of mapped reads corresponding to an individual gene by the length of the gene in kilobases (this gives the number of reads per kilobase - RPK). All the RPK values for a sample were summed together and the resulting number was divided by 1,000,000. RPK values for individual genes were divided by this number in order to give TPM.

Protein and bulk RNA expression

Protein and bulk RNA expression data was retrieved from the Human Protein Atlas (proteinallas.org) and the Human Proteome Map (humanproteomemap.org). Please note that these publicly available bulk RNA and protein data sets were generated independently from the previously described scRNAseq datasets. RNA expression was categorized into the groups used in Figure 5 by the following schema, consistent with that used by Hikmet *et al* (20). Low expression = 1-10% of the maximum value, medium expression = 11-50% of the maximum value, and high expression = 51-100% of the maximum value. Categorizations for protein expression correspond to the categorizations given by the Human Protein Atlas and the Human Proteome Map.

Results

Viral entry RNA expression

Published scRNAseq datasets in human testicular (18) and non-human primate ovarian (19) tissue were used to assess the cell-type specific expression pattern of SARS-CoV-2 entry receptor *ACE2* and entry-associated protease *TMPRSS2* and their co-expression. The expression and co-expression of alternative receptor *BSG* and protease *CTSL* were also investigated. Additionally, a novel RNAseq data set was available for human cumulus cells and provided supplementary data on gene expression levels in that cell type.

Human testicular cells

From scRNAseq datasets, we annotated eleven previously identified cell types of the testes including somatic niche cells (Leydig cells, endothelial cells, myoid cells and macrophages) and germ cells (differentiating spermatogonia, early primary spermatocytes, late primary spermatocytes, round spermatids, elongated spermatids, and sperm; **Figure 1A-B**). We were also able to annotate a cluster of spermatogonial stem cells (**Figure 1A-B**). Sertoli cells are underrepresented owing to cell size filtering in the original study (18) and therefore did not cluster separately.

Consistent with independent analyses (21), we found *ACE2* expression in myoid cells (average expression = 2.2, 5.3% of cells with mRNA transcripts), spermatogonial stem cells (average expression = 1.0, 4.8% of cells with mRNA transcripts) and leydig cells (average expression = 1.0, 2.3% of cells with mRNA transcripts; **Figure 1C**). *TMPRSS2* expression was detected predominantly in elongated spermatids (average expression = 2.5, 15.0% of cells with mRNA transcripts) and to a lesser extent in spermatogonial stem cells (average expression = 1.5, 9.3% of cells with mRNA transcripts; **Figure 1C**). Despite a small proportion of spermatogonial stem cells expressing *ACE2* and *TMPRSS2*, cells co-expressing both genes were extremely rare (0.05% of cells; **Figure 1C**). We did not observe a correlation between *ACE2* and *TMPRSS2* expression in any of the annotated cell types or testicular cells broadly (Pearson correlation value = -0.01). Alternative receptors and proteases may mediate viral entry in these cells. *BSG* was more broadly expressed across testicular cell types than *ACE2* and was co-expressed with *CTSL* in early and late primary spermatocytes (78.7% and 90.8% of cells with mRNA transcripts, respectively). Lower levels of *BSG* and *CTSL* co-expression were also detected in Leydig cells, myoid cells, endothelial cells and differentiating spermatogonia (**Figure 1C**).

Non-human primate ovarian cells

To our knowledge, scRNA data in human ovaries (specifically, the outer ovarian cortex) is not yet readily available, however data in non-human primate ovaries has been made

public. Consistent with previous analyses (19), we identified seven ovarian cell types including oocytes and six somatic cell types (stromal cells, granulosa cells, smooth muscle cells, natural killer cells, macrophages and endothelial cells; **Figure 2A-B**). *BSG* was expressed in somatic ovarian cells, while *ACE2* and *TMPRSS2* expression was restricted to germ cells (**Figure 2C**). A moderate correlation between *ACE2* and *TMPRSS2* in ovarian cells (Pearson correlation value = 0.51) was driven entirely by expression in oocytes (**Figure 2D**). As in the original study (19), a principle component analysis within the oocyte cluster was used to identify four subtypes of oocytes with distinct gene expression profiles (**Figure 3A**). These four clusters represented sequential developmental stages of folliculogenesis (i.e. oocytes within primordial, primary, secondary and antral follicles) as is shown by the distribution of markers known to be involved in follicular development (**Figure 3B**). The degree of *ACE2* and *TMPRSS2* co-expression increased with oocyte maturity: primordial (17.0% of cells, Pearson correlation value = 0.21), primary (39.0%, Pearson correlation value = 0.21), secondary (25.0%, Pearson correlation value = 0.13) and antral (62.1%, Pearson correlation value = 0.37; **Figure 3B-C**). Unlike lung tissue (5), *ACE2* (average expression = 2.3, 100% of cells) is more broadly expressed in oocytes than *TMPRSS2* (average expression = 2.2, 37.0% of cells) suggesting that the latter may be a potential limiting factor for infection of the female gamete (**Figure 3B**). Expression of the alternative protease cathepsin L (*CTSL*) could not be interrogated as it was not captured in the original study (absent from combined expression matrix; 19). *BSG* transcripts were detected in 100% of primordial oocytes with diminishing expression across maturing oocytes (**Figure 3B**).

Human cumulus cells

Analysis of the entire transcriptome of 18 human cumulus cell samples using RNAseq, allowed evaluation of the expression of *ACE2*, *TMPRSS2*, *CTSL* and *BSG*. Of note, each sample was comprised of all of the cumulus cells associated with a single oocyte and consequently it was not possible to determine whether genes were co-expressed in individual cells. Consistent with the non-human primate scRNAseq results, *ACE2* was shown to have wide expression. Transcripts from this gene were detectable in 18/18 samples and a 'transcripts per million (tpm)' score of 24.03 was recorded. Also concordant with the primate results, expression of *TMPRSS2* was found to be very low in human cumulus cells. Transcripts were completely undetectable in 15/18 samples and only expressed at a low levels in the other three samples (averaging 0.13 tpm). In contrast, *BSG* and *CTSL* transcripts were detectable in all samples (18/18) and had expression levels of 374.19 tpm and 105.45 tpm, respectively.

Viral entry protein expression

Proteomic expression of *ACE2* (**Figure 4A**), *TMPRSS2* (**Figure 4B**), *BSG* (**Figure 4C**) and *CTSL* (**Figure 4D**) was queried from two publicly available databases, the Human Protein Atlas (HPA) and the Human Proteome Map. The HPA measures protein expression by immunohistochemistry (IHC), while the Human Proteome Map uses a mass spectrometry-based profiling platform. Protein expression in ovarian, testicular, endometrial and placental tissue was correlated with bulk cell RNA expression data

from four publicly available datasets: the HPA and GTEx datasets, which both use RNA-seq, the FANTOM5 dataset, which uses cap analysis gene expression (CAGE) and the HPA consensus dataset, a consensus dataset of the other three RNA datasets. The low sequencing depth of scRNAseq datasets means that many cells will have no expression by chance, therefore bulk RNAseq datasets, sequenced at higher depths, can serve as a useful point of comparison.

