scientific reports

Check for updates

OPEN

SARS-CoV-2 receptor ACE2 is co-expressed with genes related to transmembrane serine proteases, viral entry, immunity and cellular stress

Wasco Wruck & James Adjaye✉

The COVID-19 pandemic resulting from the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which emerged in December 2019 in Wuhan in China has placed immense burden on national economies and global health. At present neither vaccination nor therapies are available. Here, we performed a meta-analysis of RNA-sequencing data from three studies employing human lung epithelial cells. Of these one focused on lung epithelial cells infected with SARS-CoV-2. We aimed at identifying genes co-expressed with angiotensin I converting enzyme 2 (ACE2) the human cell entry receptor of SARS-CoV-2, and unveiled several genes correlated or inversely correlated with high significance, among the most significant of these was the transmembrane serine protease 4 (TMPRSS4). Serine proteases are known to be involved in the infection process by priming the virus spike protein. Pathway analysis revealed virus infection amongst the most significantly correlated pathways. Gene Ontologies revealed regulation of viral life cycle, immune responses, pro-inflammatory responses- several interleukins such as IL6, IL1, IL20 and IL33, IFI16 regulating the interferon response to a virus, chemo-attraction of macrophages, and cellular stress resulting from activated Reactive Oxygen Species. We believe that this dataset will aid in a better understanding of the molecular mechanism(s) underlying COVID-19.

Severe acute respiratory disease COVID-19 is a result of infections with the coronavirus SARS-CoV-2 first reported in the Chinese city Wuhan, Province Hubei, in December 2019 and has since 11 March 2020 been designated as a pandemic by WHO. The origin of the virus is most likely zoonotic[1,2] but the exact species transferring it is still under investigation as some studies suggest that it was transferred from pangolins[3] or bats[4]. Tilocca et al. analysed the SARS-CoV-2 nucleocapsid and envelop proteins and found besides highest similarities with bat and pangolin also considerable similarities with dog, cat, cattle and other species[5,6]. Based on this, they suggest that earlier contact to similar viruses hosted by other species might be responsible for protection or—in case of multiple contacts—for antibody defense enhancement[7]. At the end of April 2020, the number of globally confirmed cases of COVID-19 exceeded 3 million and recorded deaths beyond 200,000 in the real-time statistics of the Johns Hopkins University[8]. Due to many unreported and asymptomatic cases, the infection fatality rate (IFR) is difficult to determine however Verity et al. estimate approximately 0.66% (0·39–1·33) in China[9]. The age-associated IFR is approximately 7.8% for those above 80 years[9]. Drugs have been re-purposed for stabilizing COVID-19, but these are not effective therapies[10], examples include hydroxy-chloroquine (Malaria)[11,12] and nelfinavir (HIV)[13]. However, remdesivir which was designed for Ebola treatment[11,12] was at least shown to shorten the time to recovery and to reduce infection in the lower respiratory tract[14]. Another treatment option is to indirectly immunize individuals with plasma from convalescent COVID-19 patients[15]. Further approaches aim at mimicking the human virus cell entry receptor ACE2[4] with human recombinant soluble ACE2 (hrsACE2)[16]. The cell entry receptor ACE2 associates with transmembrane proteases which prime the spike protein of the virus. Hoffmann et al. assigned this task to the protein TMPRSS2[17] which they identified in the predecessor virus SARS-CoV and showed it is the same for SARS-CoV-2. The protease can be inhibited by existing compounds to interrupt further propagation of the virus in the human host. Several other publications confirmed the role

Institute for Stem Cell Research and Regenerative Medicine, Medical Faculty, Heinrich Heine University, Moorenstr.5, 40225 Düsseldorf, Germany. ✉email: james.adjaye@med.uni-duesseldorf.de

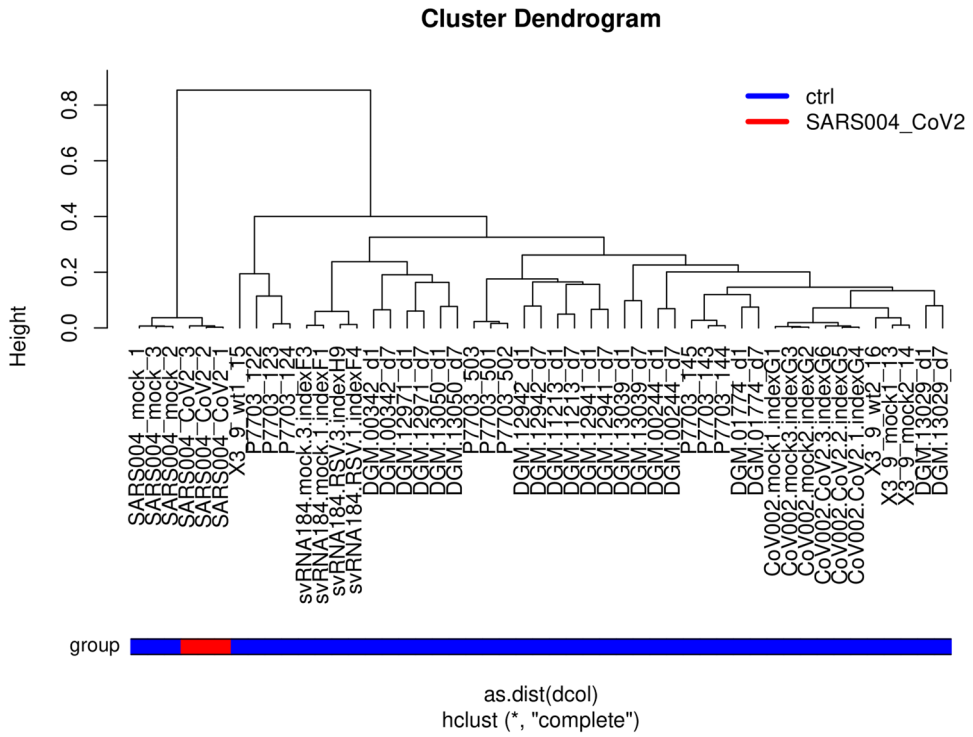

Figure 1. Lung epithelial cells (labeled "SARS004") infected with SARS-CoV-2 cluster together with mock-infected lung epithelial cells but separated from all other lung epithelial cells.

Dataset	Description	Reference
GSE147507	Primary human bronchial epithelial cells and lung adenocarcinoma infected with SARS-CoV-2 or Mock	Blanco-Melo, D. et al. Cell 181, 1036–1045.e9 (2020)
GSE146482	Human bronchial epithelium cell line BEAS-2B	Mukherjee SP et al. unpublished
GSE85121	Small airway epithelium brushing	Staudt MR et al. Respir Res 2018 May 14;19(1):78. PMID: 29,754,582

Table 1. Datasets used in this focussed meta-analysis.

of TMPRSS2 including a study by Matsuyama et al. showing enhanced SARS-CoV-2 infection in TMPRSS2 expressing VeroE6 cells[18] and a study by Sungnak et al. reporting co-expression of ACE2 and TMPRSS2 in multiple tissues[19]. Collin et al. found that SARS-CoV-2 may enter the human body through the ocular surface epithelium mediated by ACE2 and TMPRSS2[20].

Interestingly, ACE2 has been reported to be down-regulated following lung injury by Imai et al.[21] and in the previous virus SAR-CoV by Haga et al.[22] via a mechanism involving TNF-α-converting enzyme (TACE) shedding of the ectodomain of ACE2.

Here, we describe a meta-analysis focussing on the transcriptome data from human lung epithelial cells including samples infected with SARS-CoV-2 from a study described by Blanco Melo et al.[23]. We directed the exploration to co-expression with the known SARS-CoV-2 receptor ACE2. The analysis revealed a signature consisting of 72 genes significantly co-expressed with ACE2 either with positive or negative Pearson correlation. Of the transmembrane serine proteases, the most significantly co-expressed with ACE2 was TMPRSS4, suggesting it to be a putative druggable target.

Results

Cluster analysis of SARS-CoV-2 infected cells compared to other non-infected lung cells. Figure 1 shows a hierarchical cluster analysis of all samples used in this meta-analysis. Lung epithelial cells (labeled "SARS004" in Fig. 1) infected with SARS-CoV-2 cluster together with mock-infected lung epithelial cells but separated from all other lung epithelial cells and lung carcinoma cell lines which together we consider as control in this analysis. The table of Pearson correlation co-efficients (suppl. Table 2) reflects the grouping implicated by the hierarchical clustering: SARS-CoV-2 samples have highest correlation (r = 0.9884–0.9936) to the mock-infected SARS cells (Table 1).

