Vascular inflammation and endothelial injury in SARS-CoV-2 infection:

The overlooked regulatory cascades implicated by the *ACE2* gene cluster

Claire L Shovlin, MB BChir PhD FRCP[1,2]* and Marcela P. Vizcaychipi[4,5] MD PhD FRCA

[1]NHLI Vascular Science, Imperial College London, UK;

[2]Respiratory Medicine, Imperial College Healthcare NHS Trust, London UK;

[3]and Intensive Care, Chelsea & Westminster NHS Foundation Trust, London, UK.

[4]Department of Surgery and Cancer, Imperial College London, UK

Corresponding author
Claire L. Shovlin, MB, BChir, MA, PhD FRCP, Professor of Practice (Clinical and Molecular Medicine), NHLI Vascular Science, Imperial Centre for Translational and Experimental Medicine, Imperial College London, Hammersmith Campus, Du Cane Road, London W12 0NN, UK.
email c.shovlin@imperial.ac.uk; Phone 0208 383 1178.

Word Count 1626

Abstract Word Count 243

KEY POINTS

Question: In contrast to other "respiratory" viruses, why does SARS-CoV-2 result in such multi-systemic, life-threatening complications, including the vascular inflammation, endothelial injury and pulmonary thrombotic microangiopathy for which evidence is accumulating?

Findings: With ongoing viral infection, the particularly strong interaction between SARS-CoV-2 and ACE2, followed by internalisation, operates to profoundly deplete cell surface ACE2, an essential protein under feedback regulation to restore protein expression through upregulated gene transcription. Data residing unflagged in experimental repositories indicate that the *ACE2* gene lies in a co-regulated cluster with *PIR* and *VEGFD*, sharing "double-elite" enhancers, implying homeostatic host responses to restore ACE2 expression would generate damaging reactive oxygen species in the setting of impaired host defences, and separately impact on the ability of the pulmonary endothelium to respond appropriately.

Meaning: For patients who fail to suppress replication of the SARS-CoV-2 virus within a few days, the *ACE2* co-regulated gene cluster predicts delayed responses that would contribute to subsequent catastrophic deteriorations, and merit further attention.

ABSTRACT

COVID-19 has presented physicians with an unprecedented number of challenges and mortality. The basic question is why, in contrast to other "respiratory" viruses, SARS-CoV-2 infection can result in such multi-systemic, life-threatening complications and a severe pulmonary vasculopathy. It is widely known that SARS-CoV-2 uses membrane-bound angiotensin-converting enzyme 2 (ACE2) as a receptor, resulting in internalisation of the complex by the host cell. We discuss the evidence that failure to suppress coronaviral replication within 5 days results in sustained downregulation of ACE2 protein expression, and that ACE2 is under negative-feedback regulation. We then expose openly-available experimental repository data that demonstrate the gene for ACE2 lies in a novel cluster of interegulated genes on the X chromosome including *PIR* encoding pirin (quercetin 2,3-dioxygenase), and *VEGFD* encoding the predominantly lung-expressed vascular endothelial growth factor D. The five double-elite enhancer/promoters that are known to be operational, and shared read-through lncRNA transcripts, imply that ongoing SARS-CoV-2 infection will reduce host defences to reactive oxygen species, directly generate superoxide $O_2^{\cdot -}$ and H_2O_2 (a "ROS storm"), and impair pulmonary endothelial homeostasis. Published cellular responses to oxidative stress complete the loop to pathophysiology observed in severe COVID-19. Thus for patients who fail to rapidly suppress viral replication, the newly-appreciated *ACE2* co-regulated cluster predicts delayed responses that would account for catastrophic deteriorations. We conclude that ACE2 homeostatic drives provide a unified understanding which should help optimise therapeutic approaches during the wait until safe, effective vaccines and antiviral therapies for SARS-CoV-2 are delivered.

