Upregulation of Human Endogenous Retroviruses in Bronchoalveolar Lavage Fluid of COVID-19 Patients

Konstantina Kitsou
Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Greece

Anastasia Kotanidou
1st Department of Critical Care & Pulmonary Services, Medical School, National and Kapodistrian University of Athnens, Greece

Dimitrios Paraskevis
Department of Hygiene, Epidemiology and Medical Statistics, Medical School,National and Kapodistrian University of Athens, Greece

Timokratis Karamitros
Unit of Bioinformatics and Applied Genomics, Department of Microbiology, Hellenic Pasteur Institute, Athens, Greece

Aris Katzourakis
Department of Zoology, University of Oxford, United Kingdom

Richard Tedder
Imperial College London, United Kingdom

Tara Hurst
Birmingham City University, Birmingham, United Kingdom

Spyros Sapounas
National Public Health Organization, Athens, Greece

Athanassios Kotsinas
Department of Histology and Embryology, School of Medicine, National and Kapodistrian University of Athens, Greece

Vassilis Gorgoulis
Department of Histology and Embryology, School of Medicine, National and Kapodistrian University of Athens, Greece

Vana Spoulou
Immunobiology and Vaccinology Research Laboratory, First Department of Peadiatrics, School of Medicine, National and Kapodistrian University of Athens, Greece

Sotirios Tsiodras
4th Department of Internal Medicine, Attikon University Hospital, Medical School, National and Kapodistrian University of Athens, Greece

Pagona Lagiou

Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Greece

Gkikas Magiorkinis (gmagi@med.uoa.gr)

National and Kapodistrian University of Athens https://orcid.org/0000-0002-0141-4753

Research

Keywords: COVID-19, endogenous retroviruses, HERV-W, inflammation, senescence

Posted Date: December 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-116697/v1

Abstract

Background

Severe COVID-19 pneumonia has been associated with the development of intense inflammatory responses during the course of infections with SARS-CoV-2. Given that Human Endogenous Retroviruses (HERVs) are known to be activated during and participate in inflammatory processes, we examined whether HERV dysregulation signatures are present in COVID-19 patients.

Results

By comparing transcriptomes of Peripheral Blood Monocytes (PBMCs) and Bronchoalveolar Lavage Fluid (BALF) from patients and normal controls we have shown that HERVs are intensely dysregulated in BALF, but not in PBMCs. In particular, upregulation in the expression of multiple HERV families was detected in BALF samples of COVID-19 patients, with HERV-H being the most highly upregulated family among the families analysed. In addition, we compared the expression of HERVs in Human Bronchial Epithelial Cells (HBECs) without and after senescence induction in an oncogene-induced senescence model, in order to quantitatively measure changes in the expression of HERVs in bronchial cells during the processes of cellular senescence.

Conclusions

This apparent difference of HERV dysregulation between PBMCs and BALF warrants further studies in involvement of HERVs in inflammatory pathogenetic mechanisms as well as exploration of HERVs as potential biomarkers for disease progression. Furthermore, the increase in the expression of HERVs in senescent HBECs in comparison to non-induced HBECs provides a potential link for increased COVID-19 severity and mortality in aged populations.

Background

Human Endogenous Retroviruses (HERVs) are the remnants of ancient retroviral infections that infiltrated the germlines of our deep in time ancestors and integrated within their genomes. Through Mendelian inheritance and evolutionary processes spanning millions of years they now occupy around 8% of the human genome (1). Their evolutionary history based on analyses of molecular sequences suggests that at least 30 independent colonization events generated an equivalent number of closely related retroviral integrations known as families (2).

The vast majority of HERV integrations have accumulated mutations that effectively incapacitated their proliferative functionality, however these integrants can still be transcribed and encode proteins (3). HERV expression is usually silenced through numerous post-transcriptional mechanisms, but can be upregulated in certain diseases and conditions such as cancer and inflammation (4,5). The retroviral involvement in inflammatory processes has been shown to occur through over-production of nucleic

acids and proteins that interfere with a variety of innate immune response and inflammatory pathways (6).

COVID-19 emerged in late 2019 in Wuhan City and from China spread around the globe generating the most intensive pandemic responses within the last 50 years leading to significant socioeconomic disruption. The disease is caused by a zoonotic coronavirus, now known as SARS-CoV-2, which in a proportion of individuals will cause severe pneumonia, respiratory and multi-organ failure referred to as Corona virus infectious disease 2019 (COVID-19) (7). For severe COVID-19 it has been suggested that a critical component of the pathophysiology is a severe inflammatory response driven by cytokine release (8).

Here we aimed to explore whether HERV expression is upregulated in patients with COVID-19 as this could be consistent with an involvement of HERVs in inflammatory responses such as the production of interferonogenic nucleic acids and inflammatory proteins. In the initial studies, in which the COVID-19 patients' RNAseq data were produced, the authors correlated the increased expression of proinflammatory cytokines and cytokine receptors in Bronchoalveolar Lavage Fluid (BALF) samples of COVID-19 patients when compared to healthy subjects, and related this finding to the cytokine storm and disease severity in these patients (9,10). In addition, we aimed to determine the differences between the expression of HERVs after oncogene-induced senescence of Human Bronchial Epithelial Cells (HBECs) in a non-malignant senescence model and the expression of HERVs in non-induced HBECs (11,12)In this way, we aim to determine the effect of senescence on the expression of HERVs as a potential component of the increased inflammatory burden of senescence in bronchial cells (12). We hypothesized that the expression of HERVs is augmented in the state of senescence and this could offer a plausible explanation for the increased COVID-19 severity and mortality observed as the patient age increases.

We find that multiple HERV families are upregulated in BALF, but not in Peripheral Blood Monocytes (PBMCs), in patients with COVID-19 compared to healthy individuals. The findings merit further study regarding the potential involvement of HERVs in COVID-19.