There was generally agreement between the expression data from the two protein datasets and the RNA datasets. In ovarian tissue, the mass spectrometry found high levels of *ACE2* protein expression in ovarian cells while the IHC method did not detect the protein in either follicular or stromal ovarian cells (**Figure 4A**). In ovarian cells, *TMPRSS2* was not detected by either protein profiling platform, which was supported by the external RNA datasets (**Figure 4B**). These results are consistent with our internal RNAseq results in human cumulus cells showing broad expression of *ACE2* but not *TMPRSS2*. *BSG* expression on a protein level was detected by mass spectrometry but not IHC in ovarian cells, which was supported by RNA expression data (**Figure 4C**). *CTSL* showed moderate and high RNA expression but no protein expression (**Figure 4D**).

In testicular tissue, *ACE2* (**Figure 4A**) and *BSG* (**Figure 4C**) both showed medium or high expression across all platforms. The underrepresentation of Sertoli cells in the scRNAseq data may account for some of the discrepancy in the *ACE2* expression profile between these two datasets. It is also worth noting that baseline *ACE2* expression is low, even in tissues known to be infected (5), and therefore given the low sequencing depth of scRNAseq platforms, detection of mRNA transcripts in few cells may still indicate functional activity (as is suggested by protein expression data). *TMPRSS2* was only found to be expressed at a low level in the RNA datasets and was not detectable by either protein platform (**Figure 4B**). *CTSL* was detected by mass spectrometry and all RNA analyses but not IHC (**Figure 4D**). In endometrial tissues, only *BSG* (**Figure 4C**) was detected at a protein level. In placental tissues, there was both RNA and protein evidence of robust expression of *BSG* (**Figure 4C**) and *CTSL* (**Figure 4D**), but not *ACE2* (**Figure 4A**) or *TMPRSS2* (**Figure 4B**).

Human reproductive cell lines

In addition to primary tissues, human cell lines may serve as valuable models for studying viral replication and effects on different tissue types. The Cell Atlas provides RNAseq expression data for 64 different human cell lines (22). Of those 64, three were of particular interest to this study: EFO-21, a metastatic ovarian serous cystadenocarcinoma cell line; AN3-Ca, an endometrial carcinoma line; and BEWO, a metastatic choriocarcinoma line. The disease states of the cell lines may perturb expression and thus undermine conclusions about expression in healthy human tissues. Even so, expression of SARS-CoV-2 viral entry proteins in the cell lines would establish whether or not they were susceptible to viral infection and appropriate models for further study. RNAseq expression levels for *ACE2*, *BSG*, *TMPRSS2* and *CTSL* was queried for all three lines. Expression data given by the Cell Atlas was in the form of normalized

expression levels (NX), where the Cell Atlas takes an NX of 1 as the lower threshold for expression.

Of the three cell lines, only BEWO had detectable expression of all five proteins of interest, with low levels of *ACE2* and *TMPRSS2* expression (NX = 2.4 and 1.3, respectively) and higher levels expressed of *BSG* and *CTSL* (NX = 87.4 and 15.4, respectively). EFO-21 and AN3-Ca both showed *BSG* and *CTSL* expression (EFO-21: NX = 61.7 and 96, respectively; AN3-Ca: NX = 92.3 and 4.3, respectively), but not *ACE2* or *TMPRSS2*. As such, the BEWO cell line seems a good potential model system for investigating any impact(s) of SARS-CoV-2 infection on placental tissue. EFO-21 and AN3-Ca could be viable model lines if *CTSL* is determined capable of functionally replacing *TMPRSS2* for viral S protein priming. These lines may serve as good candidates for investigating the potential functional interchangeability of *TMPRSS2* and *CTSL*.

Discussion

It is not yet clear what effects, if any, the COVID-19 pandemic will have on male and female fertility. Therefore, we decided to look at the expression patterns of known viral host entry proteins to gain an insight into the possible biological consequences of SARS-CoV-2 infection on reproduction. Our results from scRNAseq data suggest that sperm cells may not be at increased risk of *ACE2* and *TMPRSS2* mediated viral entry and spread, given the lack of co-expression in any testicular cell type, although we were unable to interrogate Sertoli cells. Separately, bulk RNAseq and protein platforms indicated moderate *ACE2* expression in testicular tissue, while expression of *TMPRSS2* was found to be low or undetectable. Alternative receptor *BSG* and protease cathepsin L were broadly expressed in testicular cell clusters as is true in other tissue types studied to date (7). However, it is not yet empirically determined whether cathepsin L can functionally replace the key role of *TMPRSS2* in viral S protein priming(4,5). Our findings are consistent with a recent study, showing an absence of SARS-CoV-2 RNA in semen and testicular biopsy specimens from infected men (23). Taken together, these results suggest that long-term effects of SARS-CoV-2 on male reproductive function are unlikely and that the risk of infection of sperm and subsequent transmission to embryos during fertilization is low.

At present, relatively little scRNAseq data is available from human ovarian tissue. Fan *et al* (2019) sequenced the inner cortex of ovaries from women undergoing ovarian tissue cryopreservation (24), however the inner cortex is largely devoid of germ cells. More recently, Wagner *et al* (2020) performed scRNAseq of the outer cortex, containing both germinal and somatic cell types, producing data that could potentially serve as an external validation dataset if expression matrices were made available (25). As a result of the paucity of scRNAseq data from the human ovary, this study and others have used information from closely related non-human primates as a proxy for expression in our own species (19). Naturally, such analyses have limitations and, while potentially indicative, cannot be considered conclusive.

In the current study, based on scRNAseq data, a subpopulation of oocytes in non-human primate ovarian tissue was found to co-express *ACE2* and *TMPRSS2*, while limited co-expression was observed in other oocytes and was absent from ovarian somatic cells. Co-expression of *ACE2* and *TMPRSS2* appeared to increase as follicles progressed through development: primordial follicles had minimal co-expression, while 62% of antral follicles had co-expression of *ACE2* and *TMPRSS2* (Pearson correlation value = 0.37). These results are relatively reassuring concerning the long-term effects of SARS-CoV-2 on female reproduction. Considering that the cells predicted to have the greatest susceptibility to infection are oocytes in antral follicles, which are either ovulated or atrophy within several days of appearance each cycle, a sustained impact on the female gamete seems unlikely. Additionally, it is important to note that oocytes within antral follicles are entirely surrounded within a mass of cumulus cells (a specialized class of granulosa cell). The non-human primate scRNAseq data suggests that such cells do not co-express *ACE2* and *TMPRSS2* and may therefore provide a physical barrier to infection. Our transcriptomic data from human cumulus cells indicates that *ACE2* is expressed in this cell type. Although a relatively small sample was assessed (cumulus cells from 18 oocytes in 9 women), these findings are concordant with previous research (26). However, transcripts of *TMPRSS2* are either entirely absent or exist at very low levels in cumulus cells and therefore a low risk of infection is predicted.