During 2020, the world has faced unexpected health, social and economic damage due to SARS-CoV-2 infection, with over 16 million cases, more than half a million deaths, and continuing high mortality rates.[1] The multifaceted severe presentations of COVID-19 have centred on viral pneumonitis/adult respiratory distress syndrome (ARDS), the "cytokine storm" of dysregulated inflammation, and multi-organ failure, with the leading cause of death now recognised as thromboembolism.[2,3] There is evidence for a hypercoagulable state in excess of that observed in sepsis,[4] with vascular inflammation, endothelial injury and thrombotic microangiopathy assuming prominence.[5-8]

The question is why are these occurring? Many inflammatory and signalling cascades have been proposed, though as yet there are no answers to the basic question of why SARS-CoV-2 infection, as for SARS-CoV, can result in such multi-systemic, life-threatening complications, differing in individual cases, or why thrombosis plays such a pronounced role in the life-threatening phenotypes.

Here we draw attention to the physiological and homeostatic responses employed by cells in response to coronavirus infection that point towards defective endovascular defences to reactive oxygen species, a "ROS storm", and specifically impaired pulmonary vascular homeostasis.

The SARS-CoV-2 "receptor" angiotensin-converting enzyme 2 (ACE2)

SARS-CoV-2, SARS-CoV and related SARS viruses are so named due to their ability to generate a severe acute respiratory syndrome (SARS) in a proportion of infected individuals. They are unusual compared to more familiar viral pathogens as they use membrane-bound angiotensin-converting enzyme 2 (ACE2) as a receptor.[9,10] ACE2 is an essential component of the renin-angiotensin-aldosterone system (RAAS), normally catalysing angiotensin II degradation, with multiple cardiovascular-protective consequences.[11] There is some evidence from other species that expression of the ACE2 receptor varies with age[12] and sex.[13]

On binding to ACE2, the SARS-CoV-2 Spike (S) envelope protein is cleaved to subunits S1 and S2 by host proteases.[9] Specific amino acids in the S1 C-terminal domain (CTD) and host ACE2 form a network of hydrogen-bond and salt bridge interactions resulting in internalization of the complex by the host cell.[9,10] The SARS-CoV-2-ACE2 interactions are 4-fold stronger than between SARS-CoV and ACE2.[10]

The kinetics of recycling ACE2 following viral-ACE2 complex internalisation are not yet reported for SARS-CoV-2, but those of a related human coronavirus, HCoV-NL63, were examined by Dijkman and colleagues in Rhesus monkey epithelial cells.[14] They demonstrated that cell surface expression of ACE2 starts to reduce 3 days post infection, becoming almost undetectable in infected cells from day 5 to the end of their study period, and therefore of unknown duration.[14]

This is for an essential protein where strong evolutionary constraints have served to maintain physiological activity: The evidence is multifactorial,[11] and most compelling *ACE2* is a "loss-of-function intolerant" gene with a very low frequency of null alleles,[15,16] particularly in males who only have a single copy due to its X-chromosome location. Transcription upregulation is observed within hours for many proteins sequestered in pathophysiological settings.[17-19] Specifically for *ACE2,* mRNA transcription is regulated by an unknown negative-feedback signal attributed to ACE2 activity.[20] Thus when ACE2 activity is reduced, homeostatic responses operate to restore the protein, apparently via upregulation of gene transcription.[20]

The *ACE 2* gene enhancers

Upregulation of gene transcription is enhanced by regulatory DNA sequences that modulate distant target gene expression, often by long-range physical chromosomal interactions. Several enhancers regulate *ACE2* transcription in man, and two (*GH0XJ015596* and *GH0XJ01557*) meet what are considered to be quite stringent GeneHancer definitions for 'double-elite' enhancer-promoter interactions.[21,22] These restrict 'elite' enhancer assignment to a candidate regulatory element derived

from more than one experimental data source, and 'double elite' promoter-enhancer interactions to those meeting likelihood-based score thresholds using expression quantitative trait loci, again derived from multiple sources.[21,22] Fewer than 10% of identified enhancer-promoter interactions fulfil the GeneHancer 'double elite' status.[21] The validity of the stringent 'double elite' enhancer–gene network is supported by the method's capture of 38 gene pairings where each of the pair represented a different causal gene for the same disease (hypergeometric probability $P<6.0\times10^{-35}$).[21]