Analysis of genes with correlated and anti-correlated expression with ACE2. Building on the knowledge about ACE2 as receptor of the SARS-CoV-2 virus we set out to indentify genes with highly correlated expression with ACE2 with the aim of elucidating the molecular mechanisms underlying COVID-19. Figure 2 shows the genes most significantly (Bonferoni-adjusted p < 1E−1) correlated (red to yellow in last column) or anti-correlated (green) with ACE2 (full table in suppl. Table 3). The transmembrane serine protease 4 (TMPRSS4) is one of the most significantly correlated (r = 0.9142, p = 4.59E−20) with ACE2 therefore implying a major role in priming the SARS-CoV-2 spike protein. CXCL17 (r = 0.9273, p = 1.1E−21), ABCA12 (r = 0.9256, p = 1-92E−21) and ATP10B (r = 0.9193, p = 1.14E−20) had marginally higher correlation with ACE2 while another transmembrane protease TMPRS11E (r = 0.9121, p = 7.91E-20) had a slightly lower correlation. The expression of CXCL17 is probably due to a reaction to the infection by chemo-attracting macrophages[24]. The role of the ATP binding cassette subfamily A member 12 (ABCA12) is not fully elucidated with respect to COVID-19 but assumed to transport lipid via lipid granules to the intracellular space and transporting specific proteases – in the case of harlequin ichtyosis associated with desquamation[25]. Current knowledge on ATP10B is scarce. However, Wilk et al. (Table 3 in their publication) report Atp10b gene expression levels as highly inversely (negative) correlated with influenza gene expression changes in infected C57BL/6 J mice[26]. In Fig. 3a a cluster analysis and gene expression heatmap of the 72 most significantly (Bonferoni-adjusted p < 1E−11) correlated and anti-correlated genes with ACE2 shows close clustering of the serine protease TMPRSS4 with ACE2. Also in this analysis of 72 genes, SARS-CoV-2 (red color bar) cluster together with mock-infected SARS lung epithelial cells and separated from the other lung cells (blue color bar indicates control). In the heatmap presented in Fig. 3b, TMPRSS family members TMPRSS11D/E and TMPRSS4 show close clustering with ACE2 but also TMPRSS2 and TMPRSS13 have similar expression in all experiments, especially in the SARS-CoV-2 infected samples.

Pathway analysis of genes co-regulated with ACE2. In order to investigate the functionality of genes interacting with *ACE2* we filtered genes correlated with ACE2 by Bonferoni-adjusted p value < 0.05 and Pearson correlation coefficient > 0.6. 1891 genes fulfilled these criteria and were subjected to over-representation analysis of KEGG pathways[27]. The most significantly over-represented pathways associated with the 1891 genes correlated with *ACE2* (Fig. 4a, suppl. Table 4) are for example, *Bacterial invasion of epithelial cells (q = 4.4E−06), Human papillomavirus infection (q = 0.0006), Transcriptional misregulation in cancer (q = 0.0006)* and *Endocytosis (q = 0.002)*. This reflects the mechanisms of virus infection via invasion of epithelial cells and endocytosis.

Pathway analysis of genes anti-correlated with ACE2. Analogously to the positively correlated genes we also examined the negatively correlated genes with ACE2 by filtering for Bonferoni-adjusted p value < 0.05 and Pearson correlation coefficient < − 0.6. 1993 genes passed these filtering criteria and were subjected to over-representation analysis of KEGG pathways[27]. The most significantly over-represented pathways in the 1993 genes negatively correlated with *ACE2* (Fig. 4b, suppl. Table 5) are *DNA replication (q = 1E−12), Metabolic pathways (q = 1.86E.08), Cell cycle (q = 1.1E−05), Fanconi anemia pathway (q = 1.24E−05), Mismatch repair (q = 9.89E−05)* and *Homologouos recombination (q = 0.0046)*. Many of these pathways are associated with DNA processing or repair. That these are over-represented in genes down-regulated upon infection with SARS-CoV-2 is in line with reports about interferon and interferon stimulated genes (ISGs) inhibiting virus replication[28]. This would be a defense against the attempts of the virus to recruit the host's DNA repair and homologous recombination mechanisms[29].

GO analysis of genes co-regulated with ACE2. We furthermore assessed the GOs over-represented in the 1891 genes positively and the 1993 genes negatively correlated with *ACE2*. Table 2 shows a selection of significant GOs from all three categories Biological Process (BP, Fig. 5a), Cellular Component (CC) and Molecular Function (MF), suppl. Table 6 provides the full table and suppl. Table 7 provides the full table for the 1993 genes negatively correlated with *ACE2*. Amongst the GO-BPs, metabolic processes are the most significant. GO-BP terms such as *Interspecies interaction between organisms, Cytokine production* and *positive regulation of viral process* reflect activated mechanisms post-viral infection. Interestingly, we found *regulation of coagulation* amongst the GO-BPs what may help elucidate reports about co-agulation in acro-ischemic COVID-19 patients[30]. In the GO-CCs, the terms *intracellular* and *membrane-bounded organelle* are most significant. In the GO-MFs, the terms *metal ion binding* and *protein binding* emerge as most significant due probably reflecting the binding of the virus proteins to the host cells. For the full gene lists associated with these terms refer to suppl. Tables 5 and 6.

Immune system associated GOs co-regulated with ACE2. Table 3 and Fig. 5b show GOs (all Biological Processes) related to the immune system over-represented in genes correlated with *ACE2*. Myeloid cells involved in the immune response (GO:0,002,275, p = 5E−07) as well as T-cells (GO:0,050,870, p = 0.02) are activated. Chemokines, in particular interleukin-1 are produced (GO:0,032,642, p = 0.0049, GO:0,032,652, p = 0.0049) also IL33 and TNF. Additionally, positive regulation of the innate immune response was prominent (GO:0,045,089, p = 0.0079). *CXCL17* was the gene with the highest correlation with *ACE2* and is involved in immune system process—as described above by chemo-attracting macrophages[24]. Among the most significantly *ACE2*-correlated genes was *IFI16* which is associated with several immune system GOs and known as regulator of the interferon response to viruses[31] and will be described in more detail in the next section about *Protein interaction networks*. Also in the protein-interaction network of the most significantly ACE2-correlated genes was the interleukin 20 receptor B (IL20RB) which appeared in several GOs listed in Table 3. With respect to viruses, there is meagre knowledge on IL20RB, however a study reported over-expression in the pneumonia causing avian influenza A H7N9 virus[32].