The documented lack of SARS-CoV (a coronavirus identified in 2003 that also uses *ACE2* for viral entry; 4) replication in ovarian tissue, as determined via pathology and immunohistochemical analyses, potentially supports our hypothesis that ovarian somatic tissues have little susceptibility to SARS-CoV-2 infection (12). As of yet, no autopsy reports in persons with laboratory confirmed COVID-19 at time of death have assessed SARS-Cov-2 replication in the reproductive system. Rather, infected tissues known to cause death in patients such as the lungs and heart have been the primary focus (27,28). Moreover, current pathology data is largely based upon samples from post-menopausal women and therefore cannot address the potential activity of SARS-CoV-2 in mature oocytes (since post-menopausal women have none). Further experimental data in ovarian tissue from women of reproductive age with COVID-19 is needed to confirm whether the female germ cells can be infected. It is also worth noting that detection of *ACE2* and *TMPRSS2* mRNA transcripts in mature oocytes may not reflect active production of the associated proteins. Translation of mRNA gradually decreases during oocyte maturation, allowing a large pool of RNA transcripts to accumulate, necessary to sustain the embryo prior to genome activation (29).

Given that the impact of SARS-CoV-2 on preimplantation development, implantation and early pregnancy is largely unknown and considering that there is a theoretical possibility of viral infection of mature oocytes (due to co-expression of *ACE2* and *TMPRSS2*) there may be concerns over the initiation of pregnancy during the COVID-19 pandemic. However, the data presented here suggests that procedures in which cumulus cell-enclosed oocytes are collected and fertilized outside the female reproductive tract (e.g. *in vitro* fertilization) are likely to pose very little risk and may have the potential to reduce or entirely eliminate exposure of susceptible cell types to

infection. The risk of infection for primordial follicles (the ovarian reserve) appears to be low (*ACE2* and *TMPRSS2* expression correlation value = 0.21), indicating that any possible effects of COVID-19 on fertility may be transient.

So far, published data also provides some reassurance concerning aspects of reproductive risk occurring at later stages, after initiation of a pregnancy. Publicly available developmental scRNAseq databases reveal no co-expression of *ACE2* and *TMPRSS2* in fetal tissues including liver, thymus, skin, bone marrow and yolk sac (6). Expression of *ACE2* has been documented in the placenta, but co-expression with *TMPRSS2* has not been reported, suggesting that placental cells may act as a barrier to vertical SARS-CoV-2 transmission (7). These conclusions are supported by the lack of *ACE2* and *TMPRSS2* protein expression in endometrial and placental tissue (HPA/IHC data sets) described in this study. The limited clinical data available has also indicated a lack of vertical transmission in newborns of pregnant women who had confirmed COVID-19 in the third trimester (30). Additionally, obstetric outcomes for babies born to mothers infected with SARS-CoV-2 have not shown any impact of COVID-19 on pregnancy (31).

Conclusion

The results presented here can only be considered indicative and require verification. Our results cannot rule out the possibility that proteases other than *TMPRSS2* (including *CTSL*), may facilitate viral entry in some cells such as primordial oocytes (100% of which have detected *ACE2* and *BSG* transcripts). Nonetheless, they add to a growing body of evidence suggesting that most aspects of reproduction during the COVID-19 pandemic are unlikely to be associated with increased risks of clinical complications. The amount of data currently available is small, but thus far there is no indication of an increased incidence of severe disease amongst pregnant women and obstetric outcomes for babies born to mothers infected with SARS-CoV-2 appear to be within the normal range (30). Furthermore, viral RNA has not been detected in semen or testicular biopsy specimens from infected men (23). The results of the current study suggest that sperm and cumulus-enclosed oocytes are unlikely to be susceptible to infection by SARS-CoV-2 and consequently it should be possible to undertake oocyte collection and fertilization, followed by *in vitro* embryo culture and cryopreservation with minimal risk to the embryos produced. Indeed, considering the lack of knowledge concerning the risk of infection during *in vivo* preimplantation development, implantation and early pregnancy, it is possible that IVF may represent a safer reproductive strategy than natural conception at this time. Given that the success rates of IVF treatments decline at an ever accelerating rate as the age of the female patient increases (falling by approximately 0.3% per month from the mid-thirties, according to data from the Society of Assisted Reproductive Technologies and the Human Fertilisation and Embryology Authority), it is imperative that delays to fertility treatments are minimized. Provided adequate precautions can be instituted to ensure the safety of patients attending fertility clinics, as well as clinical staff, it seems reasonable to consider the reintroduction of IVF treatments (at a minimum cycles involving cryopreservation of embryos) in countries

where the ability of patients to access such procedures has been denied or severely limited.

Acknowledgements

DW received funding from the NIHR Oxford Biomedical Research Centre.

References

1. World Health Organization. Novel Coronavirus (2019-nCoV): Situation Report 97 [Internet]. 2020; Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200416-sitrep-87-covid-19.pdf?sfvrsn=9523115a_2
2. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 2020;367(6485):1444–8.
3. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;579(7798):270–3.
4. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* [Internet] 2020 [cited 2020 Apr 17]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7102627/>
5. Wang X, Dhindsa R, Povysil G, Zoghbi A, Motelow J, Hostyk J, et al. Transcriptional Inhibition of Host Viral Entry Proteins as a Therapeutic Strategy for SARS-CoV-2. 2020 [cited 2020 Apr 17]; Available from: <https://www.preprints.org/manuscript/202003.0360/v1>
6. Sungnak W, Huang N, Bécavin C, Berg M, Queen R, Litvinukova M, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine* 2020;1–7.
7. Sungnak W, Huang N, Bécavin C, Berg M, Network HLB. SARS-CoV-2 Entry Genes Are Most Highly Expressed in Nasal Goblet and Ciliated Cells within Human Airways. *arXiv:200306122 [q-bio]* [Internet] 2020 [cited 2020 Apr 17]; Available from: <http://arxiv.org/abs/2003.06122>
8. Bangalore S, Sharma A, Slotwiner A, Yatskar L, Harari R, Shah B, et al. ST-Segment Elevation in Patients with Covid-19 — A Case Series. *New England Journal of Medicine* 2020;0(0):null.
9. Rabb H. Kidney diseases in the time of COVID-19: major challenges to patient care [Internet]. 2020 [cited 2020 Apr 17]; Available from: <https://www.jci.org/articles/view/138871/pdf>