GH0XJ015596 and *GH0XJ015579* flank *ACE2* exons 1 and 9 respectively (*Supplementary Figure 1*). Both were among ~22,000 enhancers identified by high-resolution, high throughput chromosome conformation capture (Hi-C) libraries from human blood cells- hematopoietic progenitors and lymphoblastoid cell lines, with the libraries enriched for long-range promoter contacts,[23] and also identified by the ENCODE laboratories[24] and Ensembl.[25] While the anatomical nature of GeneHancer enhancer interactions are defined and visible to all through tracks on the human genome browser from the University of Santa Cruz (UCSC),[21,22,26,27] mechanisms are not well studied. For instance it is not known whether a long non-coding (lnc)RNA that overlaps *GH0XJ015596* (*lnc-ACE2-1, Supplementary Figure 2*)[22,29] plays a functional role in any cell type

Why stimuli to upregulate ACE2 gene transcription will have wider consequences:

The *ACE2* co-regulated cluster of genes

The region on the human X-chromosome that contains the *ACE2* gene is illustrated in *Figure 1A*, and annotated by position on both human genome reference sequence "build 37" (GRCh37/hg19), and "build 38" (38/hg38). *ACE2* is encoded by the reverse (-) strand, and is sited between *CLTRN* and *PIR*, which is adjacent to *VEGF-D*.

These genes, which encode proteins with better-known names and functions, share enhancers and other regulatory elements with *ACE2* (*Figure 1B*). Homeostatic responses attempting to restore

expression of cell surface ACE in the setting of ongoing SARS-CoV-2 viral replication predict upregulation and/or perturbation of these critical proteins, as discussed in the next section.

***a) ACE2, PIR* and quercetin:**

The *ACE2* enhancers *GH0XJ015596* and *GH0XJ015579* have double-elite interactions (confirmed by Hi-C[23] and expression quantitative trait loci from GTEx version v6p[30]), with the promoter of *PIR* which encodes quercetin 2,3-dioxygenase ("pirin"), and is near-ubiquitously expressed.[30] The substrate of pirin is quercetin (3,3',4'5,7-penthydroxyflavone), a flavonoid considered to be one of the most potent antioxidants of plant origin.[31,32] Quercetin was identified as a potential COVID-19 therapy by supercomputer modelling of the viral S-human ACE2 receptor interface,[33] and by genomics-guided tracing of SARS-CoV-2 targets.[34] Quercetin has electron-donating properties due to a phenolic hydroxyl group essential for scavenging free radicals such as superoxide anion $O_2^{\cdot -}$ and these antioxidant activities are the reasons cited for two randomised controlled trials in COVID-19 that are currently recruiting on www.clintrials.gov.

Quercetin 2,3-dioxygenase, the protein product of *PIR*, converts the 'antioxidant' quercetin to *o*-quinone, with an *o*-semiquinone radical intermediate which reacts with oxygen to generate the reactive oxygen species (ROS) superoxide $O_2^{\cdot -}$ and H_2O_2.[32] Hence oxidation of quercetin by up-regulated quercetin 2,3-dioxygenase is predicted to both remove an important antioxidant and directly generate ROS.[32] Consequences will reflect the residual capacity of host anti-oxidant defences.

***b) ACE2 , VEGFD,* and vascular endothelial growth factor D**

The next "downstream" (3') gene is *VEGFD* (*Figure 1A*). There are double-elite interactions between three *PIR* locus-encoded enhancers (*GH0XJ015492, GH0XJ015473,* and *GH0XJ015471*) and the promoter of *VEGFD.*[21,22] Additionally, there is a long *PIR* and *VEGD* read-through RNA transcript that will not produce full length protein (lnc-VEGFD-1, *Figure 2*):[22,26-29] Any change in lnc-VEGFD-1

transcription, including that mediated by the *ACE2* locus-encoded enhancers interacting with the *PIR* promoter, would alter the balance between the read-through and *VEGD* protein-coding transcripts.

This is also relevant to COVID-19 since *VEGFD* encodes vascular endothelial growth factor D which is predominantly expressed in the lungs,[30] modifies differentiation, proliferation and permeability of vascular and lymphatic endothelial cells,[44] is a therapeutic target for lymphangioleiomyomatosis,[45] and independently predicts all-cause mortality in patients with suspected or known coronary artery disease.[46]

c) Additional cluster genes.