symbol	cor.p	cor.BH	cor.bonferoni	cor
ACE2	0.00E+00	0.00E+00	0.00E+00	1.0000
CXCL17	1.10E-21	7.53E-18	1.51E-17	0.9273
ABCA12	1.92E-21	8.77E-18	2.63E-17	0.9255
ATP10B	1.14E-20	3.92E-17	1.57E-16	0.9193
TMPRSS4	4.59E-20	1.26E-16	6.30E-16	0.9142
TMPRSS11E	7.91E-20	1.81E-16	1.09E-15	0.9121
CASP4	1.52E-19	2.98E-16	2.09E-15	0.9095
FUT3	3.20E-19	5.49E-16	4.39E-15	0.9065
GABRB3	4.99E-19	7.61E-16	6.85E-15	-0.9046
C14orf132	9.13E-19	1.20E-15	1.25E-14	-0.9020
TMPRSS11D	9.61E-19	1.20E-15	1.32E-14	0.9018
INHBB	1.27E-18	1.45E-15	1.75E-14	-0.9006
CDH13	1.89E-18	1.75E-15	2.60E-14	0.8988
MMP1	1.91E-18	1.75E-15	2.62E-14	0.8987
GNA15	1.91E-18	1.75E-15	2.63E-14	0.8987
DUOXA2	2.77E-18	2.38E-15	3.81E-14	0.8970
IL20RB	3.20E-18	2.59E-15	4.40E-14	0.8964
RPH3AL	3.70E-18	2.66E-15	5.08E-14	-0.8957
QSOX2	3.71E-18	2.66E-15	5.09E-14	-0.8957
LIMA1	3.87E-18	2.66E-15	5.32E-14	0.8955
ZNF750	4.47E-18	2.92E-15	6.14E-14	0.8948
NT5M	5.79E-18	3.62E-15	7.96E-14	-0.8936
HIP1	6.40E-18	3.82E-15	8.79E-14	-0.8931
TSPAN1	6.71E-18	3.84E-15	9.22E-14	0.8929
PIGR	8.73E-18	4.80E-15	1.20E-13	0.8916
CYP24A1	2.27E-17	1.15E-14	3.12E-13	-0.8868
ALDH3B1	2.30E-17	1.15E-14	3.16E-13	-0.8868
FOXA2	2.34E-17	1.15E-14	3.22E-13	-0.8867
HNF1A	2.78E-17	1.32E-14	3.82E-13	-0.8858
NMNAT1	3.09E-17	1.36E-14	4.25E-13	0.8852
DUOX2	3.14E-17	1.36E-14	4.31E-13	0.8852
PMP22	3.17E-17	1.36E-14	4.36E-13	-0.8851
TMEM121	3.58E-17	1.49E-14	4.92E-13	-0.8845
TUB	3.68E-17	1.49E-14	5.05E-13	-0.8843
BLNK	4.61E-17	1.81E-14	6.34E-13	0.8832
TRAF3IP2	6.20E-17	2.35E-14	8.52E-13	0.8816
PTHLH	6.32E-17	2.35E-14	8.69E-13	0.8815
WNT7B	8.74E-17	3.10E-14	1.20E-12	-0.8797
RNF39	8.80E-17	3.10E-14	1.21E-12	0.8797
ASB2	9.75E-17	3.35E-14	1.34E-12	0.8791
SPP1	1.02E-16	3.43E-14	1.41E-12	-0.8789
GPR68	1.19E-16	3.90E-14	1.64E-12	0.8780
PAQR5	1.23E-16	3.94E-14	1.69E-12	-0.8778
IGFBP4	1.33E-16	4.16E-14	1.83E-12	-0.8774
SAMD12	1.42E-16	4.32E-14	1.95E-12	0.8771
C2orf72	1.62E-16	4.64E-14	2.22E-12	-0.8763
PARP8	1.62E-16	4.64E-14	2.22E-12	0.8763
DDR1	1.63E-16	4.64E-14	2.24E-12	0.8763
NUP62CL	1.65E-16	4.64E-14	2.27E-12	-0.8762
PCDH7	1.87E-16	5.14E-14	2.57E-12	0.8755
C19orf33	1.91E-16	5.15E-14	2.63E-12	0.8754
TMEM229B	1.98E-16	5.24E-14	2.73E-12	0.8752
HNF4A	2.05E-16	5.32E-14	2.82E-12	-0.8750
MYRF	2.39E-16	6.07E-14	3.28E-12	-0.8741
SGCE	2.59E-16	6.46E-14	3.56E-12	-0.8737
IRAK3	2.63E-16	6.46E-14	3.62E-12	0.8736
CNNM3	2.80E-16	6.74E-14	3.84E-12	-0.8732
HK2	3.28E-16	7.76E-14	4.50E-12	0.8723
CRMP1	3.49E-16	8.06E-14	4.80E-12	-0.8719
IFI16	3.52E-16	8.06E-14	4.84E-12	0.8719
HLA-DOB	3.62E-16	8.15E-14	4.97E-12	0.8717
DAPP1	4.37E-16	9.57E-14	6.01E-12	0.8706
PCSK6	4.39E-16	9.57E-14	6.03E-12	-0.8706
CCDC102A	5.40E-16	1.15E-13	7.42E-12	-0.8694
S100A11	5.44E-16	1.15E-13	7.47E-12	0.8693
TRIM9	5.53E-16	1.15E-13	7.59E-12	-0.8692
SGK2	6.16E-16	1.26E-13	8.46E-12	-0.8686
EYA2	6.46E-16	1.30E-13	8.87E-12	0.8683
FANCC	6.54E-16	1.30E-13	8.98E-12	-0.8682
CHST10	6.71E-16	1.30E-13	9.22E-12	-0.8681
TMEM63C	6.76E-16	1.30E-13	9.28E-12	0.8680
ARNT2	6.83E-16	1.30E-13	9.39E-12	-0.8680
CSGALNACT1	7.14E-16	1.34E-13	9.81E-12	-0.8677

Figure 2. Most significantly (Bonferoni-adjusted p < 1E−11) correlated (red to yellow in last column) or anti-correlated (green) genes with ACE2. The transmembrane serine protease 4 (TMPRSS4) is one of the most significantly correlated (r = 0.9142, p = 4.59E−20) with ACE2 suggesting a major role in priming the SARS-CoV2 spike protein. CXCL17 (r = 0.9273, p = 1.1E−21), ABCA12 (r = 0.9256, p = 1−92E−21) and ATP10B (r = 0.9193, p = 1.14E−20) had marginally higher correlation with ACE2 while another transmembrane protease TMPRS11E (r = 0.9121, p = 7.91E−20) had slightly lower correlation.

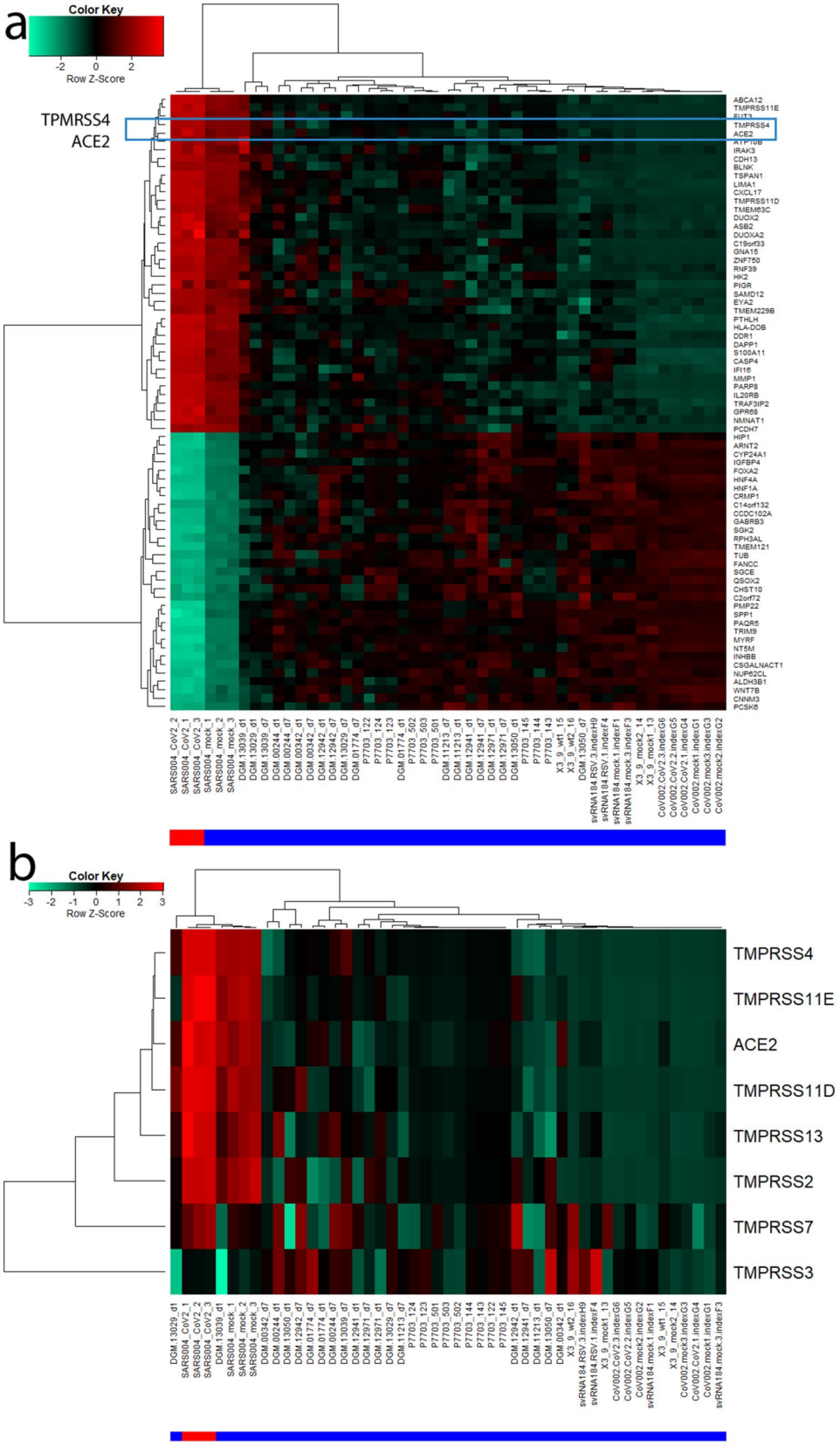

Figure 3. (**a**) Cluster analysis and gene expression heatmap of 72 most significantly (Bonferroni-adjusted $p < 1E{-}11$) correlated and anti.correlated genes with ACE2 shows close clustering of the serine protease TMPRSS4 with ACE2. (**b**) Heatmap of TMPRSS family members shows close clustering of TMPRSS11D/E and TMPRSS4 with ACE2 but also TMPRSS2 and TMPRSS13 have similar expression, particularly in SARS-CoV-2 infected samples. The red color bar indicates SARS-CoV-2, blue color bar control.