10. Goyal P, Choi JJ, Pinheiro LC, Schenck EJ, Chen R, Jabri A, et al. Clinical Characteristics of Covid-19 in New York City. *N Engl J Med* 2020;NEJMc2010419.
11. Qi J, Zhou Y, Hua J, Zhang L, Bian J, Liu B, et al. The scRNA-seq expression profiling of the receptor ACE2 and the cellular protease TMPRSS2 reveals human organs susceptible to COVID-19 infection. *bioRxiv* 2020;2020.04.16.045690.
12. Ding Y, He L, Zhang Q, Huang Z, Che X, Hou J, et al. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. *The Journal of Pathology* 2004;203(2):622–30.
13. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* [Internet] 2020 [cited 2020 Apr 18];Available from: <https://jamanetwork.com/journals/jama/fullarticle/2762997>
14. Peng L, Liu J, Xu W, Luo Q, Deng K, Lin B, et al. 2019 Novel Coronavirus can be detected in urine, blood, anal swabs and oropharyngeal swabs samples. *medRxiv* 2020;2020.02.21.20026179.
15. Wang K, Chen W, Zhou Y-S, Lian J-Q, Zhang Z, Du P, et al. SARS-CoV-2 invades host cells via a novel route: CD147-spike protein [Internet]. *Microbiology*; 2020 [cited 2020 Apr 17]. Available from: <http://biorxiv.org/lookup/doi/10.1101/2020.03.14.988345>
16. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications* 2020;11(1):1–12.
17. Griswold MD. Spermatogenesis: The Commitment to Meiosis. *Physiol Rev* 2016;96(1):1–17.
18. Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, et al. The adult human testis transcriptional cell atlas. *Cell Research* 2018;28(12):1141–57.
19. Wang S, Zheng Y, Li J, Yu Y, Zhang W, Song M, et al. Single-Cell Transcriptomic Atlas of Primate Ovarian Aging. *Cell* 2020;180(3):585–600.e19.
20. Hikmet F, Méar L, Uhlén M, Lindskog C. The protein expression profile of ACE2 in human tissues. *bioRxiv* 2020;2020.03.31.016048.
21. Sungnak W, Huang N, Bécavin C, Berg M, Network HLB. SARS-CoV-2 Entry Genes Are Most Highly Expressed in Nasal Goblet and Ciliated Cells within Human Airways. *arXiv:200306122 [q-bio]* [Internet] 2020 [cited 2020 Apr 17];Available from: <http://arxiv.org/abs/2003.06122>
22. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, et al. A subcellular map of the human proteome. *Science* 2017;356(6340).

23. Song C, Wang Y, Li W, Hu B, Chen G, Xia P, et al. Detection of 2019 novel coronavirus in semen and testicular biopsy specimen of COVID-19 patients. medRxiv 2020;2020.03.31.20042333.
24. Fan X, Bialecka M, Moustakas I, Lam E, Torrens-Juaneda V, Borggreven NV, et al. Single-cell reconstruction of follicular remodeling in the human adult ovary. Nat Commun 2019;10(1):3164.
25. Wagner M, Yoshihara M, Douagi I, Damdimopoulos A, Panula S, Petropoulos S, et al. Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. Nature Communications 2020;11(1):1–15.
26. Cavallo IK, Dela Cruz C, Oliveira ML, Del Puerto HL, Dias JA, Lobach VN, et al. Angiotensin-(1-7) in human follicular fluid correlates with oocyte maturation. Hum Reprod 2017;32(6):1318–24.
27. Yao XH, Li TY, He ZC, Ping YF, Liu HW, Yu SC, et al. [A pathological report of three COVID-19 cases by minimally invasive autopsies]. Zhonghua Bing Li Xue Za Zhi 2020;49(0):E009.
28. Li H, Liu L, Zhang D, Xu J, Dai H, Tang N, et al. SARS-CoV-2 and viral sepsis: observations and hypotheses. The Lancet [Internet] 2020 [cited 2020 Apr 24];0(0). Available from: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30920-X/abstract](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30920-X/abstract)
29. Jansova D, Tetkova A, Koncicka M, Kubelka M, Susor A. Localization of RNA and translation in the mammalian oocyte and embryo. PLoS One [Internet] 2018 [cited 2020 Apr 18];13(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5846722/>
30. Chen H, Guo J, Wang C, Luo F, Yu X, Zhang W, et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records. The Lancet 2020;395(10226):809–15.
31. Chen L, Li Q, Zheng D, Jiang H, Wei Y, Zou L, et al. Clinical Characteristics of Pregnant Women with Covid-19 in Wuhan, China. New England Journal of Medicine 2020;0(0):null.

Figure Legends

Figure 1: Clustering of distinct testicular cell populations and transcriptional signatures. Clustering was performed on the normalized and scaled data (**A**) and clusters were manually assigned cell types based on well-known markers (**B**). Legend bars reflect the normalized and log-transformed gene expression values for each cell based on the raw expression matrix (mRNA transcript counts) and therefore contain zero or positive values. The dot plot (**C**) depicts the scaled (Pearson residuals) and centered (mean zero) expression of an average cell in each cluster and therefore contains negative and positive values. The average expression reflects the mean expression in each cluster compared to all other cells. The size of the dot reflects the

percentage of cells with mRNA transcripts detected. VIM = somatic cell marker. CD14 = macrophage marker, VWF = endothelial cell marker, DLK1 = Leydig cell marker, ACTA2 = myoid cell marker. STRA8 = retinoic acid target gene that marks the transition between differentiating spermatogonia and spermatocytes, SYCP3 = meiosis marker, ZPBP = Spermatid structure protein, PRM2 = nuclear condensation/protamine repackaging factor, ID4 = spermatogonial stem cell marker. S'cytes = spermatocytes, S'gonial = spermatogonial, ACE2 = angiotensin-converting enzyme 2, TMPRSS2 = transmembrane serine protease 2, BSG = receptor basigin, CTSL = cysteine protease cathepsin L.