There are additional regulatory interactions between the *ACE2*-encoded enhancers and the preceding genes, *CLTRN* (also known as *TMEMM27)* and *CA5B* (*Figure 2B*). *CLTRN* encodes collectrin which has roles in renal proximal tubule amino acid transporter trafficking, insulin exocytosis, and neuropsychiatric stability.[47] *CA5B* which is on the (+) strand (*Figure 2A*), encodes the mitochondrial-expressed carbonic anhydrase 5B, a metalloproteinase that generates bicarbonate for metabolic liver enzymes: defective hepatic bicarbonate production leads to biochemical findings[48] that could be relevant to the alkalosis observed in some COVID-19 patients.

Notably, none of the major proteins or other species encoded by the contiguous genes have been linked based on curated pathways, protein–interactions, or text mining databases from over 20 million PubMed publications (*Supplementary Figure 3*).

Relevance of *ACE2* gene cluster to host responses following SARS-CoV-2 infection

The section above outlines how cellular homeostatic feedback loops, in the setting of a continued stimulus to upregulate *ACE2* transcription in infected cells with severely depleted cell surface ACE2,[15] will lead to the COVID-19 reminiscent scenario outlined in *Figure 2*.

This linking of experimentally validated data provides a compelling prediction of the delayed, life-threatening features observed following SARS-CoV-2 infection. It depends on only one assumption- that relatively normal host transcription will be in operation in some SARS-CoV-2 infected cells, some of the time. This seems a reasonable assumption based on evidence from other viruses: Many interfere with host transcription, for instance, H1N1 influenza (A/WSN/33) viral infection results in a transient host transcriptional shutdown, with host RNA polymerase II gene occupancy depleted genome-wide.[49] There has been no evidence to date for a similar pattern in coronaviruses, but they do have RNA-dependent RNA polymerases, and therefore might interfere with RNA transcription. However, for H1N1 influenza (A/WSN/33), escape occurs after several hours, when the virus switches to replication mode.[50] Given the prolonged time-course for COVID-19 infection, it would seem reasonable to assume that relatively normal host transcription will be in operation in some to most infected cells.

SUMMARY

In these unprecedented times, there are substantial existing resources of experimental biological data, currently residing unflagged in major repositories, and available for interrogation, informed by first-hand knowledge of human physiology and pathology in the setting of COVID-19.

These indicate that for patients who fail to suppress replication of the SARS-CoV-2 virus within a few days, delayed consequence of essential, post-infective homeostatic responses by SARS-CoV-2-infected cells could account for subsequent catastrophic deteriorations, mediated by defective host responses

to reactive oxygen species, augmented generation of reactive oxygen species, and ROS-induced cellular injury. Further examination is warranted.

Acknowledgments:

We gratefully acknowledge the research laboratories and bioinformatics groups whose open source materials made these analyses feasible, particularly the UCSC Genome Browser Group of the University of California Santa Cruz; the international teams performing high-resolution capture Hi-C and generating GeneHancer; the Genotype-Tissue Expression (GTEx), ENCODE and Ensembl Consortia, and the ENCODE production laboratories.

Competing interests:

The authors have no competing interests to declare.

Ethical approvals:

None required. All data are in open source databases.

Author contributions:

Both authors performed literature searches, and designed the work based on clinical experience, particularly from MPV, and genetic experience, particularly from CLS. CLS performed the data analysis, generated the Figures and wrote the first draft. Both authors contributed to data interpretation, and manuscript revisions before joint approval.

REFERENCES

1 Johns Hopkins University Center for Systems Science and Engineering (CSSE). COVID-19 Dashboard. Available from https://coronavirus.jhu.edu/map.html [Accessed 24th July 2020].

2 Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, 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):1054–10

3 Vizcaychipi MP, Shovlin CL, Hayes M, Singh S, Christie L, Sisson A et al. Early detection of severe COVID-19 disease patterns define near real-time personalised care, bioseverity in males, and decelerating mortality rates. *Medrxiv 2020 May 11.* Available from https://medrxiv.org/cgi/content/short/2020.05.08.20088393v1 [Accessed 24th July 2020, *(manuscript in late stage peer review*].