Results

Expression of HERVs in BALF samples and PBMCs from COVID-19 patients in comparison to healthy controls

Upregulation in the expression of endogenous retroviral elements was observed in the BALF samples of COVID-19 patients in comparison to healthy controls. In particular, in BALF samples of COVID-19 patients in comparison to controls statistically significant upregulation was observed in the expression of HERV-H (28.24 fold change, p=0.038), ERV-L (9.92 fold change, p=0.036), HERV-E (9.22 fold change, p=0.015), HERV-I (4.51 fold change, p=0.036), HERV-W (12.96 fold change, p=0.007), HERV-9 (17.87 fold change, p=0.036), HERV-K (HML-3) (4.76 fold change, p=0.01) and HERV-K (HML-5) (4.36 fold change, p=0.02). No statistically significant changes were observed in the expression of HERV-K (HML-1), HERV-K (HML-2), HERV-K (HML-4), HERV-K (HML-6) and HERV-FRD. **(Figure 1)**

The results of the sensitivity analyses per dataset analysed and per housekeeping gene are available in the Supplementary Table. **(Additional Files 1, 2 and 3)**

Regarding the expression of endogenous retroviral elements in the PBMCs of COVID-19 patients compared to healthy controls statistically significant downregulation was observed in the expression of ERV-L (0.81 fold change, p=0.045), HERV-FRD (0.61 fold change, p=0.019), HERV-H (0.77 fold change, p=0.01), and HERV-I (0.73 fold change, p=0.013). No statistically significant changes were observed in the expression of HERV-K (HML-1), HERV-K (HML-2), HERV-K (HML-3), HERV-K (HML-4), HERV-K (HML-5), HERV-K (HML-6), HERV-W, HERV-9 and HERV-E. **(Figure 2)**

The results of the sensitivity analyses per housekeeping gene are available in the Supplementary Table. **(Additional Files 1 and 4)**

Expression of HERVs in non-induced and induced (senescent) HBECs

Upregulation in the expression of endogenous retroviral elements was recognised in the senescence induced HBECs in comparison to non-induced HBECs. After the normalization, statistically significant upregulation in the induced HBECs was detected for HERV-K (HML-1) (1.74 fold change, p<0.001), HERV-K (HML-2) (2.11 fold change, p<0.001), HERV-K (HML-3) (1.25 fold change, p=0.002), HERV-K (HML-5) (1.56 fold change, p<0.001), HERV-K (HML-6) (1.2 fold change, p<0.001), HERV-W (1.78 fold change, p<0.001), HERV-9 (1.63 fold change, p<0.001), ERV-L (2.73 fold change, p<0.001) and HERV-E (3.77 fold change, p<0.001). No statistically significant differences were observed in the expression of HERV-H and HERV-FRD.

Downregulation in the expression of senescent HBECs in comparison to non-induced HBECs was recognised in the expression of HERV-K (HML-4) (0.65 fold change, p=0.001) and in the expression of HERV-I (0.75 fold change, p=0.007). **(Figure 3)**

The results of the sensitivity analyses per housekeeping gene are available in the Supplementary Table. **(Additional Files 1 and 5)**

Discussion

SARS-CoV-2 infection is either mild or asymptomatic in the majority of the infections (7,13). Age and underlying comorbidities like pulmonary and/or cardiovascular disease are major risk factors for severe COVID-19 (14). However, the underlying causes of the severity of the infection are not fully understood. Here we have explored whether HERVs are dysregulated in patients with COVID-19 as this could provide evidence in favour of the hypothesis that they could be implicated in the severity of the disease. Based on transcriptomic data we have found that multiple HERV families are upregulated in BALF. Interestingly, BALF samples include a wide variety of cells such as alveolar macrophages, lymphocytes, neutrophils, eosinophils and respiratory epithelial cells (15), but this increment cannot be attributed to upregulation in the transcription in PBMCs as this finding was not reproduced in these cells. Our findings combined with

data showing that entry factors for SARS-CoV-2 are co-expressed with innate immunity genes in respiratory cells suggests that HERV upregulation in BALF might indeed be relevant to local rather than systematic immunity responses (16).

While HERVs are not commonly expressed throughout life, they have on the other hand been shown to have a dual role with respect to inflammation. They are upregulated by inflammatory pathways (17), but also ERV proteins and nucleic acids may trigger inflammatory responses (18). For example, we have previously shown that IFI-16 has a broad range of ssDNA targets, thus if HERVs or other transposable elements are upregulated and reverse transcribed as a result of COVID-19, this could amplify inflammatory responses (19). On the other hand, SARS-CoV-1 and SARS-CoV-2 alike, have been found to delay the production of interferons (20,21), a mechanism leading to reduced anti-viral responses and hence has been speculated to promote the cytokine storm that has been described in severe COVID-19. In some patients developing severe COVID-19, this immune dysregulation has been linked to the presence of IgG auto-antibodies against type I INFs that pre-existed before the infection and to inborn errors in the regulation of type I INF innate immunity (22,23). It is reasonable to assume that the intricate mechanisms behind the inflammatory responses to the SARS-CoV-2 infection and its development to severe COVID-19 with cytokine storm in patients prone to severe illness is the result of the aberrant balance in the transcription of pro-inflammatory, anti-inflammatory and anti-viral immune cytokines, which may be highly influenced by the effect of exogenous and endogenous viral pressure on innate immunity. Besides, the coexistence of external pathogens with endogenous retroelements seem to modulate the endogenous retroviral transcription (24), thus one can hypothesise that another way in which this cytokine dysregulation occurs in COVID-19 is the presence of SARS-CoV-2 as a trigger for the enhanced transcription of endogenous retroviruses.

In our analyses HERV-W was one of the most highly upregulated family (up to about x13 times higher). Interestingly, HERV-W *env* has been shown to be associated with proinflammatory transcriptional signatures and induce proinflammatory responses (25). Whether the upregulation of HERV-W RNA seen in transcriptomes results in production of HERV-W proteins, or production of peptides from the retroviral immunosuppressive domain, remains to be determined, as the finding of the upregulation at an RNA transcription level might not always result in increased protein translation.