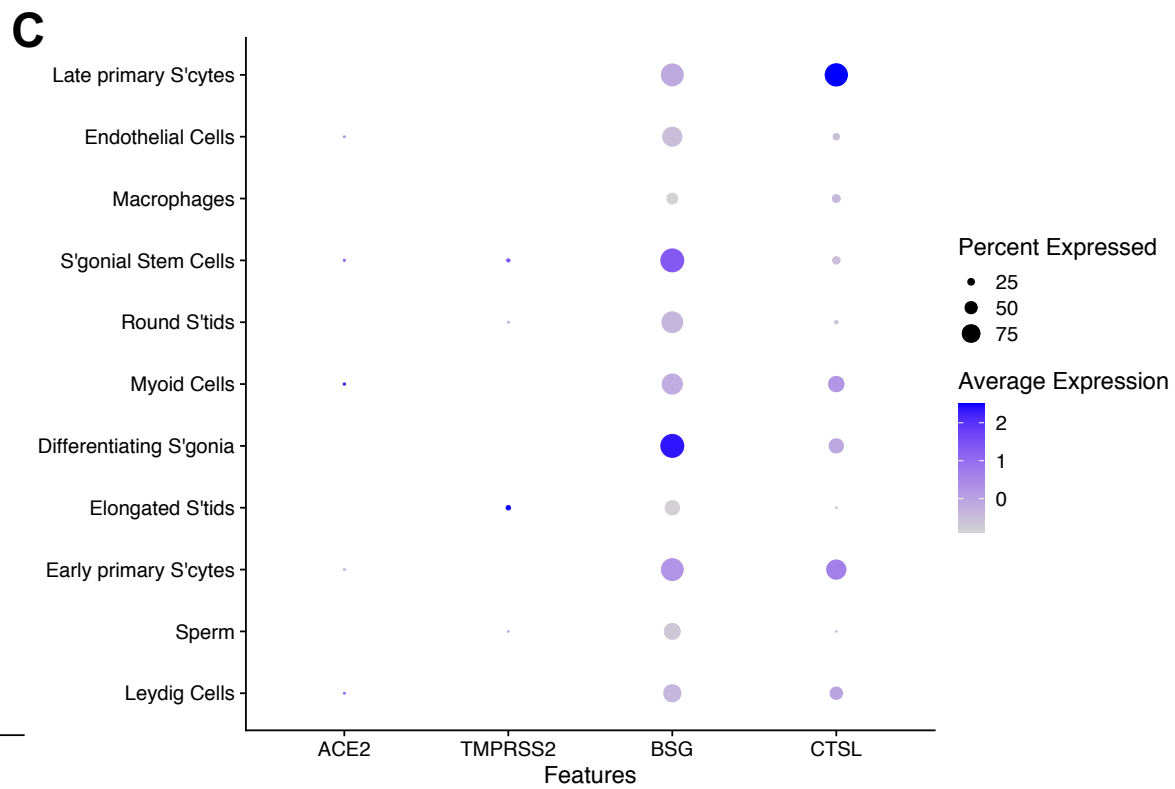
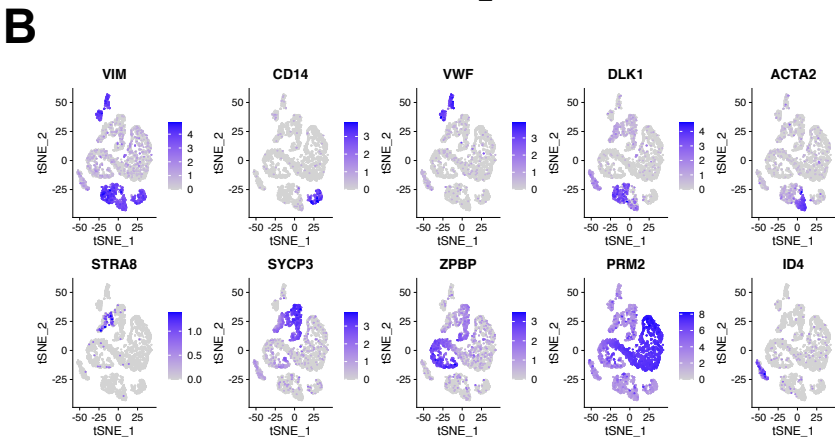
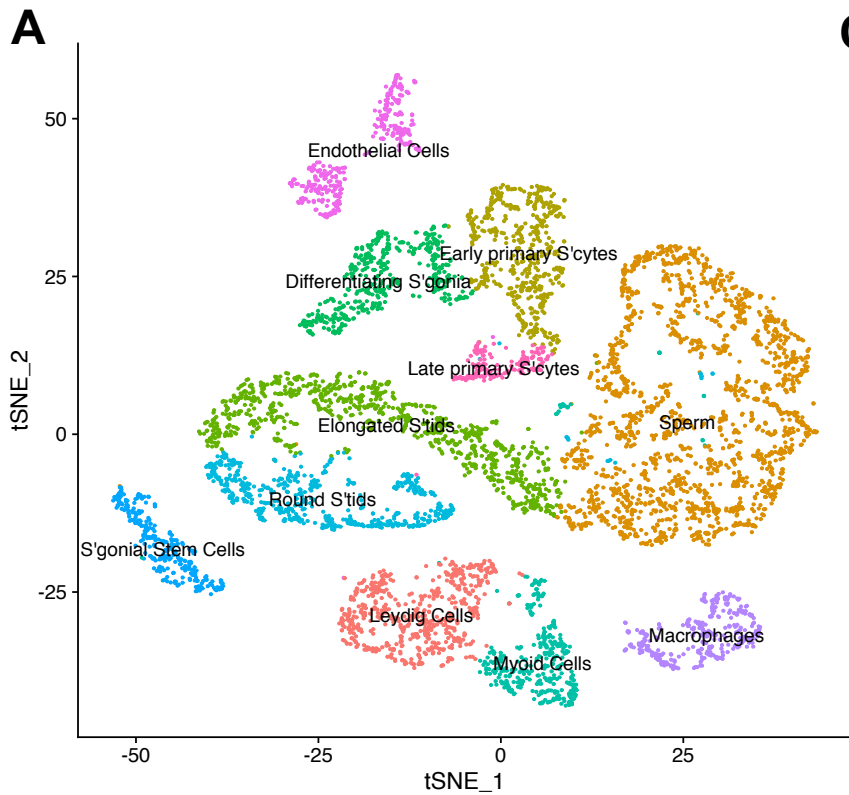
Figure 2: Clustering of distinct ovarian cell populations and transcriptional signatures. Clustering was performed on the normalized and scaled data (A) and clusters were manually assigned cell types based on well-known markers (B). Legend bars reflect the normalized and log-transformed gene expression values for each cell based on the raw expression matrix (mRNA transcript counts) and therefore contain zero or positive values. The dot plot (C) depicts the scaled (Pearson residuals) and centered (mean zero) expression of an average cell in each cluster and therefore contains negative and positive values. The average expression reflects the mean expression in each cluster compared to all other cells. The size of the dot reflects the percentage of cells with mRNA transcripts detected. The Feature Scatter plot (D) shows the Pearson correlation value for *ACE2* and *TMPRSS2* expression in all ovarian cells with color coded cell types.