4 Iba T, Levy JH, Connors JM, Warkentin TE, Thachil J, Levi M. The unique characteristics of COVID-19 coagulopathy. *Crit Care.* 2020;24(1):360.

5 Patel BV, Arachchillage DJ, Ridge CA, Bianchi P, Doyle JF, Garfield B et al. Pulmonary Angiopathy in Severe COVID-19: Physiologic, Imaging and Hematologic Observations. *Am J Respir Crit Care Med*. 2020 Jul 15. Epub ahead of print. Available from https://doi.org/10.1164/rccm.202004-1412OC [Accessed 24th July 2020].

6 Fox SE, Akmatbekov A, Harbert JL, Li G, Brown JQ, Vander Heide RS. Pulmonary and Cardiac Pathology in African American patients with COVID-19: an autopsy series from New Orleans. *Lancet Respir Med.* 2020 Jul;8(7):681-686.

7 Magro C, Mulvey JJ, Berlin D, Nuovo G, Salvatore S, Harp J, et al. Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: A report of five cases. *Transl Res.* 2020 Jun;220:1-13.

8 Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N Engl J Med.* 2020;383(2):120-128

9 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.* 2020 Apr 16;181(2):271-280.e8.

10 Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell.* 2020; 181(4):894-904.

11 South AM, Diz D, Chappell MC. COVID-19, ACE2 and the Cardiovascular Consequences. *Am J Physiol Heart Circ Physiol.* 2020 May 1;318(5):H1084-H1090.

12 Song R, Preston G, Yosypiv IV. Ontogeny of angiotensin-converting enzyme 2. *Pediatr Res*. 2012; 71:13–19.

13 Xie X, Chen J, Wang X, Zhang F, Liu Y. Age- and gender-related difference of ACE2 expression in rat lung. *Life Sci.* 2006; 78:2166–2171.

14 Dijkman R, Jebbink MF, Deijs M, Milewska A, Pyrc K, Buelow E, et al. Replication-dependent downregulation of cellular angiotensin-converting enzyme 2 protein expression by human coronavirus NL63. *J Gen Virol.* 2012 Sep;93(Pt 9):1924-1929.

15 Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

16 Shovlin CL, Vizcaychipi MP. COVID-19 genomic susceptibility: definition of ACE2 variants relevant to human infection with SARS-COV-2. *MedRxiv* 2020 May 18. Available from https://www.medrxiv.org/content/10.1101/2020.05.12.20098160v3 [Accessed 24th July 2020].

17 Noyes WD, Garby L. Rate of haptoglobin in synthesis in normal man. Determinations by the return to normal levels following hemoglobin infusion. *Scand J Clin Lab Invest* 1967:20(1)33-38, DOI: 10.1080/00365516709076917

18 Morris AR, Stanton DL, Roman D, Liu AC. Systems Level Understanding of Circadian Integration with Cell Physiology. *J Mol Biol*. 2020;432(12):3547-3564.

19 Mollet IG, Patel D, Govani FS, Giess A, Paschalaki K, Periyasamy M et al. Low dose iron treatments induce a DNA damage response in human endothelial cells within minutes. *PLoS One.* 2016 Feb 11;11(2):e0147990.

20 Ferrario CM, Jessup J, Chappell MC, Averill DB, Brosnihan KB, Tallant EA et al. Effect of angiotensin-converting enzyme inhibition and angiotensin II receptor blockers on cardiac angiotensin-converting enzyme 2. *Circulation.* 2005 May 24;111(20):2605-10.

21 Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet.* 2015; 47, 598–606.

22 Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford).* 2017 Jan 1;2017. doi: 10.1093/database/bax028.

23 Stelzer G, Rosen R, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics*. 2016 Jun 20;54:1.30.1-1.30.33.

24 Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* 2018; 46(D1):D794-D801

25 Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J et al. Ensembl 2020. *Nucleic Acids Res.* 2020; 48 (D1):D682–D688.