Crucially, HERV expression seems to have an association with age, which is the strongest risk factor for severe COVID-19. Differential expression of HERVs has been described when comparing older with younger adults (26). Upregulation of endogenous retroviruses as a result of aging has been studied in mice whereas failure of epigenetic mechanisms seems to be the driving force for upregulation of endogenous retroviruses with age (27). Hence, it could be hypothesized that a possible lack of epigenetic regulation of HERVs in older individuals, may result in HERV expression overdrive during COVID-19 which in turn results in massive pro-inflammatory responses (28). This association between aging and increased expression of endogenous retroviral elements is confirmed by our findings. In the frameworks of this study an upregulation in the expression of most HERV families included in our analyses is demonstrated in senescence induced HBECs in comparison to non-induced. These results indicate that

the increased expression of endogenous retroviral elements in senescent bronchial epithelial cells possibly induces local chronic inflammation, as a hallmark of immunosenescence and enhance the hypothesis of local rather than systematic inflammatory reactions. This hypothesis provides a plausible explanation for the increased COVID-19 severity and mortality as age increases.

Furthermore, recently, COVID-19 exposure has been linked to hyperinflammatory shock in peadiatric patients with manifestations that resemble Kawasaki Disease (both classic and incomplete), thus COVID-19 was linked to an unexpected increase of cases of the syndrome (29). Intriguingly, in a case series of Kawasaki peadiatric patients linked to the COVID-19 epidemic, HERV levels were detectable in the microbiology results of one of the eight afflicted children described in this case series (29). This finding could potentially indicate a further role of HERVs as an inflammatory trigger upon SARS-CoV-2 infection.

The main limitation of our study is the small number of samples; thus, our findings need to be replicated in more patients, but also need to be followed up by functional studies. BALF and PBMCs samples are not from the same patients, thus we have a limitation with respect to direct associations of the expression of endogenous retroviral elements between BALF and PBMCs in COVID-19 and different sequencing procedures were used for the production of the data included in this study. On the other hand, it is striking that HERVs were dysregulated in BALF in every patient studied, while the same pattern could not be observed in PBMCs from any patient studied. Finally, we used an oncogene-induced senescence model by which one cannot make direct deductions to endogenous retroviral transcription in COVID-19, but the use of this model aimed to demonstrate the potential causal link between ageing and increased COVID-19 severity, as the accumulation of senescent cells in lungs is associated with ageing and the senescent cells per se are considered to contribute to the ageing-associated inflammatory burden (30,31).

Due to the small number of patients analysed we cannot study potential inter-person differences with regards to the induction of transcription of endogenous retroelements in COVID-19. Thus, future studies with numerous patients could potentially elucidate any potential dose-dependent aspect of the upregulation of HERVs in COVID-19.

We find that multiple HERV families are upregulated in BALF, but not in PBMCs, in patients with COVID-19 compared to healthy individuals. Furthermore, we were able to identify upregulation in the expression of HERVs in senescence induced HBECs in comparison to non-induced cells, a fact that indicates the potential role of increased endogenous retroviral expression as a mediator of inflammatory reactions in older individuals that are at increased risk due to the disease. The findings merit further study regarding the potential involvement of HERVs in COVID-19. It thus seems feasible that should HERV expression be aetiologically linked to severe COVID-19 this expression could be a therapeutic target that would minimize the likelihood of severe Covid-19 and death. It also remains to be seen if HERV expression could act as a potential marker of disease severity.

Conclusions

In this study, we recognised dysregulated expression of endogenous retroviral elements in BALF samples, but not in PBMCs of COVID-19 patients. At the same time, we were able to identify upregulated expression of multiple HERV families in senescence induced HBECs in comparison to non-induced, a fact that could possibly explain the differences in disease severity among age groups. These results indicate that HERV expression might play a pathophysiological role in local inflammatory pathways in lungs afflicted by SARS-CoV-2 and their expression could be a potential therapeutic target.

Methods

In order to compare the expression of endogenous retroviral elements in the BALF and PBMCs of COVID-19 patients and healthy individuals, online available data were utilized. RNA sequencing (RNAseq) data from BALF samples of 4 COVID-19 patients (for three of whom two biological duplicates were available for analysis) and PBMCs of 3 COVID-19 patients and 3 healthy donors were downloaded in .fastq format from National Genomics Data Center- Genome Sequence Archive (NGDC-GSA) with accession numbers: CRR119890, CRR119891, CRR119892, CRR119893, CRR119894, CRR119895, CRR119896, CRR119897, CRR125445, CRR125446, HRR057164, HRR057168, HRR057172). RNAseq data from BALF samples of healthy individuals were downloaded from Sequence Read Archive (SRA) with accession numbers: SRR10571724, SRR10571730, SRR10571732. All data were paired-end, except HRR057164 and HRR057172 that were single-end. The platforms used for the production of the raw data were BGISEQ-500 for CRR119890, CRR119891, CRR119892 and CRR119893 (reads of 100 bases length), Illumina MiSeq for the production of CRR119894, CRR119895, CRR119896 and CRR119897 (reads of 150 bases length), Illumina NovaSeq 5000 for the production of CRR125445 and CRR125446 (reads of 150 bases length), Illumina HiSeq 2500 for the production of HRR057164 (reads of 76 bases length), HRR057168 and HRR057172 (reads of 100 bases length) and finally Illumina HiSeq 2000 for the production of SRR10571724, SRR10571730 and SRR10571732 (reads of 50 bases length). Datasets with accession numbers: CRR119894 and CRR119895, CRR119896 and CRR119897, HRR057164 and HRR057168 are biological duplicates and belong to same patient respectively.