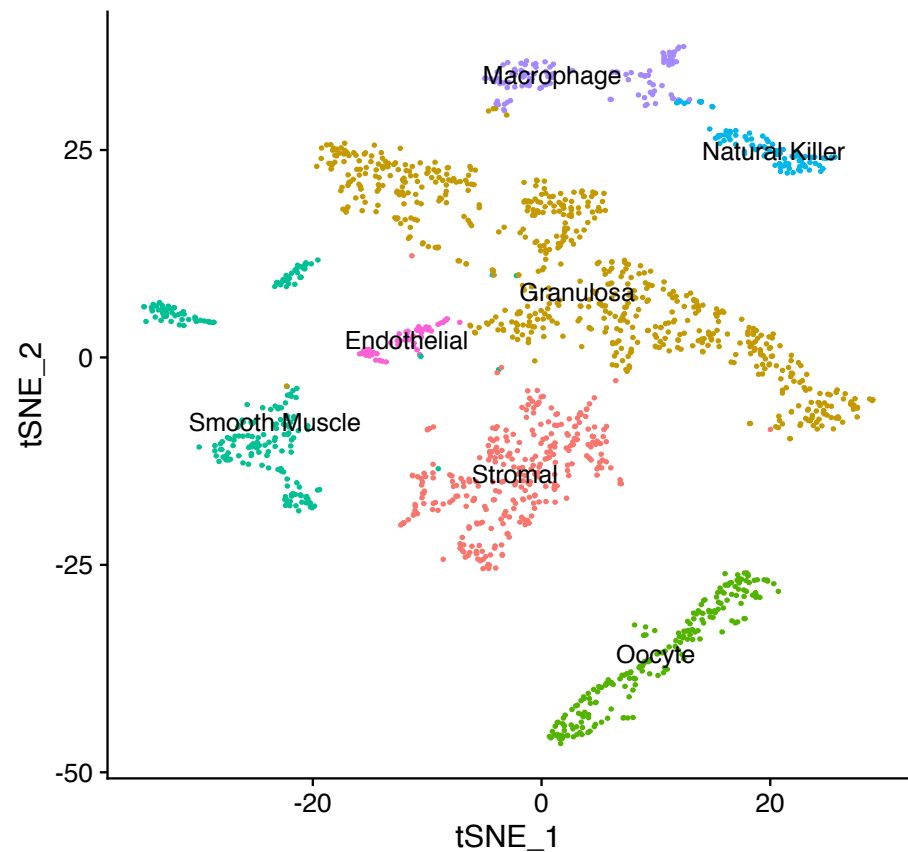
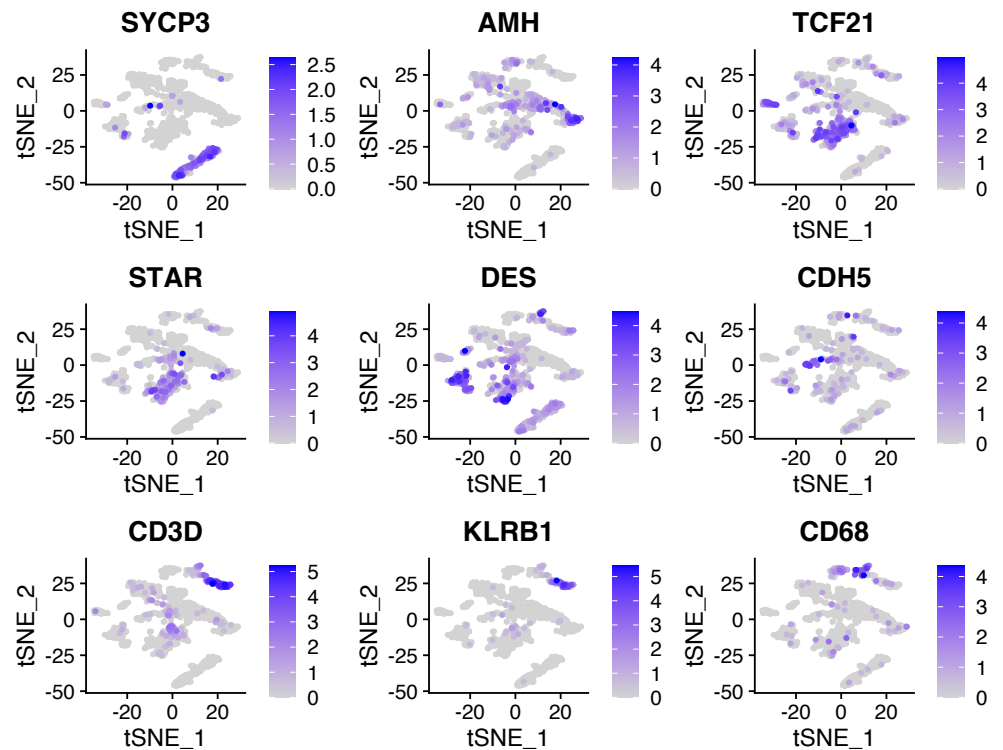
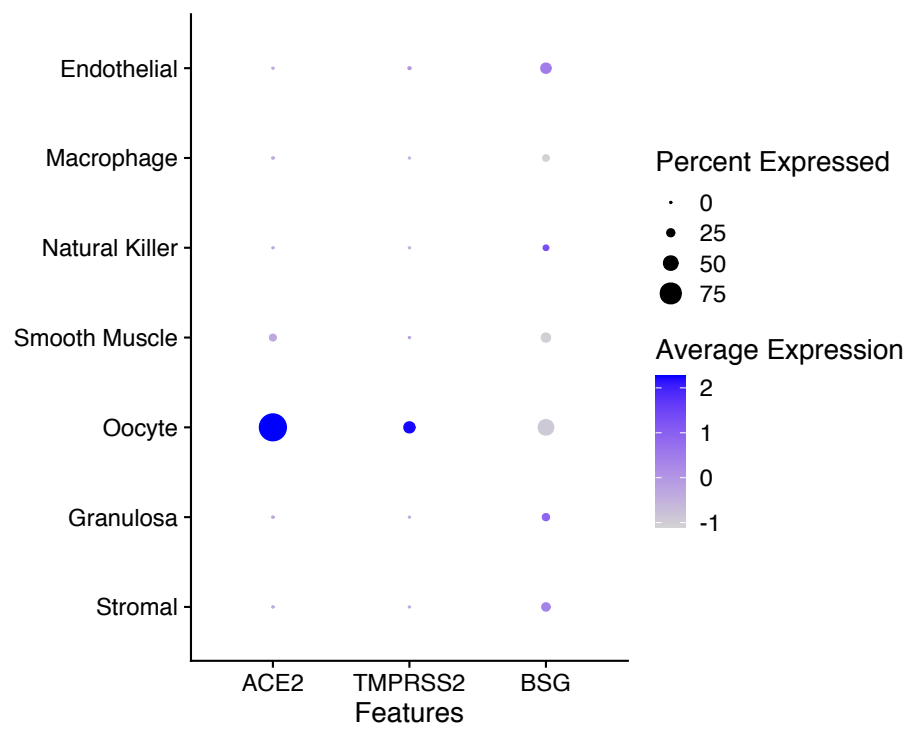
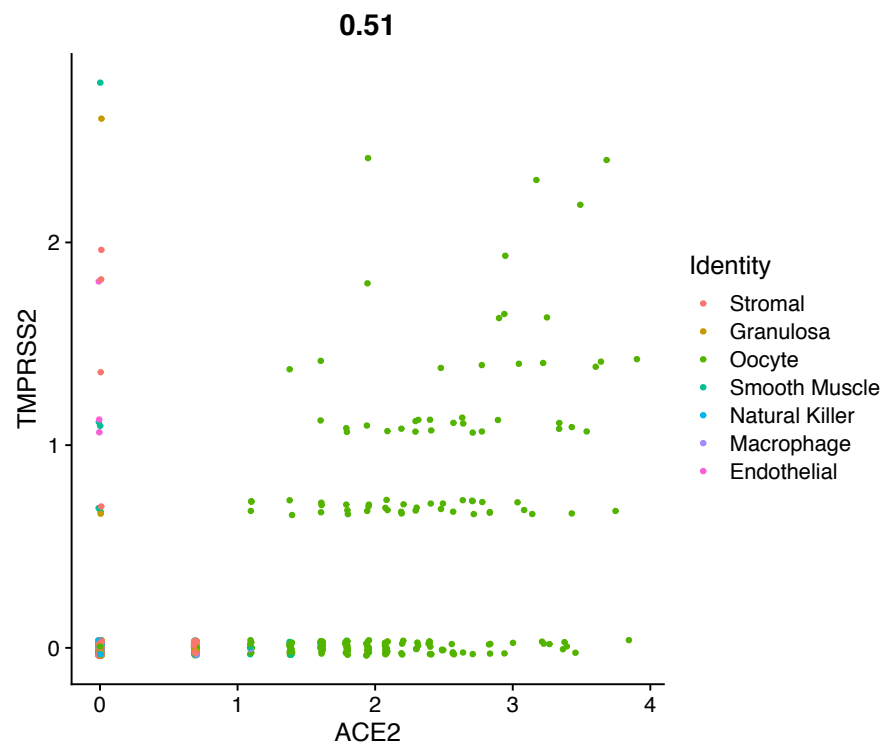
Figure 3: Clustering of distinct stages of folliculogenesis. Principal components analysis (PCA) revealed four stages of folliculogenesis within the oocyte cluster (A). PC1 is plotted against PC2 and clusters were manually assigned cell types using known markers of follicular development. The dot plot (B) depicts the expression distribution of follicular development markers and *TMPRSS2* and *ACE2* expression across oocyte sub-clusters. ATP6 and COX2 are markers of primordial follicles and are progressively downregulated during folliculogenesis. ZP1, GDF9 and BMP15 are promoters of follicular development and are progressively upregulated during folliculogenesis. Average expression represents the scaled (Pearson residuals) and centered (mean zero) expression of an average cell in each cluster and therefore contains negative and positive values. The size of the dot reflects the percentage of cells with detected mRNA transcripts. Primordial = oocytes within primordial follicles (also known as the ovarian reserve). Primary = oocytes within primary follicles. Secondary = oocytes within secondary follicles. Antral = oocytes within antral follicles.