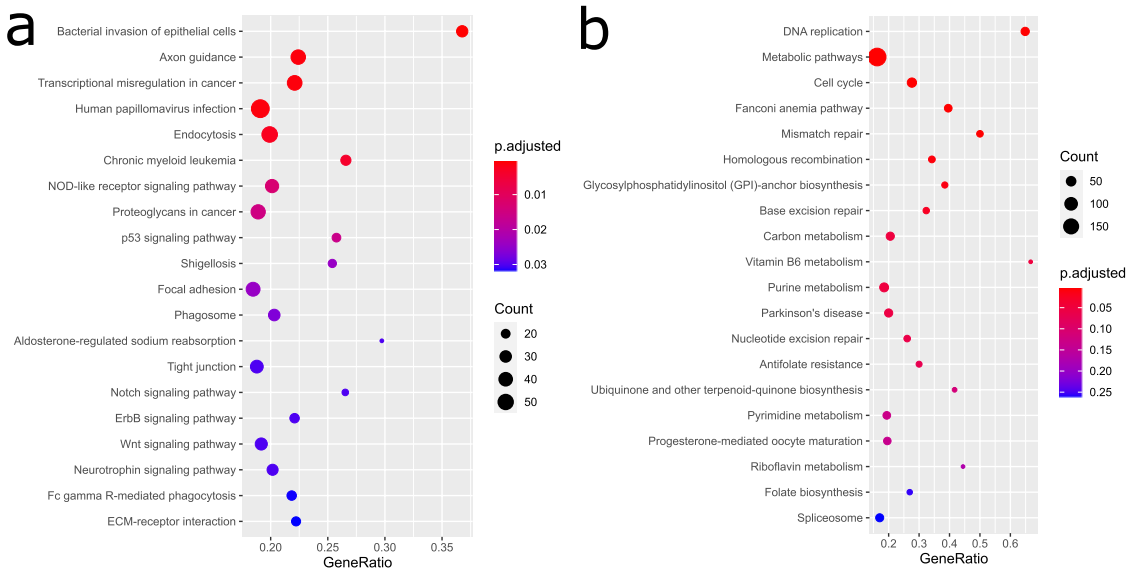

Figure 4. (**a**) The five most significantly overrepresented pathways correlated with ACE2 are *Human papillomavirus infection, Bacterial invasion of epithelial cells, Endocytosis, Axon Guidance* and *Transcriptional mis-regulation in cancer*. (**b**) The six most significantly overrepresented pathways anti-correlated with ACE2 are *DNA replication, Metabolic pathways, Cell cycle, Fanconi anemia pathway, Mismatch repair* and *Homologous recombination*. Many of these pathways are associated with DNA processing or repair. That these are down-regulated upon infection with SARS-CoV-2 is in line with reports about interferon and interferon stimulated genes (ISGs) inhibiting virus replication[16]. This would be a defense against the attempts of the virus to recruit the host's DNA repair and homologous recombination mechanisms as Gillespie et al. report[29].

GOs associated with inflammation and reactive oxygen species (ROS) amongst genes co-regulated with ACE2. Table 4 and Fig. 5c show GOs (all Biological Processes) related to inflammation and ROS over-represented in genes correlated with ACE2. The GO, *positive regulation of inflammatory response* (p = 0.0039) would imply that an inflammatory process is induced which finally leads to apoptosis—as demonstrated by the GO *inflammatory cell apoptotic process* (p = 0.042). The virus receptor ACE2 is involved in the *positive regulation of reactive oxygen species metabolic process* (p = 0.0185). Induction of ROS by a respiratory virus and subsequent inflammation has been reported Jamaluddin et al.[33].

Protein-interaction networks. We further restricted the set of ACE2-correlated or -anti-correlated genes by drastically filtering with Bonferoni-adjusted p < 1E−11 in order to construct a human readable protein interaction network of the most significant proteins (Fig. 6). The protein-interaction network generated from correlated and anti-correlated genes with *ACE2* shows *IFI16 (r = 0.8719), LIMA1 (r = 0.8955), CNNM3 (r = − 0.8732), HNF4A (r = − 0.8750), TRAF3IP2 (r = 0.8816), ASB2 (r = 0.8791)* and *FANCC* (r = − 0.8682) as hub genes (interactors from the BioGrid database are marked in red, original *ACE2*-correlated/anti-correlated genes are marked in green, hub genes and *ACE2* are highlighted with yellow shading). Interferon plays a major role in the host response to a virus and Thompson et al. reported that IFI16 – one of the hub genes in our network—controls the interferon response to DNA and RNA viruses[31]. Lin and Richardson review that *LIMA1* (formerly EPLIN) – which is known to enhance bundling of actin filaments[34]—mediates the interaction between Cadherins and Actin in the context of adherens junctions—playing a role in measles virus infection—via trans-binding with molecular interactors on adjacent cells[35]. In line with this, LIMA1 is connected with E-cadherin (CDH1) in the network and also Cadherin 13 (CDH13) is part of it and among the most significantly correlated genes to ACE2. The connection of ACE2 to calmodulin 1 (CALM1) is based on a publication by Lambert et al.[36] in which they show that CALM1 interacts with the corona virus receptor ACE2 and inhibits shedding of its ectodomain[36]. CALM1 inhibitors in turn can reverse this process so that the ACE2 ectodomain is shed, and is partially mediated by a metalloproteinase[36]. The direct connection between CALM1 and LIMA1 was found in a large-scale interactome study[37]. The involvement of the hub gene Fanconi anemia complementation group C (FANCC), although not experimentally proven, might reflect recruitment of DNA repair and homologous recombination mechanisms from the host by the virus.

Role of TMPRSS4 and other TMPRSS gene family members. A pivotal result of this meta-analysis is the transmembrane serine protease TMPRSS4 which was one of the most significantly correlated genes with ACE2. Additionally TMPRSS11E (r = 0.9121, Bonferoni-corrected p = 1.09E−15) and TMPRS11D (r = 0.9018, Bonferoni-corrected p = 1.32E−14) from the same gene family were found more significant than TMPRSS2 which however was still significantly correlated with ACE2 (r = 0.767, Bonferoni-corrected p = 1.79E−6). The assignment of the SARS-CoV-2 spike protein priming functionality to TMPRSS2 was based on the assumption that it would be the same as for its predecessor SARS-CoV[17] and was confirmed by several publications[19,20].

GO_BP term	P value	GO_CC term	P value	GO_MF term	P value
Regulation of primary metabolic process	5.23E−11	Intracellular	3.35E−27	Metal ion binding	4.98E−08
Regulation of cellular metabolic process	6.04E−09	Membrane−bounded organelle	8.95E−23	Protein binding	1.84E−05
Organic substance biosynthetic process	8.56E−07	Extracellular exosome	9.79E−18	Catalytic activity, acting on a protein	6.44E−03
Positive regulation of cellular process	2.09E−06	Extracellular organelle	1.21E−17	Metallopeptidase activity	4.99E−02
Negative regulation of cellular metabolic process	2.69E−05	Cytoplasm	2.07E−14		
Gene expression	3.24E−05	Vesicle	8.85E−14		
Interspecies interaction between organisms	3.30E−05	Ion binding	5.22E−08		
Cell junction assembly	1.20E−04	Membrane raft	2.03E−04		
Cytokine production	2.21E−04	Membrane region	5.95E−04		
Regulation of nitrogen compound metabolic process	3.90E−04	Extracellular region	1.33E−03		
Cellular component organization or biogenesis	4.77E−04	Plasma membrane	5.48E−03		
Positive regulation of viral process	4.86E−04	Whole membrane	8.94E−03		
Establishment of localization	6.03E−04				
Regulation of cell junction assembly	7.11E−04				
Amyloid precursor protein metabolic process	7.30E−04				
Signaling	7.61E−04				
Regulation of symbiosis, encompassing mutualism through parasitism	9.15E−04				
Organic substance transport	9.22E−04				
Regulation of viral life cycle	1.04E−03				
Regulation of biological quality	1.31E−03				
Regulation of cellular component organization	1.69E−03				
Vesicle-mediated transport	2.18E−03				
Cellular metabolic process	2.21E−03				
Cell–cell junction organization	2.53E−03				
Positive regulation of multicellular organismal process	2.74E−03				
Viral process	2.81E−03				
Endocytosis	3.08E−03				
Cellular nitrogen compound metabolic process	3.48E−03				
Positive regulation of cellular component biogenesis	5.04E−03				
Regulation of biological process	5.23E−03				
Catabolic process	6.34E−03				
Positive regulation of biological process	1.09E−02				
Entry into host cell	1.27E−02				
Entry into other organism involved in symbiotic interaction	1.27E−02				
Proteolysis	1.72E−02				
Positive regulation of reactive oxygen species metabolic process	1.85E−02				
Regulation of defense response	2.06E−02				
Regulation of multi-organism process	2.09E−02				
Regulation of multicellular organismal process	2.12E−02				
Regulation of viral entry into host cell	2.49E−02				
Gap junction assembly	2.51E−02				
Cell proliferation	2.52E−02				
Regulation of response to stress	2.57E−02				
Receptor biosynthetic process	2.68E−02				
Organonitrogen compound catabolic process	2.79E−02				
Macromolecule metabolic process	2.91E−02				
Response to stress	3.85E−02				
Regulation of cardiac muscle contraction	3.93E−02				
Zinc ion binding	4.07E−02				
Organonitrogen compound metabolic process	4.45E−02				

Table 2. Selected over-represented GOs in genes significantly correlated with ACE2.