Patients with the samples with accession numbers CRR119891, CRR119892, CRR119893, CRR119894, CRR119895, CRR119896 and CRR119897 are described as severe cases in the initial work (9). The patient whose sample has the accession numbers HRR057164 and HRR057168 is a deceased male patient with an ICU admission history due to COVID-19 and the patient whose sample is HRR057172 is a female patient with COVID-19 that did not need ICU admission (10).

In order to examine the effect of senescence on the expression of ERVs as a potential causal link between ageing and increased COVID-19 severity, we conducted an analysis on the expression of HERV families in data from a non-malignant epithelial oncogene-induced senescence model of production of senescent HBECs as was described by Komseli et al (11). In a nutshell, in order to study the precancerous and cancerous phases of tumorigenesis in epithelial cells, the researchers developed this platform where they studied the stages from non-induced human bronchial cells to senescent (induced) and then to “escaped” (cancerous) cells. Non-induced and induced (senescent) HBECs online available RNAseq data were

retrieved from NCBI-SRA with accession numbers: SRR6261633, SRR6261634, SRR6261635, SRR6261636, SRR6261637, SRR6261638, SRR6261639, SRR6261640, SRR6261641, SRR6261642, SRR6261643, SRR6261644, SRR6261645, SRR6261646, SRR6261647, SRR6261648, SRR6261649, SRR6261650, SRR6261651, SRR6261652, SRR6261653, SRR6261654, SRR6261655, SRR6261656. For the production of these data Illumina NextSeq500, producing paired-end data (reads of 38 bases length).

Bioinformatics Analysis

Bowtie2 with default settings for single-end and paired-end data was used accordingly for the alignment of these data to hg19 human genome assembly (32). Samtools view command with the -q option was used, in order for the mapping quality of the reads included in the alignments to be over 20 (33). Samtools sort and index commands were used with default settings. IGV was used for the visualisation of the mapping alignments (34). FastQC was used for the reads included in this study (35).

For the identification of the expression of endogenous retroviral elements, the coordinates of HERV-K (HML-2), HERV-H, HERV-W HERV-L, HERV-E, HERV-I, HERV-9, HERV-FRD, HERV-K (HML-1), HERV-K (HML-3), HERV-K (HML-4), HERVK (HML-5) and HERV-K (HML-6) elements were found in the existing literature (36–39). HERV-K (HML-2), HERV-H and HERV-W elements' coordinates were set in reference genome hg19. Regarding the rest of the endogenous retroviral elements, their coordinates were referring to hg38 assembly and the UCSC Batch Coordinate Conversion (liftOver) online tool was used for the conversion to hg19 coordinates (40).

Bedtools multicov command was used for the quantification of endogenous retroviral expression in each of the datasets used (41). The -f option was used in order to reassure that at least 80% of the read length were overlapping with the HERVs elements' coordinates. For this reason, based on the coordinates we used and the length of endogenous retroviral elements, we considered a length of 9000bp for HML-3, HML-5 and HML-6, 10000bp for HML-1, HML-2 and HML-4, HERV-W, HERV-L, HERV-E, HERV-I and HERV-FRD, a length 12000bp for HERV-9 and 17000bp for HERV-H for the calculation of -f for each virus, in order to increase the sensitivity of the detection. In fact, we created histograms for the distributions for the lengths of the retroelements included in this study and chose to use these lengths that were closer to the longest elements of each virus in order to ensure that reads aligned in longer viral elements would be detected and that we minimize the possibility of losing aligned reads.

Taking into account the different read lengths and different depths of sequencing among the datasets used in this analysis, we utilised the expression of widely used housekeeping genes for the normalization of our raw reads in each of the datasets. Transcription of HERVs was normalized by means of four housekeeping genes, succinate dehydrogenase (SDHA), hypoxanthine phosphoribosyl transferase 1 (HPRT1), RING-box protein 1 (RBX1) and Ras Related GTP Binding A (RRAGA). Bedtools multicov command was used for the quantification of the expression of these genes with the -f option for the detection of reads that overlap with the genes coordinates by at least 80%, the same way as with the endogenous retroviral elements.

Finally, the sum of the raw read counts from each ERV in each dataset was calculated and was normalized by dividing this number by the expression (in raw reads) of each of the housekeeping genes used, in the respective dataset. For the three out of four COVID-19 patients for whom two biological duplicates were available for analysis the average of the expression, corrected with each of the housekeeping genes used, was calculated and was considered as the patients' expression of the elements tested. Finally, in order to combine the information of the four housekeeping genes used for the needs of this analysis the median of the expression corrected with each of the housekeeping genes was calculated (normalized expression value) and these values were used for the main analyses in this work.

The fold change in the expression of the endogenous retroviral elements in COVID-19 BALF samples and PBMCs was calculated as the ratio of the mean normalized expression value of ERVs in COVID-19 patients to the mean normalized expression value of ERVs in healthy individuals. Respectively, the fold change for the expression between non-induced and induced HBECs was calculated as the ratio of the mean normalized expression value of ERVs in induced (senescent) HBECs to the mean normalized expression value of ERVs in non-induced HBECs.

Also, we have conducted sensitivity analyses with regards to the comparisons per housekeeping gene correction (separately for each one of the housekeeping genes used) for the analyses in BALF, PBMCs and HBEC datasets. Furthermore, per datasets comparisons with regards to the BALF datasets was also performed. The results for these analyses are included in the Supplementary Material provided. **(Additional File 1)**

Statistical Analysis

IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp. was used for the statistical analysis. The normalized expression value of each ERV was log-transformed (ln[normalized expression value]) in order to perform independent sample t-test, in order to identify statistically significant differences between the BALF samples and PBMCs of COVID-19 patients and healthy controls as well as to identify statistically significant differences in the expression of the studied HERVs between HBECs without and with oncogene-induced senescence. StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP. was used for the creation of the figures presented in this work.