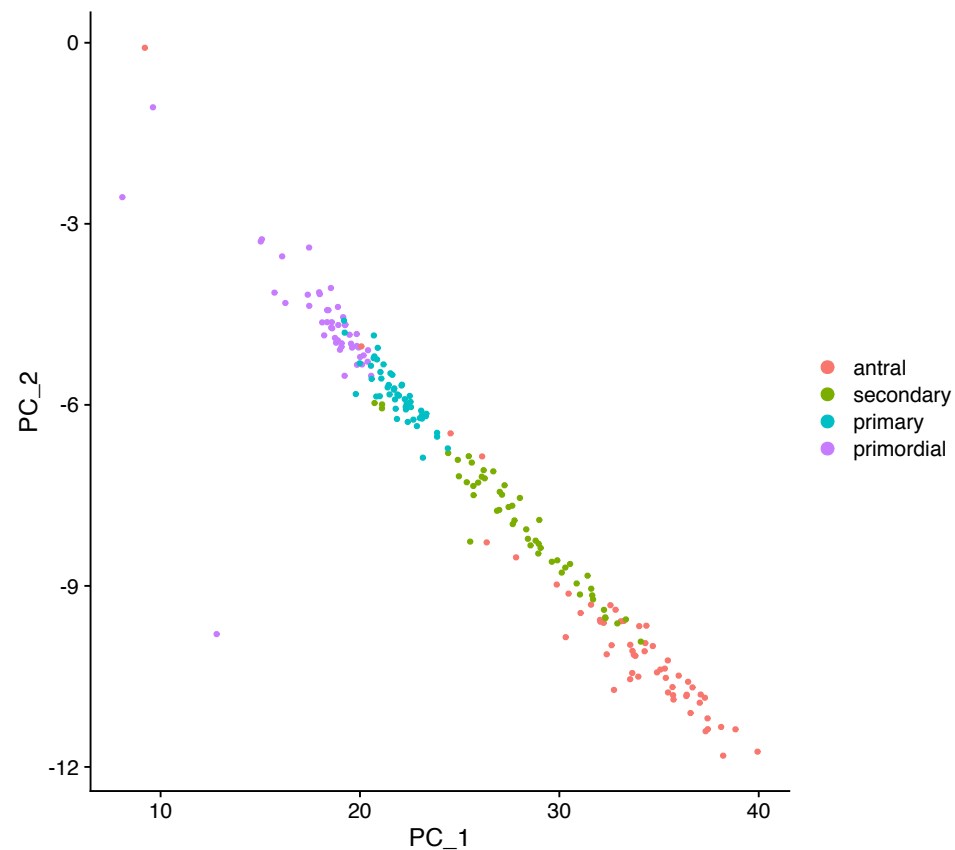
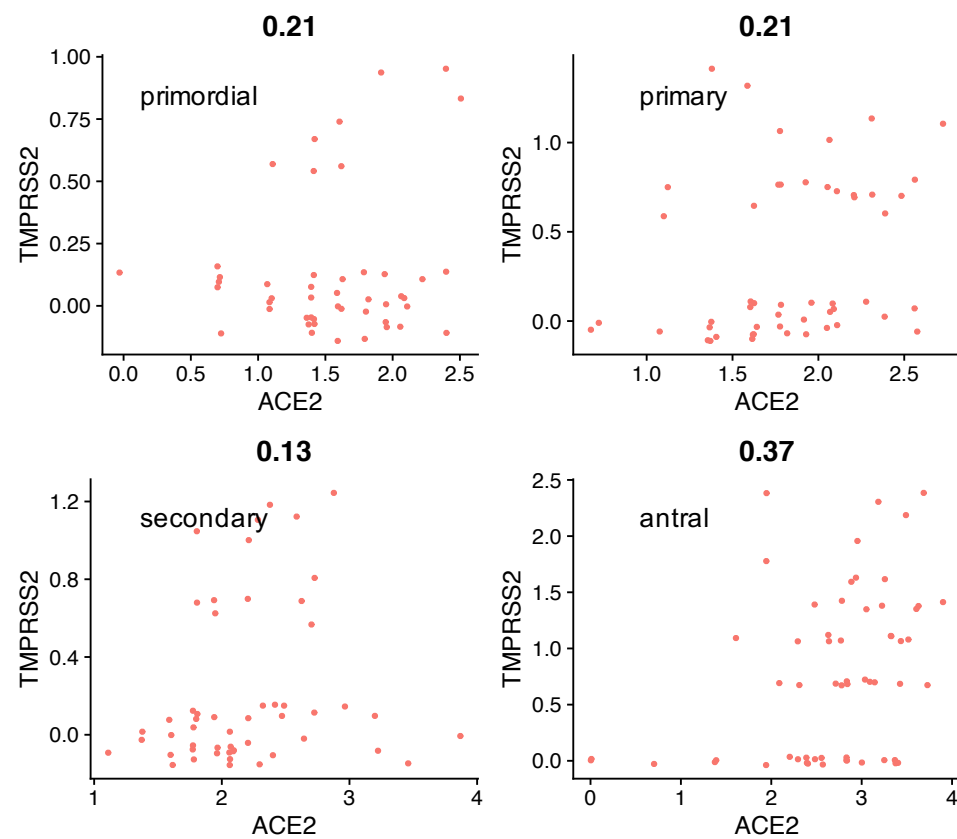
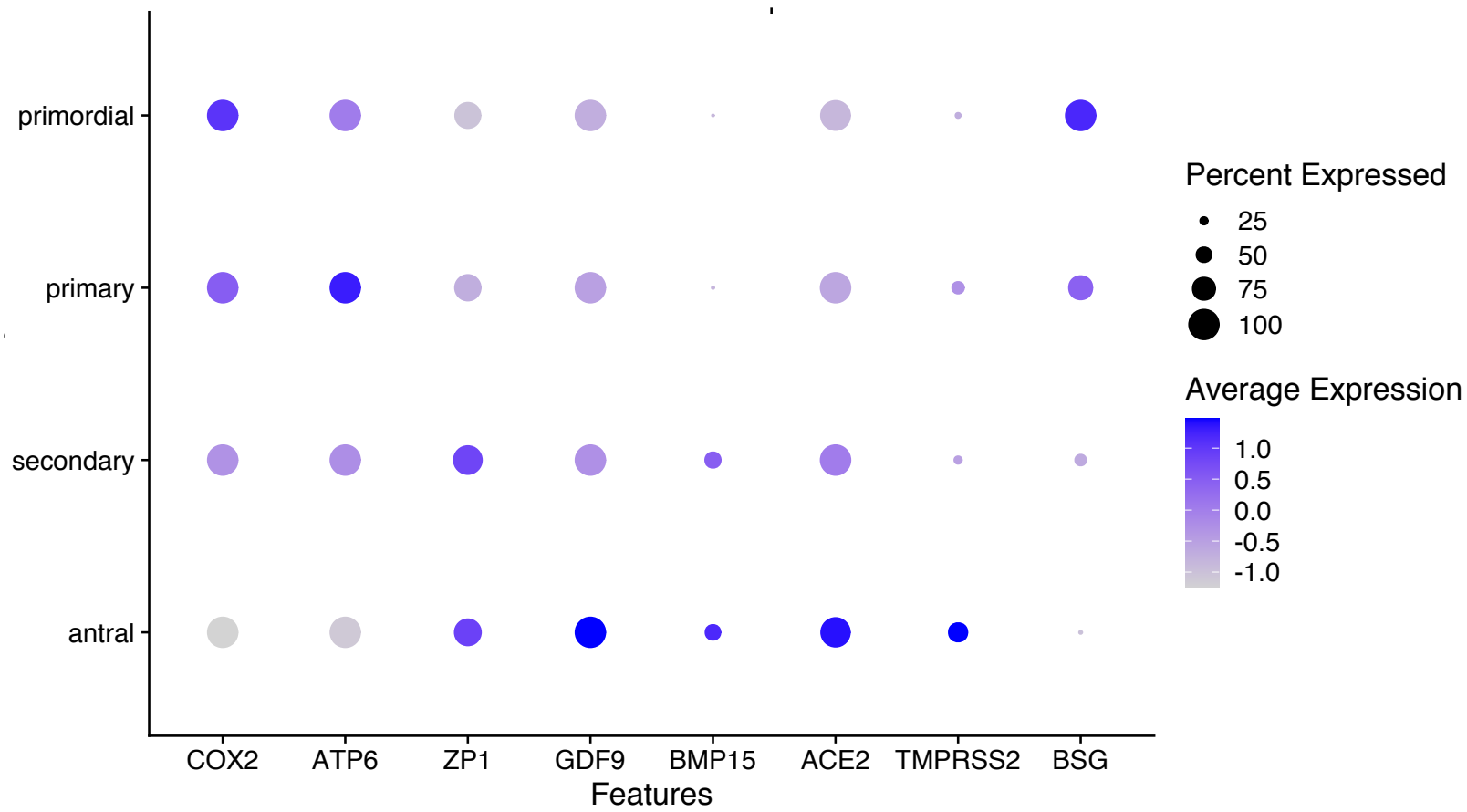
Figure 4. Protein expression of viral host entry proteins in human reproductive tissues. Proteomic expression of *ACE2* (A), *TMPRSS2* (B), *BSG* (C) and *CTSL* (D) was queried from two publicly available databases, the Human Protein Atlas and the Human Proteome Map and correlated with RNA expression data from four different publicly available datasets. Expression was categorized as no expression (white space), low (small circle), medium (medium circle), or high (large circle). The expression of some proteins was not assessed in certain tissues, denoted by 'N/A.' HPA = Human Protein Atlas, GTEx = Genome-based Tissue Expression, CAGE = Cap Analysis Gene