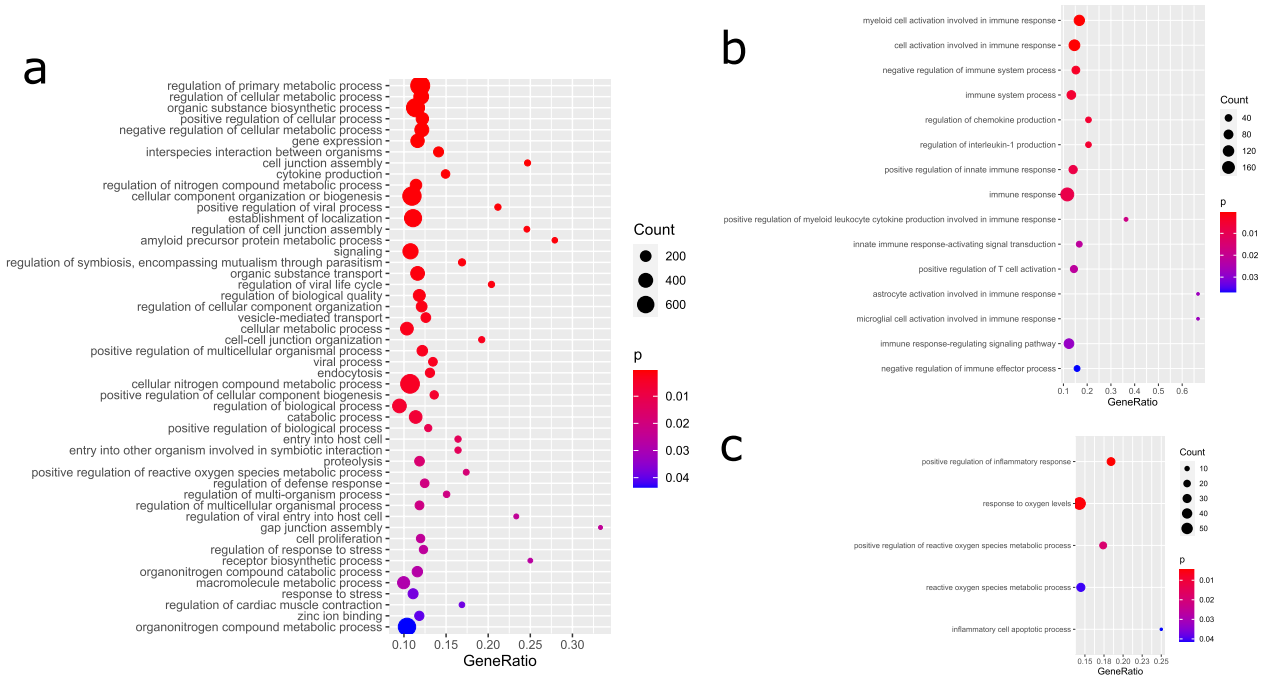

Figure 5. GO analysis reflects virus entry and immune response involving ROS and inflammation. (**a**) Selected GOs (Biological processes) shows interaction of virus and host, cell-junction organization, endocytosis, reaction involving cytokine production. (**b**) Immunity-related GOs illustrate the immune response involving activation of myeloid cells and T-cells and interleukin-1 and other chemokine production. (**c**) GOs associated with ROS and inflammation demonstrate involvement of ROS and inflammation leading to apoptosis.

Term	P value
Myeloid cell activation involved in immune response	5.3978E−07
Cell activation involved in immune response	4.3656E−05
Negative regulation of immune system process	0.00331363
Immune system process	0.0048604
Regulation of chemokine production	0.00486129
Regulation of interleukin-1 production	0.00486129
Positive regulation of innate immune response	0.0078574
Immune response	0.00855493
Positive regulation of myeloid leukocyte cytokine production involved in immune response	0.01809849
Innate immune response-activating signal transduction	0.02190672
Positive regulation of T cell activation	0.02209419
Astrocyte activation involved in immune response	0.02764374
Microglial cell activation involved in immune response	0.02764374
Immune response-regulating signaling pathway	0.0286071
Negative regulation of immune effector process	0.03757784

Table 3. GOs (all biological process) related to immune system in genes correlated with *ACE2*.

Term	P value
Positive regulation of inflammatory response	0.0039263
Response to oxygen levels	0.00438988
Positive regulation of reactive oxygen species metabolic process	0.01851891
Reactive oxygen species metabolic process	0.04088524
Inflammatory cell apoptotic process	0.04206722

Table 4. GOs (all biological process) related to inflammation and reactive oxygen species (ROS) in genes correlated with *ACE2*.

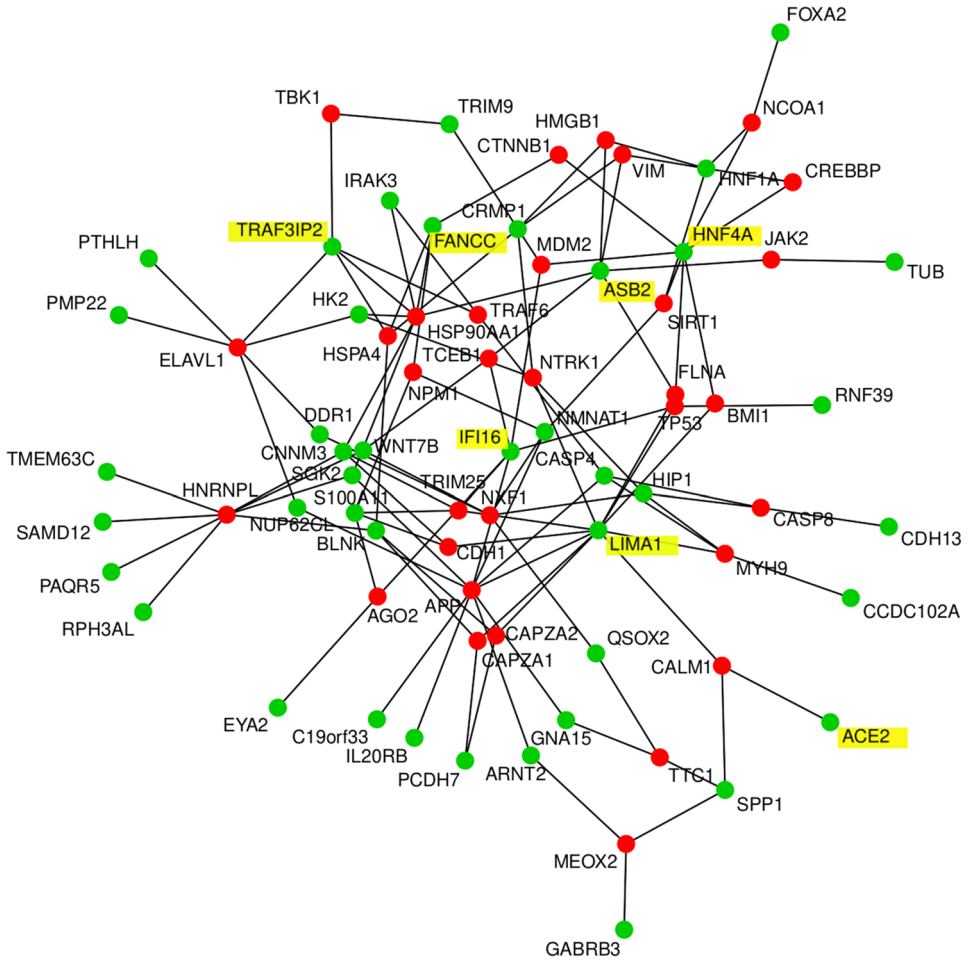

Figure 6. Protein interaction network of genes most significantly (Bonferoni-adjusted p < 1E−11) correlated and anti-correlated genes with ACE2 shows IFI16 (r = 0.8719), LIMA1 (r = 0.8955), CNNM3 (r = − 0.8732), HNF4A (r = − 0.8750), TRAF3IP2 (r = 0.8816), ASB2 (r = 0.8791) and FANCC (r = − 0.8682) as hub genes. Genes found as interactors in the BioGrid database are marked in red, the original geneset of *ACE2*-correlated genes is marked in green, hub genes and *ACE2* have yellow shading.