Abbreviations

COVID-19 (Corona Virus Disease 2019)

SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus)

HERVs (Human Endogenous Retroviruses)

PBMCs (Peripheral Blood Monocytes)

BALF (Bronchoalveolar Lavage Fluid)

HBECs (Human Bronchial Epithelial Cells)

SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2)

SDHA (Succinate Dehydrogenase)

HPRT1 (Hypoxanthine Phosphoribosyl Transferase 1)

RBX-1 (RING-box protein-1)

RRAGA (Ras related GTP binding A)

HERV-W *env* (HERV-W envelope protein)

hg19 (human genome, version 19)

hg38 (human genome, version 38)

ERV (Endogenous retrovirus)

Declarations

Ethics Approval and Consent to Participate

Not applicable

Consent for Publication

Not applicable

Availability of Data and Materials

The datasets of the PBMCs from COVID-19 patients and healthy volunteers, as well as the data of BALF from COVID-19 patients analysed during the current study are available in National Genomics Data Center-Genome Sequence Archive (NGDC-GSA), in BIG Data Center (https://bigd.big.ac.cn/), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences with accession numbers CRR119890, CRR119891, CRR119892, CRR119893, CRR119894, CRR119895, CRR119896, CRR119897, CRR125445 and CRR125446 (https://bigd.big.ac.cn/gsa/browse/CRA002390) (9).

Datasets of COVID-19 patients BALF samples used in this study with accession numbers HRR057164, HRR057168, HRR057172 are available in National Genomics Data Center-Genome Sequence Archive (NGDC-GSA), in BIG Data Center (https://bigd.big.ac.cn/), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (https://bigd.big.ac.cn/bioproject/browse/PRJCA002273).

The datasets of BALF from healthy controls analysed in the present work are available in NCBI-Sequence read Archive (SRA), under accession numbers: SRR10571724 (https://www.ncbi.nlm.nih.gov/sra/?term=srr10571724), SRR10571730 (https://www.ncbi.nlm.nih.gov/sra/?term=srr10571730),

SRR10571732 (https://www.ncbi.nlm.nih.gov/sra/?term=srr10571732).

The datasets of the induced and non-induced HBECs analysed in the present work are available in NCBI-Sequence read Archive (SRA), under accession numbers: SRR6261633, SRR6261634, SRR6261635, SRR6261636 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367915), SRR6261637, SRR6261638, SRR6261639, SRR6261640 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367916), SRR6261641, SRR6261642, SRR6261643, SRR6261644 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367917), SRR6261645, SRR626163346, SRR6261647, SRR6261648 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367918), SRR6261649, SRR6261650, SRR6261651, SRR6261652 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367919), SRR6261653, SRR6261654, SRR6261655, SRR6261656 (https://www.ncbi.nlm.nih.gov/sra?term=SRX3367920).

Competing Interests

The authors declare that they have no competing interests.

Funding

Emblematic action to handle SARS-CoV-2 infection: Epidemiological study in

Greece via extensive testing for viral and antibody detection, sequencing

of the virome and genetic analysis of the carriers (GSRT No 2020ΣΕ01300001).

Authors' contributions

GM and KK designed this work, analysed and interpreted the data used in this study and drafted the first manuscript. AK, DP, TK, AK, RT, TH, SS, AK, VG, VS, ST and PL interpreted the data and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The study is supported by:

GSRT No 2020ΣΕ01300001

Emblematic action to handle SARS-CoV-2 infection: Epidemiological study in

Greece via extensive testing for viral and antibody detection, sequencing

of the virome and genetic analysis of the carriers (GSRT No 2020ΣΕ01300001).

References

1. Griffiths DJ. Endogenous retroviruses in the human genome sequence. *Genome Biol*. 2001;2(6):REVIEWS1017. doi:10.1186/gb-2001-2-6-reviews1017
2. De Parseval N, Heidmann T. Human endogenous retroviruses: From infectious elements to human genes. *Cytogenet Genome Res*. 2005;110(1-4):318-332. doi:10.1159/000084964
3. Chen J, Foroozesh M, Qin Z. Transactivation of human endogenous retroviruses by tumor viruses and their functions in virus-associated malignancies. *Oncogenesis*. 2019;8(1). doi:10.1038/s41389-018-0114-y
4. Johanning GL, Malouf GG, Zheng X, et al. Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype. *Sci Rep*. 2017;7:41960. doi:10.1038/srep41960
5. Greenig M. Hervs, immunity, and autoimmunity: Understanding the connection. *PeerJ*. 2019;2019(4). doi:10.7717/peerj.6711
6. Grandi N, Tramontano E. Human endogenous retroviruses are ancient acquired elements still shaping innate immune responses. *Front Immunol*. 2018;9(SEP). doi:10.3389/fimmu.2018.02039
7. Wu Z, McGoogan JM. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA - J Am Med Assoc*. April 2020. doi:10.1001/jama.2020.2648
8. Henry BM, de Oliveira MHS, Benoit S, Plebani M, Lippi G. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis. *Clin Chem Lab Med*. 2020;0(0). doi:10.1515/cclm-2020-0369
9. Xiong Y, Liu Y, Cao L, et al. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg Microbes Infect*. 2020;9(1):761-770. doi:10.1080/22221751.2020.1747363
10. Zhou Z, Ren L, Zhang L, et al. Heightened Innate Immune Responses in the Respiratory Tract of COVID-19 Patients. *Cell Host Microbe*. 2020;27(6):883-890.e2. doi:10.1016/j.chom.2020.04.017
11. Komseli ES, Pateras IS, Krejsgaard T, et al. A prototypical non-malignant epithelial model to study genome dynamics and concurrently monitor micro-RNAs and proteins in situ during oncogene-induced senescence. *BMC Genomics*. 2018;19(1). doi:10.1186/s12864-017-4375-1
12. Gorgoulis V, Adams PD, Alimonti A, et al. Cellular Senescence: Defining a Path Forward. *Cell*. 2019;179(4):813-827. doi:10.1016/j.cell.2019.10.005
13. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*. 2020;25(10). doi:10.2807/1560-7917.ES.2020.25.10.2000180
14. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229). doi:10.1016/S0140-6736(20)30566-3