26 Harrow J, Denoeud F, Frankish A, Reymond A, Chen CK, Chrast J, et al. GENCODE: producing a reference annotation for ENCODE. *Genome Biol.* 2006;7 Suppl 1:S4.1-9

27 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM et al. The human genome browser at UCSC. *Genome Res.* 2002; 12(6):996-1006

28 Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, et al. The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* 2019; 47(D1):D853-D858

29 Volders P-J, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, et al. LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* 2019;47(D1):D135-D139.

30 The Genotype-Tissue Expression (GTEx) project. The GTEx Consortium. *Nat Genet.* 2013; 45(6):580-5.

31 Erdman JW Jr, Balentine D, Arab L, Beecher G, Dwyer JT, Folts J et al. Flavonoids and heart health: proceedings of the ILSI North America Flavonoids Workshop, May 31-June 1, 2005, Washington, DC. *J Nutr.* 2007 Mar;137(3 Suppl 1):718S-737S.

32 Murakami A, Ashida H, Terao J. Multitargeted cancer prevention by quercetin. *Cancer Lett.* 2008: 269; 315-325.

33 Smith M, Smith JC. Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface. *ChemRxiv.* 2020 Mar 11. Available from https://doi.org/10.26434/chemrxiv.11871402v4 [Accessed 24th July 2020].

34 Glinsky GV: Tripartite Combination of Candidate Pandemic Mitigation Agents: Vitamin D, Quercetin, and Estradiol Manifest Properties of Medicinal Agents for Targeted Mitigation of the

COVID-19 Pandemic Defined by Genomics-Guided Tracing of SARS-CoV-2 Targets in Human Cells. *Biomedicines.* 2020 May 21;8(5):129.

35 Wu J. Tackle the free radicals damage in COVID-19. *Nitric Oxide.* 2020;102:39-41.

36 Aggarwal S, Lazrak A, Ahmad I, Yu Z, Bryant A, Mobley JA, et al. Reactive species generated by heme impair alveolar epithelial sodium channel function in acute respiratory distress syndrome. *Redox Biol.* 2020 Jun 1;36:101592.

37 McBride MA, Owen AM, Stothers CL, Hernandez A, Luan L, Burelbach KR, et al. The Metabolic Basis of Immune Dysfunction Following Sepsis and Trauma. *Front Immunol.* 2020 May 29;11:1043.

38 Mohanty JG, Nagababu E, Rifkind JM. Red blood cell oxidative stress impairs oxygen delivery and induces red blood cell aging. *Front Physiol.* 2014;5:84.

39 He P, Talukder MAH, Gao F. Oxidative stress and microvessel barrier dysfunction. *Front Physiol.* 2020 May 27;11:472.

40 Mangalmurti NS, Reilly JP, Cines DB, Hunter CA, Meyer NJ, Vaughan AE. COVID-ARDS Clarified: A Vascular Endotype? *Am J Respir Crit Care Med*. *2020* Jul 6 online ahead of print. Available from https://doi.org/10.1164/rccm.202006-2598LE. [Accessed 24th July 2020].

41 Shi W, Lv J, Lin L. Coagulopathy in COVID-19: Focus on vascular thrombotic events. *J Mol Cell Cardiol*. 2020 Jul 15;146:32-40.

42 Pendergrass KD, Gwathmey TM, Michalek RD, Grayson JM, Chappell MC. The angiotensin II-AT1 receptor stimulates reactive oxygen species within the cell nucleus. *Biochem Biophys Res Commun*. 2009;384(2):149-154.

43 Gwathmey TM, Alzayadneh EM, Pendergrass KD, Chappell MC. Novel roles of nuclear angiotensin receptors and signaling mechanisms. *Am J Physiol Regul Integr Comp Physiol*. 2012;302(5):R518-R530.

44 Leppänen VM, Jeltsch M, Anisimov A, Tvorogov D, Aho K, Kalkkinen N, et al. Structural determinants of vascular endothelial growth factor-D receptor binding and specificity. *Blood.* 2011;117(5):1507-15

45 Taveira-DaSilva AM, Jones AM, Julien-Williams P, Stylianou M, Moss J. Long-Term effect of sirolimus on serum vascular endothelial growth factor D levels in patients with lymphangioleiomyomatosis. *Chest.* 2018;153(1):124-132.