However, besides TMPRSS2 other proteases might be involved—see the heatmap in Fig. 3b which illustrates expression of transmembrane serine proteases. The highly significant co-expression of TMPRSS4 with ACE2 and its relevant role as transmembrane serine protease has enabled us to hypothesize that TMPRSS4 might also be involved in priming the SARS-CoV-2 spike protein. We therefore anticipate that inhibitors of TMPRSS4, TMPRS11D and TMPRS11E—besides those for TMPRSS2—could be a promising subject of further research.

Validation with a dataset of SARS-CoV-2 infected human bronchial organoids. We validated our results with the RNAseq dataset GSE150819 of human bronchial organoids infected with SARS-CoV-2[38]. Suppl. Table 8 shows the correlation of members of the TMPRSS gene family with ACE2 in this dataset. As in the first analysis of lung epithelial cells (Fig. 2) TMPRSS4 is also the gene most significantly correlated with ACE2 expression (p = 4.45E−05, Benjamini-Hochberg-corrected p = 0.0006, r = 0.86). Also TMPRSS2 (p = 0.014379, Benjamini-Hochberg-corrected p = 0.042562, r = 0.62) and TMPRSS11D (p = 0.000456, Benjamini-Hochberg-corrected p = 0.003185, r = 0.79) have significant p values and also Benjamini-Hochberg-corrected p values while TMPRSS11E here is not significant.

Discussion

In this meta-analysis we compared RNA-seq data of lung cells infected with SARS-CoV-2 and other lung cells with particular focus on correlated gene expression with the known SARS-CoV-2 receptor gene *ACE2*. We identified a signature of genes positively or negatively correlated with *ACE2* amongst which the most outstanding was the transmembrane serine protease *TMPRSS4*. In a recent publication Hoffmann et al.[17]. Inferred from the knowledge that the preceeding virus, SARS-CoV, uses ACE2 as receptor for entry and the serine protease TMPRSS2 for spike protein priming that the new virus SARS-CoV-2[17] would do the same. While the involvement of the receptor ACE2 appears to be established[3,16] the use of TMPRSS2 for spike protein priming appears not

fully settled as Hoffmann et al. still have to concede "that SARS-CoV-2 infection of Calu-3 cells was inhibited but not abrogated by camostat mesylate" (serine protease inhibitor with activity against TMPRSS2)[17]. The high significance (r = 0.9142, p = 4.59E−20) in our co-expression analysis with ACE2 suggests that TMPRSS4 is a considerable candidate for spike protein priming. That is in line with findings by Zang et al. who reported that TMPRSS4 besides TMPRSS2 enhances infection of small intestinal enterocytes with SARS-CoV-2[39]. However, TMPRSS4 is closely related to TMPRSS2 which both can proteolytically cleave the hemagglutinin of influenza viruses[40]. Further transmembrane serine proteases TMPRS11D (r = 0.9018, p = 9.61E−19) and TMPRS11E (r = 0.9121, p = 7.91E−20) in our analysis also emerged more significant than TMPRSS2 (r = 0.767, p = 1.3E−10). The TMPRS11 family member TMPRS11A was found to enhance viral infection with the first coronavirus SARS-CoV by spike protein cleavage in the airway[41]. Thus, we should not exclude the probability that other members of the TMPRSS gene family may be proteases for the SARS-CoV-2 spike protein. TMPRSS2 inhibitors have been proposed by Hoffmann et al.[17]—and earlier for the SARS-Cov virus by Kawase et al.[42] as working best together with cathepsin B/L inhibitors. We propose to investigate the effect of TMPRSS4 inhibitors in further research. As TMPRSS4 has been implicated in the invasion and metastasis of several cancers it has also been considered as target for cancer therapy for which a modest inhibitory effect of the above mentioned inhibitors in TMPRSS4-overexpressing SW480 cells was reported[43]. Interestingly, also *TMPRSS2* is connected with epithelial carcinogenesis as consistently over-expressed in prostate cancer[44], and later a gene fusion of *TMPRSS2* and *ERG* was reported as the predominant molecular subtype of prostate cancer[45], where *TMPRSS2* however only contributes untranslated sequence[46]. Assuming that co-expression with *ACE2* is an indication that TMPRSS4 may prime the SARS-CoV-2 spike protein we suggest that further development and testing of more effective TMPRSS4 inhibitors in in vitro and in vivo models could support the translation into clinical settings. However, we have to state the limitation that this study is a meta-analysis based solely on transcriptome and not proteome data.

Besides the identification of *TMPRSS4,* we found several significantly over-represented GOs and pathways such as Endocytosis, Papilloma virus infection and Bacterial invasion of epithelial cells for which we provide full gene lists to foster further elucidation of disease mechanisms. Genes from the constructed protein-interaction network provide a first snapshot of a comprehensive image: *IFI16* controls the interferon response to the virus[31], *LIMA1* mediating the interaction between Cadherins (*CDH1*, *CDH13*) and Actin in the context of adherens junctions potentially playing a role in virus infection, *CALM1* inhibits shedding of the ectodomain of the virus receptor *ACE2*[36].

Furthermore, GO analyses revealed several biological processes related to viral cell entry, host reaction, immune response, ROS, inflammation and apoptosis. This led us to propose a cascade of events taking place post SARS-CoV-2 entry into host cells- illustrated in Fig. 7 together with possible drug targets. The coronavirus SARS-CoV-2 docks at the receptor ACE2 on the membrane of the human epithelial cell, the early stage of infection. According to reports by Monteil et al. these processes can be blocked with recombinant hrsACE2[16]. Transmembrane serine proteases TMPRSS mediate SARS-CoV-2 cell entry via ACE2[17,47]. TMPRSS2 has been described as a mediator of ACE2-coupled endocytosis in the first SARS-CoV[48] and by a previous publication also for SARS-CoV-2[17]. However, we identified high levels of co-expression between ACE2 and TMPRSS4 and other members of the TMPRSS family and hypothesize that any of these additional family members might have the same function as the well described TMPRSS2. As a consequence, we propose that besides TMPRSS2 also other TMPRSS family members could be targets of pharmaceutical intervention warranting further research. After entering the cell, SARS-CoV-2 RNA is released, replicated and packaged again. Drugs can target the viral protease and the polymerase needed for replication[49]. Replication can further be inhibited by interferon and interferon-stimulated genes (ISG)[28] which we also found evidence for in negatively correlated replication pathways (e.g. DNA replication and homologous recombination). This depends on a healthy immune response and may be impaired in individuals with a weak immune system due to age or underlying diseases. The virus is packaged and released into the extracellular space where it can be attacked by macrophages chemo-attracted by CXCL17[24]. Also T-Cells can be involved in the immune response. We found evidence for their activation in GO analysis by associated interleukins IL1 and IL7. Although immunity is not the main focus of this manuscript it is tempting to speculate that the severity of the clinical manifestations such as the acute respiratory failure and also failure in other organs depend on the state of the immune system which decreases with age or diseases such as diabetes.The involvement of ACE2 in the renin-angiotensin system as antagonist of ACE in regulating blood pressure via Angiotensin II, vasoconstriction, dilation and its protective role against lung injury[41] are additional factors which correlate with age[50]. This is confirmed by reports from Wadmann et al.[51] about Centers for Disease Control and Prevention (CDC) data from 14 U.S. states that 50% hospitalized COVID-19 patients had pre-existing high blood pressure[51]. In their study about ACE2 in the preceeding SARS-CoV virus Imai et al.[21] found that *ACE2* protects against lung injury and acid-induced lung injury in a *Ace2*-knockout mouse can be improved by an inhibitor of the Angiotensin II receptor AT1[21]. The results of clinical studies but also statistics on hypertension and even more important statistics on pharmacological treatment of hypertension in COVID-19 patients may shed light on the discussions if treatment with ACE inhibitors and angiotensin receptor blockers (ARBs) are detrimental[52] or beneficial[53].

We conclude, that our meta-analysis of RNA-Seq data of lung cells partially infected with SARS-CoV-2 identified the transmembrane serine protease *TMPRSS4* as one of the most significantly correlated genes with the virus receptor *ACE2*. The importance of this finding is underlined by Zang et al. who simultaneously with our preprint publication reported that TMPRSS4 enhances SARS-CoV-2 infection of small intestinal enterocytes[39]. We propose that inhibitors of TMPRSS family members TMPRSS4, TMPRSS11D and TMPRSS11E besides TMPRSS2 are worthwhile testing in i*n vitro* and i*n vivo* studies. As clinicians, pathologists and scientists are struggling to get to grips with and understand the damage wrought by SARS-CoV-2 as it invades the body, we hope that our analyses and dataset will contribute to a better understanding of the molecular basis of COVID-19.