Expression, FANTOM5 = Functional Annotation of the Mammalian, IHC = Immunohistochemistry, Protein MS = Protein Mass Spectrometry.


Journal Pre-proof




A**B****C****D**

A**C****B**


A

	RNA-seq, HPA	RNA-seq, GTEx	RNA CAGE, FANTOM5	RNA normalized, HPA, GTEx, FANTOM5	Protein IHC, HPA	Protein MS, Human Proteome Map
Endometrium					N/A	N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
glandular cells	N/A	N/A	N/A	N/A		N/A
Ovary					N/A	
follicular cells	N/A	N/A	N/A	N/A		N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
Placenta					N/A	
trophoblastic cells	N/A	N/A	N/A	N/A		N/A
decidual cells	N/A	N/A	N/A	N/A		N/A
Testis					N/A	
seminiferous duct cells	N/A	N/A	N/A	N/A		N/A
Leydig cells	N/A	N/A	N/A	N/A		N/A


B

	RNA-seq, HPA	RNA-seq, GTEx	RNA CAGE, FANTOM5	RNA normalized, HPA, GTEx, FANTOM5	Protein IHC, HPA	Protein MS, Human Proteome Map
Endometrium					N/A	N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
glandular cells	N/A	N/A	N/A	N/A		N/A
Ovary					N/A	
follicular cells	N/A	N/A	N/A	N/A		N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
Placenta	N/A				N/A	
trophoblastic cells	N/A	N/A	N/A	N/A		N/A
decidual cells	N/A	N/A	N/A	N/A		N/A
Testis					N/A	
seminiferous duct cells	N/A	N/A	N/A	N/A		N/A
Leydig cells	N/A	N/A	N/A	N/A		N/A

C

	RNA-seq, HPA	RNA-seq, GTEx	RNA CAGE, FANTOM5	RNA normalized, HPA, GTEx, FANTOM5	Protein IHC, HPA	Protein MS, Human Proteome Map
Endometrium					N/A	N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
glandular cells	N/A	N/A	N/A	N/A		N/A
Ovary					N/A	
follicular cells	N/A	N/A	N/A	N/A		N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
Placenta		N/A			N/A	
trophoblastic cells	N/A	N/A	N/A	N/A		N/A
decidual cells	N/A	N/A	N/A	N/A	N/A	N/A
Testis					N/A	
seminiferous duct cells	N/A	N/A	N/A	N/A		N/A
Leydig cells	N/A	N/A	N/A	N/A		N/A

D

	RNA-seq, HPA	RNA-seq, GTEx	RNA CAGE, FANTOM5	RNA normalized, HPA, GTEx, FANTOM5	Protein IHC, HPA	Protein MS, Human Proteome Map
Endometrium					N/A	N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
glandular cells	N/A	N/A	N/A	N/A		N/A
Ovary					N/A	
follicular cells	N/A	N/A	N/A	N/A		N/A
stromal cells	N/A	N/A	N/A	N/A		N/A
Placenta		N/A			N/A	
trophoblastic cells	N/A	N/A	N/A	N/A		N/A
decidual cells	N/A	N/A	N/A	N/A		N/A
Testis					N/A	
seminiferous duct cells	N/A	N/A	N/A	N/A		N/A
Leydig cells	N/A	N/A	N/A	N/A		N/A