46 Wada H, Suzuki M, Matsuda M, Ajiro Y, Shinozaki T, Sakagami S et al. Distinct characteristics of VEGF-D and VEGF-C to predict mortality in patients with suspected or known coronary artery

disease. *J Am Heart Assoc.* 2020 May 5;9(9):e015761. Available from doi: 10.1161/JAHA.119.015761. [Accessed 24th July 2020].

47 Pillai NR, Yubero D, Shayota BJ, Oyarzabal O, Ghosh R, Sun Q et al. Loss of CLTRN function produces a neuropsychiatric disorder and a biochemical phenotype that mimics Hartnup disease. *Am J Med Genet A.* 2019 Dec;179(12):2459-2468

48 Diez-Fernandez C, Rüfenacht V, Santra S, Lund AM, Santer R, Lindner M et al. Defective hepatic bicarbonate production due to carbonic anhydrase VA deficiency leads to early-onset life-threatening metabolic crisis. *Genet Med.* 2016 Oct;18(10):991-1000.

49 Rivas HG, Schmaling SK, Gaglia MM. Shutoff of host gene expression in influenza A virus and herpesviruses: similar mechanisms and common themes. *Viruses.* 2016 Apr 16;8(4):102.

50 Bauer DLV, Tellier M, Martínez-Alonso M, Nojima T, Proudfoot NJ, Murphy S et al. Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. *Cell Rep.* 2018 May 15;23(7):2119-2129.e3.

.

FIGURE LEGENDS

Figure 1: The *ACE2* Gene Cluster

A) Cartoon representation of the X chromosome region flanking *ACE2,* drawn to scale, with coordinates for the two human genome builds in current use, GRCh37/hg19 and Build 38/hg38. Discrete genes and other loci are encoded by either the plus (+) or the minus (-) strand, and are indicated by boxes. The four *ACE2* cluster genes are denoted in darker blue, with *ACE2*-related enhancer elements indicated in orange, with major read-through RNA transcripts indicated in pink below: *PIR-FIGF*[22], known also as *lncVEGFD-1.2*[29] reads from *PIR* exon 2 to the 3' untranslated region of *VEGFD,* skipping the final exon of *PIR* and first exon of *VEGFD*. There is also a read-through transcript between *CLTRN* and *ACE2 (AC097625.2*[26]) that skips the final exon of *CLTRN* and first exon of *ACE2*. **B)** The interacting cluster of contiguous X chromosome genes (not to scale). Circle symbols are blue for protein coding genes, orange for double-elite enhancers, and purple for non-coding RNAs. Grey background indicates overlapping loci on current resolutions. Further, generally weaker interactions have been described and are illustrated for completeness below the core cluster box. Notably however, *CA5B* is regulated differently to *ACE2* (data not shown).

Figure 2: Conceptual Model of *ACE2* homeostatic responses

Black text indicates consequences of SARS-CoV-2 binding leading to ACE2 sequestration/ internalisation, in addition to subsequent consequences of viral infection, and therapies. Green text indicates homeostatic responses to increase synthesis of the full length ACE2 mRNA, and thus protein. Off target consequences are predicted to increase *PIR* transcription, and thus impair reactive oxygen species (ROS) scavenging and increase ROS as described in the text. VEGFD consequences are expected to be pulmonary endothelial-specific, and additional cluster genes are not illustrated for clarity.