15. Meyer KC, Raghu G, Baughman RP, et al. An official American Thoracic Society clinical practice guideline: The clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease. *Am J Respir Crit Care Med*. 2012;185(9):1004-1014. doi:10.1164/rccm.201202-0320ST
16. Sungnak W, Huang N, Bécavin C, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med*. 2020:1-7. doi:10.1038/s41591-020-0868-6
17. Manghera M, Ferguson-Parry J, Lin R, Douville RN. NF-κB and IRF1 Induce Endogenous Retrovirus K Expression via Interferon-Stimulated Response Elements in Its 5' Long Terminal Repeat. *J Virol*. 2016;90(20):9338-9349. doi:10.1128/jvi.01503-16
18. Chiappinelli KB, Strissel PL, Desrichard A, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell*. 2015;162(5):974-986. doi:10.1016/j.cell.2015.07.011
19. Hurst TP, Aswad A, Karamitros T, Katzourakis A, Smith AL, Magiorkinis G. Interferon-inducible protein 16 (IFI16) has a broad-spectrum binding ability against ss DNA targets: An evolutionary hypothesis for antiretroviral checkpoint. *Front Microbiol*. 2019;10(JULY). doi:10.3389/fmicb.2019.01426
20. Blanco-Melo D, Nilsson-Payant BE, Liu WC, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. 2020;181(5):1036-1045.e9. doi:10.1016/j.cell.2020.04.026
21. Menachery VD, Eisfeld AJ, Schäfer A, et al. Pathogenic influenza viruses and coronaviruses utilize similar and contrasting approaches to control interferon-stimulated gene responses. *MBio*. 2014;5(3). doi:10.1128/mBio.01174-14
22. Bastard P, Rosen LB, Zhang Q, et al. Auto-antibodies against type I IFNs in patients with life-threatening COVID-19. *Science (80-)*. 2020;370(6515):eabd4585. doi:10.1126/science.abd4585
23. Zhang Q, Bastard P, Liu Z, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science (80-)*. 2020;370(6515):eabd4570. doi:10.1126/science.abd4570
24. Young GR, Mavrommatis B, Kassiotis G. Microarray analysis reveals global modulation of endogenous retroelement transcription by microbes. *Retrovirology*. 2014;11:59. doi:10.1186/1742-4690-11-59
25. Bhat RK, Ellestad KK, Wheatley BM, Warren R, Holt RA, Power C. Age- and disease-dependent HERV-W envelope allelic variation in brain: Association with neuroimmune gene expression. *PLoS One*. 2011;6(4). doi:10.1371/journal.pone.0019176
26. Nevalainen T, Autio A, Mishra BH, Marttila S, Jylhä M, Hurme M. Aging-associated patterns in the expression of human endogenous retroviruses. *PLoS One*. 2018;13(12). doi:10.1371/journal.pone.0207407
27. Crichton JH, Dunican DS, MacLennan M, Meehan RR, Adams IR. Defending the genome from the enemy within: Mechanisms of retrotransposon suppression in the mouse germline. *Cell Mol Life Sci*. 2014;71(9):1581-1605. doi:10.1007/s00018-013-1468-0
28. Hurst TP, Magiorkinis G. Epigenetic control of human endogenous retrovirus expression: Focus on regulation of long-terminal repeats (LTRs). *Viruses*. 2017;9(6). doi:10.3390/v9060130

29. Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet*. 2020;0(0). doi:10.1016/S0140-6736(20)31094-1
30. Meiners S, Eickelberg O, Königshoff M. Hallmarks of the ageing lung. *Eur Respir J*. 2015;45(3):807-827. doi:10.1183/09031936.00186914
31. Parikh P, Wicher S, Khandalavala K, Pabelick CM, Britt RD, Prakash YS. Cellular senescence in the lung across the age spectrum. *Am J Physiol - Lung Cell Mol Physiol*. 2019;316(5):L826-L842. doi:10.1152/ajplung.00424.2018
32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-359. doi:10.1038/nmeth.1923
33. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
34. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754
35. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed November 12, 2020.
36. Subramanian RP, Wildschutte JH, Russo C, Coffin JM. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology*. 2011;8(1):90. doi:10.1186/1742-4690-8-90
37. Wang J, Xie G, Singh M, et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature*. 2014;516(7531):405-409. doi:10.1038/nature13804
38. Grandi N, Cadeddu M, Blomberg J, Tramontano E. Contribution of type W human endogenous retroviruses to the human genome: Characterization of HERV-W proviral insertions and processed pseudogenes. *Retrovirology*. 2016;13(1):1-25. doi:10.1186/s12977-016-0301-x
39. Bendall ML, de Mulder M, Iñiguez LP, et al. Telescope: Characterization of the retrotranscriptome by accurate estimation of transposable element expression. Patro R, ed. *PLOS Comput Biol*. 2019;15(9):e1006453. doi:10.1371/journal.pcbi.1006453
40. Kent WJ, Sugnet CW, Furey TS, et al. The Human Genome Browser at UCSC. *Genome Res*. 2002;12(6):996-1006. doi:10.1101/gr.229102
41. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. doi:10.1093/bioinformatics/btq033