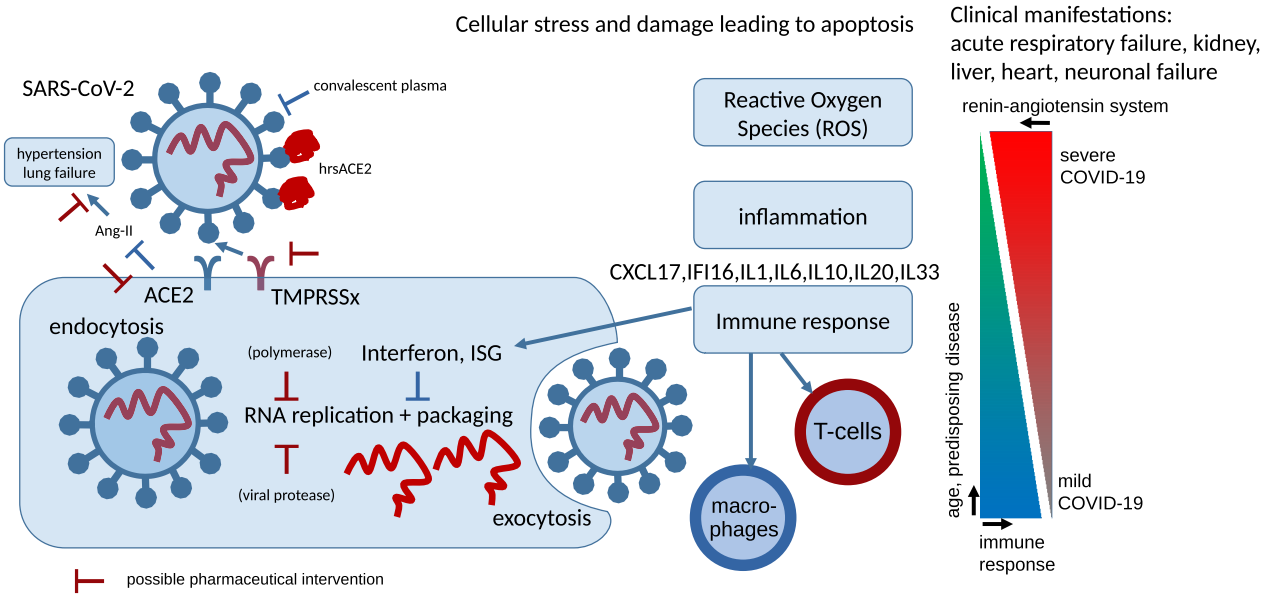

Figure 7. Scheme of SARS-CoV-2 infection. The coronavirus SARS-CoV-2 docks at the receptor ACE2 on the membrane of the human epithelial cell. Transmembrane serine proteases TMPRSSx mediate SARS-CoV-2 cell entry via ACE2. TMPRSS2 was reported for this in the first SARS-CoV and by previous publication also for SARS-CoV-2 but we hypothesize that due to co-expression with ACE2, TMPRSS4 and other members of the TMPRSS family may well perform this task. We suggest that inhibitors of TMPRSS4 and other TMPRSS family members might have therapeutic potential. Upon entry into the cell, viral RNA is released, replicated and packaged again. Replication can be inhibited by interferon and interferon stimulated genes (ISG) what we also saw in negatively correlated replication pathways (e.g. DNA replication and homologous recombination). This indicates a healthy immune response and may be impaired in persons with a weak immune system due to age or disease. The packaged virus is released from the cell and can be attacked by macrophages chemo-attracted by CXCL17—or T-cells for which we found evidence for by GO analysis and by associated interleukins IL1 and IL7. It is tempting to speculate that the severity of the clinical manifestations such as the acute respiratory failure and also failure in other organs depends on the quality of the immune system decreasing with age or diseases such as diabetes. The involvement of ACE2 in the renin-angiotensin system as antagonist of ACE in regulating blood pressure via Angiotensin II (Ang-II), vasoconstriction, dilation and its protective role against lung injury are additional factors which correlate with age[21,50,53].

Methods

Sample collection of lung cell RNA-Seq data. Next-generation sequencing datasets measured in RNA-Seq experiments with lung cells (GSE147507: Illumina NextSeq 500; GSE146482: Illumina NovaSeq 6000; GSE85121: Illumina HiSeq 2500) were downloaded from NCBI GEO (Table 1, suppl. Table 4). These datasets were provided along with studies by Blanco-Melo et al.[23] (GSE147507) and Staudt *et al.*[54] (GSE85121). A final publication related to the dataset GSE146482 is yet to materialize. From accession no. GSE85121 only small airway epithelium brushing cells were used but alveolar macrophages were excluded while from accession no. GSE147507 only human epithelial and adenocarcinoma lung cells were used but ferret cells and updates after March 24 were excluded and from accession no. GSE146482 only control epithelial cells were used but graphene oxide treated cells were excluded. After exclusion of the above mentioned datasets not fitting the target cell type, 49 samples remained useful.

Data normalization and analysis. After the excluded samples had been filtered from the downloaded RNA-Seq data data was imported into the R/Bioconductor environment[55,56]. Read counts from accesion nos. GSE147507 and GSE146482 were converted to FPKM (fragments per kilobase of exon model per million reads mapped) using trancript lengths downloaded from ENSEMBL (version GRCh38, p13). Batch effects were removed with the package sva[57] employing method ComBat[58]. Normalization was performed via the voom method[59]. Pearson correlation coefficients between samples were calculated with the R-builtin method *cor*. Dendrograms were drawn employing the dendextend package[60] filtering genes for high coefficient of variation above the 75% quantile.

The validation dataset GSE150819 of human bronchial organoids infected with SARS-CoV-2[38] downloaded from NCBI GEO was normalized with the voom method[59] before calculating correlation of each gene to ACE2 gene expression.

Detection of genes correlated with ACE2. The Pearson correlation of the normalized gene expression values for all samples was calculated between the gene *ACE2* and each other gene. The method cor.test was applied to calculate the p value for the t test for Pearson correlation. The p value was Bonferoni-corrected by divi-

sion by the number of genes and additionally adjusted via the Benjamini–Hochberg method[61]. Correlated genes were filtered with a very restrictive criterion (Bonferoni-adjusted-p < 1E−11)—in order to get human readable numbers of genes for heatmap and network generation- and conventional criteria (r > 0.6, Bonferoni-adjuste-p < 0.05) for positively correlated genes and (r < 0.6, Bonferoni-adjusted-p < 0.05) for negatively correlated genes.

Pathway and gene ontology (GO) analysis. The R package GOstats[62] was employed for over-representation analysis of positively and negatively correlated genes with the SARS-CoV-2 receptor gene *ACE2*. KEGG pathway annotations which had been downloaded from the KEGG database[27] in March 2018 were used for testing over-representation of the positively and negatively ACE2-correlated genes via the R-builtin hypergeometric test.

Dot plots showing the p value of the hypergeometric test, the ratio of significant genes to all genes in the pathway and the number of significant genes per pathway were plotted via the package *ggplot2*[63].

Protein-interaction networks. A human protein interaction network was constructed in a similar manner as we described in our previous publication[64]. However, here only direct interactors and no further interactors of interactors were extracted from the Biogrid database version 3.4.161[65] using the restrictively filtered (Bonferoni-adjusted-p < 1E−1) genes significantly correlated or anti-correlated with ACE2 gene expression. The network was reduced to the n = 30 nodes with most interactions and was plotted via the R package *network*[66] showing original genes in green and BioGrid-derived interactors in red.

Data availability

No datasets were generated during the current study. The datasets used for the meta-analysis are available at the National Center for Biotechnology Information (NCBI) Gene expression Omnibus (GEO) accessions referred to in Table 1.