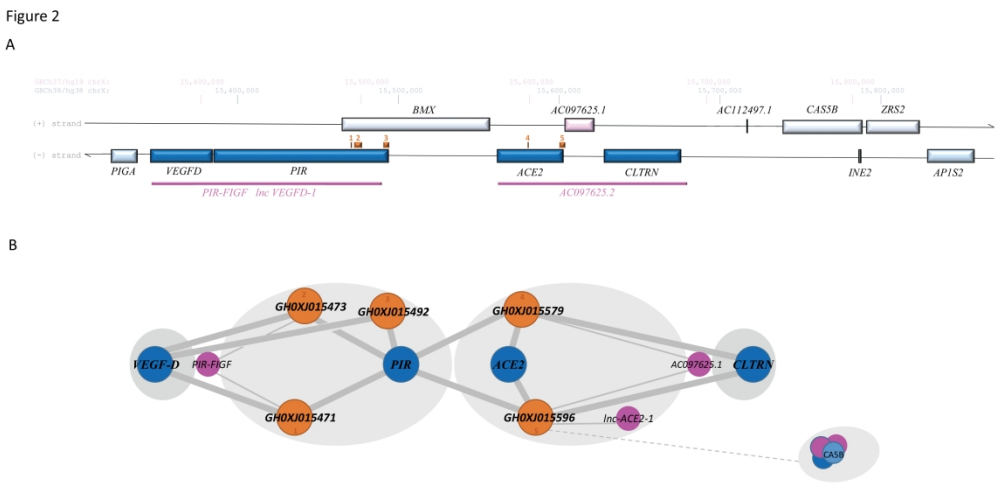

Figure 1: The *ACE2* Gene Cluster.A) Cartoon representation of the X chromosome region flanking *ACE2*, drawn to scale, with coordinates for the two human genome builds in current use, GRCh37/hg19 and Build 38/hg38. Discrete genes and other loci are encoded by either the plus (+) or the minus (-) strand, and are indicated by boxes. The four *ACE2* cluster genes are denoted in darker blue, with *ACE2*-related enhancer elements indicated in orange, with major read-through RNA transcripts indicated in pink below: *PIR-FIGF*,[22] known also as *lncVEGFD-1.2*[29] reads from *PIR* exon 2 to the 3' untranslated region of *VEGFD*, skipping the final exon of PIR and first exon of *VEGFD*. There is also a read-through transcript between *CLTRN* and *ACE2*(*AC097625.226*) that skips the final exon of CLTRN and first exon of *ACE2*. **B)** The interacting cluster of contiguous X chromosome genes (not to scale). Circle symbols are blue for protein coding genes, orange for double-elite enhancers, and purple for non-coding RNAs. Grey background indicates overlapping loci on current resolutions. Further, generally weaker interactions have been described and are illustrated for completeness below the core cluster box. Notably however, *CA5B* is regulated differently to *ACE2* (data not shown).

Figure 3

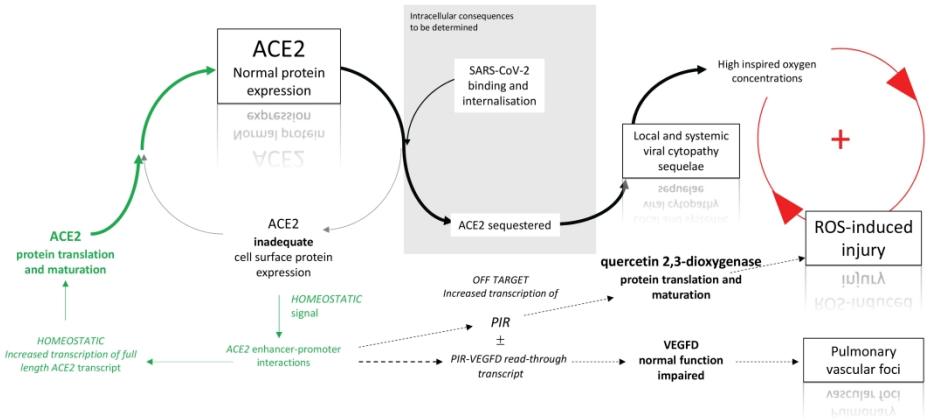

Figure 2: Conceptual Model of *ACE2* homeostatic responses. Black text indicates consequences of SARS-CoV-2 binding leading to ACE2 sequestration/ internalisation, in addition to subsequent consequences of viral infection, and therapies. Green text indicates homeostatic responses to increase synthesis of the full length ACE2 mRNA, and thus protein. Off target consequences are predicted to increase *PIR* transcription, and thus impair reactive oxygen species (ROS) scavenging and increase ROS as described in the text. *VEGFD* consequences are expected to be pulmonary endothelial-specific. Additional cluster genes are not illustrated for clarity.