Figures

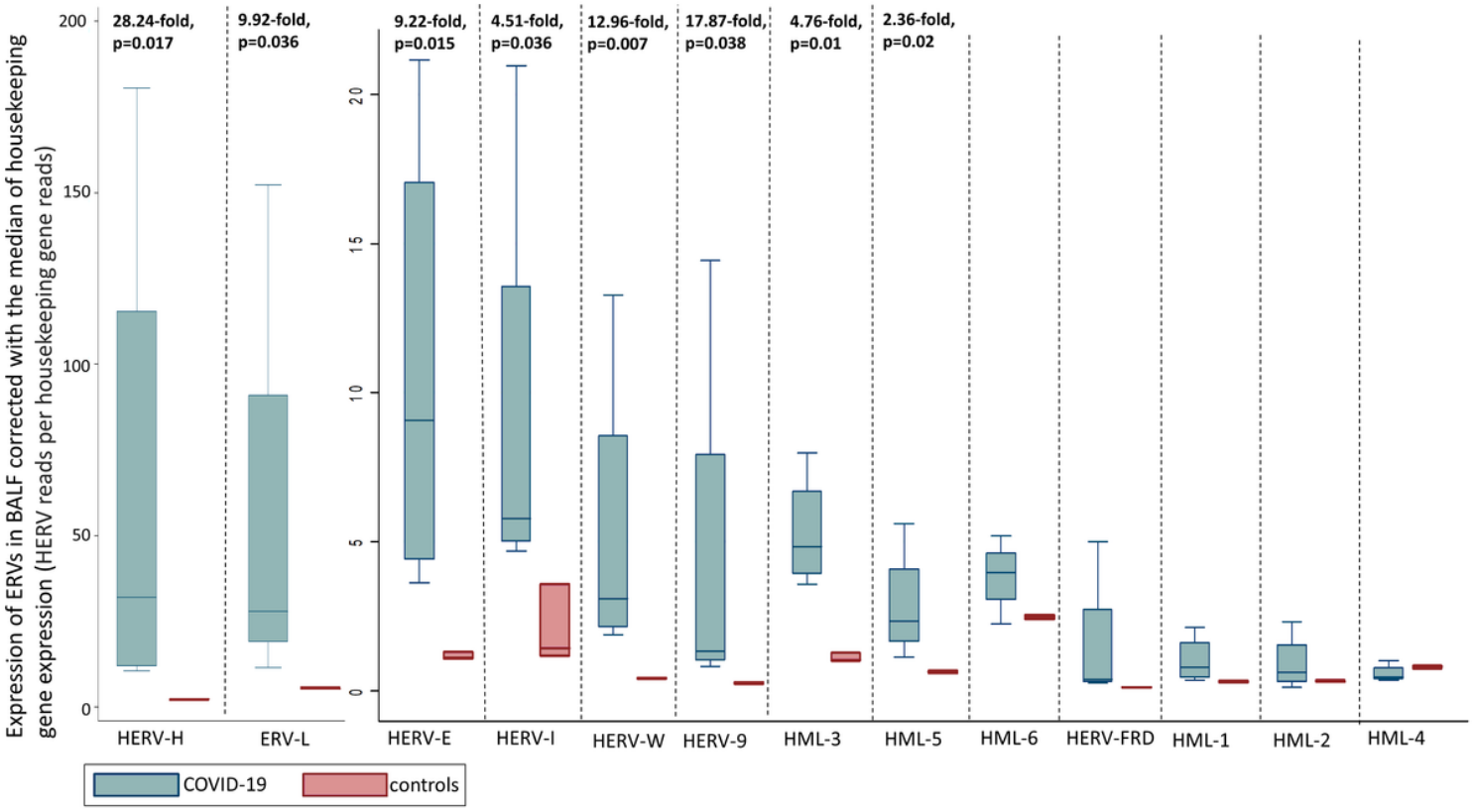

Figure 1

Transcription of HERVs in BALF samples from COVID-19 patients Legend: Transcription of HERVs in the BALF of COVID-19 patients and healthy controls corrected with the median of the transcription of four housekeeping genes (SDHA, HPRT1, RBX1 and RRAGA) expressed as HERV reads per housekeeping gene reads. We show statistically significant dysregulation of BALF compared to healthy controls as fold change.