Received: 2 June 2020; Accepted: 17 November 2020
Published online: 08 December 2020

References

1. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020).
2. Zhou, H. *et al.* A Novel Bat Coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* **30**(2196), e3-2203.e3 (2020).
3. Zhang, T., Wu, Q. & Zhang, Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr. Biol.* **30**(1346), e2-1351.e2 (2020).
4. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
5. Tilocca, B. *et al.* Comparative computational analysis of SARS-CoV-2 nucleocapsid protein epitopes in taxonomically related coronaviruses. *Microbes Infect.* **22**, 188–194 (2020).
6. Tilocca, B. *et al.* Immunoinformatic analysis of the SARS-CoV-2 envelope protein as a strategy to assess cross-protection against COVID-19. *Microbes Infect.* **22**, 182–187 (2020).
7. Tilocca, B. *et al.* Molecular basis of COVID-19 relationships in different species: a one health perspective. *Microbes Infect.* **22**, 218–220 (2020).
8. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* https://doi.org/10.1016/S1473-3099(20)30120-1 (2020).
9. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* https://doi.org/10.1016/S1473-3099(20)30243-7 (2020).
10. Wang, Y. *et al.* Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *The Lancet* https://doi.org/10.1016/S0140-6736(20)31022-9 (2020).
11. Wang, M. *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271 (2020).
12. Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J. & Hsueh, P.-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. *Int. J. Antimicrob. Agents* **55**, 105924 (2020).
13. Yamamoto, N., Matsuyama, S., Hoshino, T. & Yamamoto, N. Nelfinavir inhibits replication of severe acute respiratory syndrome coronavirus 2 in vitro. *BioRxiv*. https://doi.org/10.1101/2020.04.06.026476 (2020)
14. Beigel, J. H. *et al.* Remdesivir for the Treatment of Covid-19—final report. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2007764 (2020).
15. Casadevall, A. & Pirofski, L. The convalescent sera option for containing COVID-19. *J. Clin. Invest.* **130**, 1545–1548 (2020).
16. Monteil, V. *et al.* Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell* https://doi.org/10.1016/j.cell.2020.04.004 (2020).
17. Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* https://doi.org/10.1016/j.cell.2020.02.052 (2020).
18. Matsuyama, S. *et al.* Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc. Natl. Acad. Sci.* **117**, 7001–7003 (2020).
19. Sungnak, W. *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020).
20. Collin, J. *et al.* Co-expression of SARS-CoV-2 entry genes in the superficial adult human conjunctival, limbal and corneal epithelium suggests an additional route of entry via the ocular surface. *Ocul. Surf.* https://doi.org/10.1016/j.jtos.2020.05.013 (2020).
21. Imai, Y. *et al.* Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature* **436**, 112–116 (2005).
22. Haga, S. *et al.* Modulation of TNF- -converting enzyme by the spike protein of SARS-CoV and ACE2 induces TNF- production and facilitates viral entry. *Proc. Natl. Acad. Sci.* **105**, 7809–7814 (2008).
23. Blanco-Melo, D. *et al.* Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**(1036), e9-1045.e9 (2020).
24. Maravillas-Montero, J. L. *et al.* Cutting edge: GPR35/CXCR8 is the receptor of the mucosal chemokine CXCL17. *J. Immunol. Baltim. Md* **1950**(194), 29–33 (2015).

25. Thomas, A. C., Tattersall, D., Norgett, E. E., O'Toole, E. A. & Kelsell, D. P. Premature terminal differentiation and a reduction in specific proteases associated with loss of ABCA12 in Harlequin ichthyosis. *Am. J. Pathol.* **174**, 970–978 (2009).
26. Wilk, E. *et al.* RNAseq expression analysis of resistant and susceptible mice after influenza A virus infection identifies novel genes associated with virus replication and important for host resistance to infection. *BMC Genomics* **16**, 655 (2015).
27. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595 (2019).
28. Schoggins, J. W. & Rice, C. M. Interferon-stimulated genes and their antiviral effector functions. *Curr. Opin. Virol.* **1**, 519–525 (2011).
29. Gillespie, K. A., Mehta, K. P., Laimins, L. A. & Moody, C. A. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. *J. Virol.* **86**, 9520–9526 (2012).
30. Zhang, Y. *et al.* Clinical and coagulation characteristics of 7 patients with critical COVID-2019 pneumonia and acro-ischemia. *Zhonghua Xue Ye Xue Za Zhi Zhonghua Xueyexue Zazhi* **41**, E006 (2020).
31. Thompson, M. R. *et al.* Interferon γ-inducible protein (IFI) 16 transcriptionally regulates type I interferons and other interferon-stimulated genes and controls the interferon response to both DNA and RNA viruses. *J. Biol. Chem.* **289**, 23568–23581 (2014).
32. Chan, M. C. *et al.* Tropism and innate host responses of a novel avian influenza A H7N9 virus: an analysis of ex-vivo and in-vitro cultures of the human respiratory tract. *Lancet Respir. Med.* **1**, 534–542 (2013).
33. Jamaluddin, M., Tian, B., Boldogh, I., Garofalo, R. P. & Brasier, A. R. Respiratory syncytial virus infection induces a reactive oxygen species-MSK1-phospho-ser-276 RelA pathway required for cytokine expression. *J. Virol.* **83**, 10605–10615 (2009).
34. Meng, W. & Takeichi, M. Adherens junction: molecular architecture and regulation. *Cold Spring Harb. Perspect. Biol.* **1**, a002899 (2009).
35. Lin, L.-T. & Richardson, C. The host cell receptors for measles virus and their interaction with the viral hemagglutinin (H) protein. *Viruses* **8**, 250 (2016).
36. Lambert, D. W., Clarke, N. E., Hooper, N. M. & Turner, A. J. Calmodulin interacts with angiotensin-converting enzyme-2 (ACE2) and inhibits shedding of its ectodomain. *FEBS Lett.* **582**, 385–390 (2008).
37. Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).
38. Suzuki, T. *et al.* Generation of human bronchial organoids for SARS-CoV-2 research. *biorxiv*. https://doi.org/10.1101/2020.05.25.115600 (2020).
39. Zang, R. *et al.* TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* https://doi.org/10.1126/sciimmunol.abc3582 (2020).
40. Bertram, S., Glowacka, I., Steffen, I., Kühl, A. & Pöhlmann, S. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. *Rev. Med. Virol.* **20**, 298–310 (2010).
41. Kam, Y.-W. *et al.* Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro. *PLoS ONE* **4**, e7870 (2009).
42. Kawase, M., Shirato, K., van der Hoek, L., Taguchi, F. & Matsuyama, S. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J. Virol.* **86**, 6537–6545 (2012).
43. de Aberasturi, A. L. & Calvo, A. TMPRSS4: an emerging potential therapeutic target in cancer. *Br. J. Cancer* **112**, 4–8 (2015).
44. Bugge, T. H., Antalis, T. M. & Wu, Q. Type II transmembrane serine proteases. *J. Biol. Chem.* **284**, 23177–23181 (2009).
45. Tomlins, S. A. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648 (2005).
46. Tomlins, S. A. *et al.* Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia N. Y. N* **10**, 177–188 (2008).
47. Meng, T. *et al.* The insert sequence in SARS-CoV-2 enhances spike protein cleavage by TMPRSS. *biorxiv*. https://doi.org/10.1101/2020.02.08.926006 (2020).
48. Glowacka, I. *et al.* Differential downregulation of ACE2 by the spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus NL63. *J. Virol.* **84**, 1198–1205 (2010).
49. Huang, J., Song, W., Huang, H. & Sun, Q. Pharmacological therapeutics targeting RNA-dependent RNA polymerase, proteinase and spike protein: from mechanistic studies to clinical trials for COVID-19. *J. Clin. Med.* **9**, 1131 (2020).
50. Conti, S., Cassis, P. & Benigni, A. Aging and the renin-angiotensin system. *Hypertension* **60**, 878–883 (2012).
51. Wadman, M., Couzin-Frankel, J., Kaiser, J. & Matacic, C. A rampage through the body. *Science* **368**, 356–360 (2020).
52. Fang, L., Karakiulakis, G. & Roth, M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection?. *Lancet Respir. Med.* **8**, e21 (2020).
53. Vaduganathan, M. *et al.* Renin–angiotensin–aldosterone system inhibitors in patients with Covid-19. *N. Engl. J. Med.* **382**, 1653–1659 (2020).
54. Staudt, M. R., Salit, J., Kaner, R. J., Hollmann, C. & Crystal, R. G. Altered lung biology of healthy never smokers following acute inhalation of E-cigarettes. *Respir. Res.* **19**, 78 (2018).
55. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
56. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
57. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinforma. Oxf. Engl.* **28**, 882–883 (2012).
58. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat. Oxf. Engl.* **8**, 118–127 (2007).
59. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
60. Galili, T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720 (2015).
61. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
62. Falcon, S. & Gentleman, R. Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–258 (2007).
63. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer, Berlin, 2009).
64. Wruck, W. & Adjaye, J. Meta-analysis of human prefrontal cortex reveals activation of GFAP and decline of synaptic transmission in the aging brain. *Acta Neuropathol. Commun.* **8**, 26 (2020).
65. Chatr-aryamontri, A. *et al.* The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **45**, D369–D379 (2017).
66. Butts, C. T. Network: a package for managing relational data in R. *J. Stat. Softw.* **24**, 1–24 (2008).

Acknowledgments

James Adjaye acknowledges support from the Medical faculty of the Heinrich-Heine University, Duesseldorf.

Author contributions

W.W. analysed the data and wrote the manuscript. J.A. supervised the work, co-wrote the manuscript and gave the final approval.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-78402-2.

Correspondence and requests for materials should be addressed to J.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.