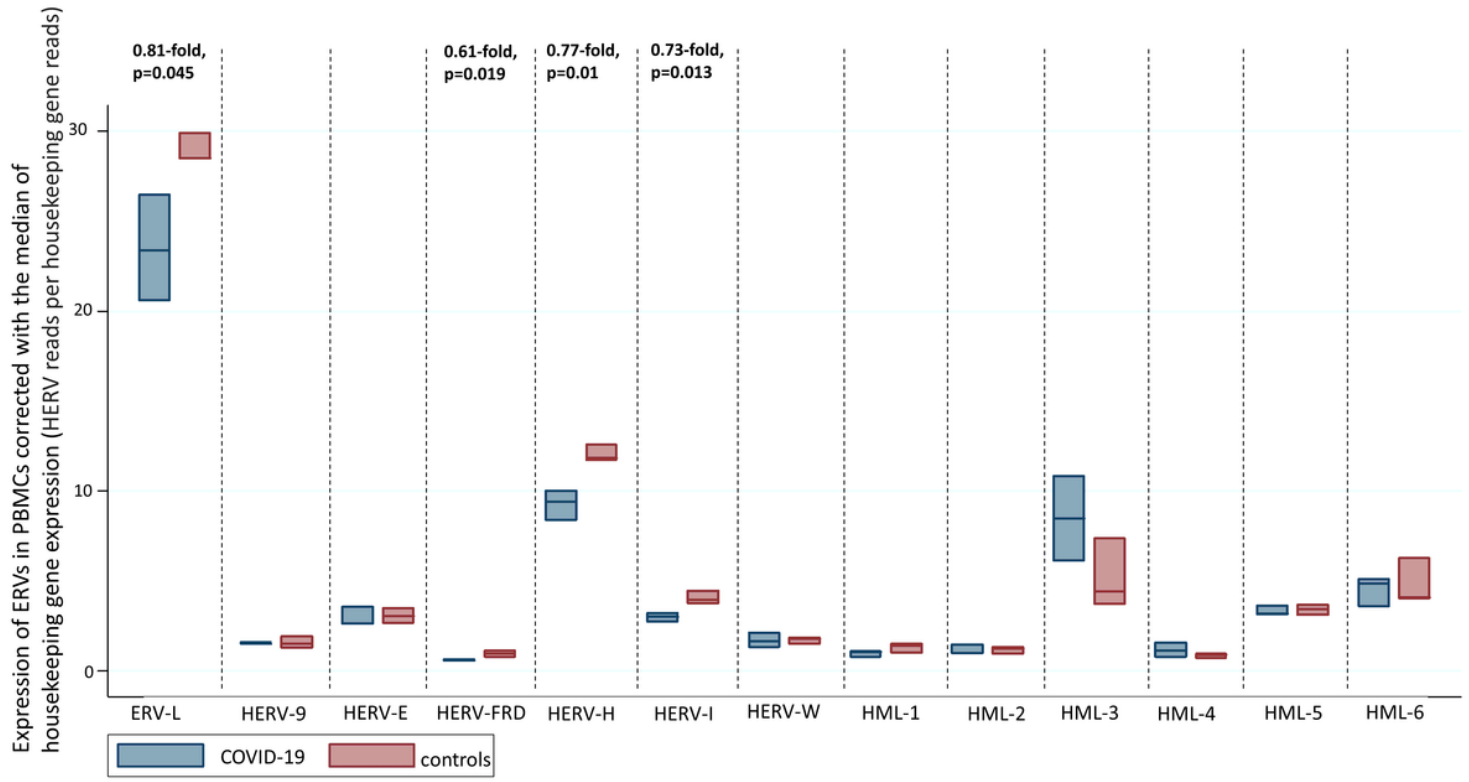

Figure 2

Transcription of HERVs in PBMCs of COVID-19 patients Legend: Transcription of HERVs in the PBMCs of COVID-19 patients and healthy controls corrected with the median of the transcription of four housekeeping genes (SDHA, HPRT1, RBX1 and RRAGA) expressed as HERV reads per housekeeping gene reads.

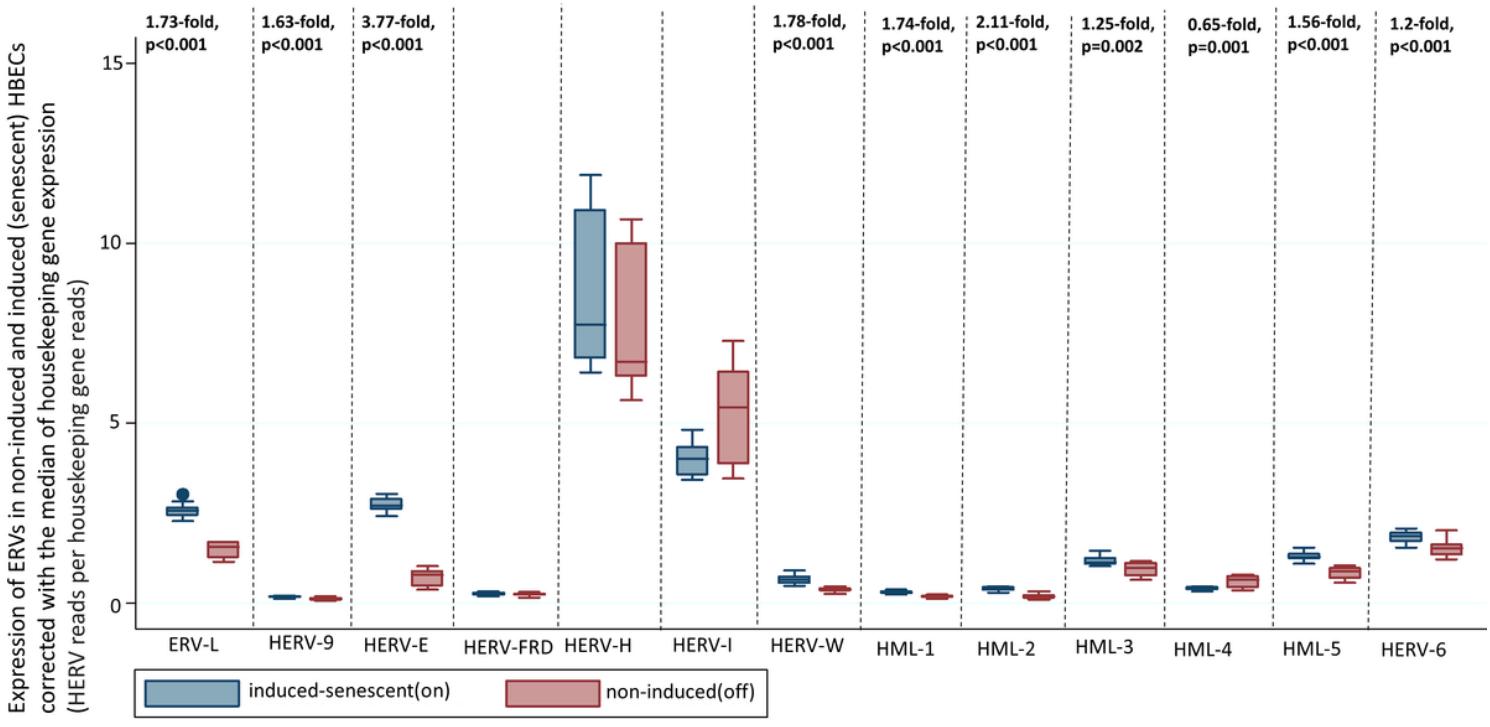

Figure 3

Transcription of HERVs in induced (senescent) HBECs Legend: Transcription of HERVs in induced (senescent) HBECs in comparison to non-induced HBECs corrected with the median of transcription of four housekeeping genes (SDHA, HPRT1, RBX1 and RRAGA) expressed as HERV reads per housekeeping gene reads.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ADDITIONALFILE1.xlsx
- ADDITIONALFILE2.pdf
- ADDITIONALFILE3.pdf
- ADDITIONALFILE4.pdf
- ADDITIONALFILE5.pdf