Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Assessing the consequences of environmental exposures on the expression of the human receptor and proteases involved in SARS-CoV-2 cell-entry

Manon Watzky¹, Maud de Dieuleveult¹, Anne Letessier¹, Claude Saint-Ruf¹, Benoit Miotto⁷

Université de Paris, Institut Cochin, INSERM, U1016, CNRS, UMR8104, F-75014, PARIS, France

ARTICLE INFO

Keywords: Coronavirus SARS-CoV-2 COVID-19 ACE2 TMPRSS2 FURIN Cathepsin genes Environmental exposure

ABSTRACT

The role of environmental condition on the infection by the novel pathogenic SARS-CoV-2 virus remains uncertain. In here, exploiting a large panel of publicly available genome-wide data, we investigated whether the human receptor ACE2 and human proteases TMPRSS2, FURIN and CATHEPSINS (B, L and V), which are involved in SARS-CoV-2 cell entry, are transcriptionally regulated by environmental cues. We report that more than 50 chemicals modulate the expression of ACE2 or human proteases important for SARS-CoV-2 cell entry. We further demonstrate that transcription factor AhR, which is commonly activated by pollutants, binds to the promoter of TMPRSS2 and enhancers and/or promoters of Cathepsin B, L and V encoding genes. Our exploratory study documents an influence of environmental exposures on the expression of genes involved in SARS-CoV-2 cell entry. These results could be conceptually and medically relevant to our understanding of the COVID-19 disease, and should be further explored in laboratory and epidemiologic studies.

1. Introduction

Coronavirus disease 2019 (or COVID-19) is caused by a coronavirus named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Oberfeld et al., 2020; Zheng, 2020). While a majority of infected individuals have minimal symptoms around 15–20% need to be hospitalised with severe respiratory complications (Cevik et al., 2020; Del Sole et al., 2020; Wolff et al., 2020). Based on first reports, several factors such as sex, age and underlying diseases (e.g. diabetes, auto-immune diseases or congenital heart diseases) were documented to cause a more severe course of the disease, and increase likelihood of death (Gallo Marin et al., 2020; Kopel et al., 2020; Wolff et al., 2020). Similar features and comorbidities were already observed for other respiratory diseases caused by coronaviruses such as the severe human acute respiratory syndrome (SARS-CoV) and the Middle East respiratory syndrome (MERS-CoV) (da Costa et al., 2020; Petrosillo et al., 2020).

It is also observed that some environmental factors cause a higher risk of infection by SARS-CoV-2 as well as to develop more severe symptoms of COVID-19. For instance, cigarette smoke increases the risk to develop severe symptoms of the disease, while, in contrast, epidemiologic reports show that smokers are under-represented in COVID-19 patients, possibly due to the protective role of nicotine against SARS-Cov-2 infection (Brake et al., 2020; Cai et al., 2020; Cai, 2020; Leung et al., 2020; Lippi and Henry, 2020; Olds and Kabbani, 2020). Several studies also report that atmospheric pollution, including particulate matters (PM2.5 and PM10) and high levels of nitrogen dioxide in the air are correlated with a higher incidence of COVID-19 and more severe forms of the disease (Bianconi et al., 2020; Boloano-Ortiz et al., 2020; Borro et al., 2020; Comunian et al., 2020; Conticini et al., 2020; Copat et al., 2020; Ogen, 2020; Yao et al., 2020; Zoran et al., 2020). These data raise the issue of environmental exposures as a confounding factor in SARS-CoV-2 infection and disease course.

It has been demonstrated that SARS-CoV-2 uses the human angiotensin-converting enzyme-2 (ACE2) receptor for infection, as already observed for other coronaviruses, including SARS-CoV and HCoV-NL63 (Hoffmann et al., 2020b; Lan et al., 2020; Letko et al., 2020; Shang et al., 2020a, 2020b; Wallis et al., 2020; Yan et al., 2020). In vitro, soluble ACE2 is sufficient to block SARS-CoV-2 infection in reconstituted human tissues (Monteil et al., 2020). Conversely, expression of human ACE2 in mouse tissues and its over-expression in human cells refractory to SARS-CoV-2 infection confers SARS-CoV-2 susceptibility (Bao et al., 2020; Ou et al., 2020; Winkler et al., 2020). It was further demonstrated that the SARS-CoV-2 spike (S) glycoprotein directly binds the receptor ACE2 (Hoffmann et al., 2020b; Lan et al., 2020; Shang et al., 2020b; Yan et al., 2020).
et al., 2020). ACE2 is expressed in many tissues including lung, liver, colon, oesophagus, intestine, heart and kidney tissues which is coherent with the variety of cell types that can be infected and might explain the variety of symptoms reported in COVID-19 patients (Cevik et al., 2020; Gkogkou et al., 2020; Hijmert et al., 2020; Hoffmann et al., 2020b; Li et al., 2020; Lukassen et al., 2020; Oberfeld et al., 2020; Qi et al., 2020; Zou et al., 2020).

The cellular entry of SARS-CoV-2 also depends on the priming (i.e. cleavage) of its transmembrane spike (S) glycoprotein by host cell proteases to promote virus-cell entry. Different proteases are important for SARS-CoV-2 cell entry, notably the transmembrane serine protease-2 (TMPRSS2) (Hoffmann et al., 2020b; Wrapp et al., 2020). It was observed that human cells over-expressing TMPRSS2 are more susceptible to coronaviruses infection than control cells, and that SARS-CoV-2 infection can be partially blocked by treatment with inhibitors of TMPRSS2 protease activity (Hoffmann et al., 2020b; Matsuyma et al., 2020; Ou et al., 2020; Shang et al., 2020a). It was also reported that Cathepsins lysosomal cysteine proteases have a major role in the virus cell entry. Chemical inhibition of Cathepsin L endosomal cysteine proteases activity (encoded by CTSL and CTSL2 also known as CTSV), and to a lesser extent inhibition of Cathepsin B (CTSB activity, impairs SARS-CoV-2 cell entry (Hoffmann et al., 2020b; Ou et al., 2020; Shang et al., 2020a; Zou et al., 2016). In addition, it was also shown that the subtilisin-like endoprotease FURIN cleaves the multibasic cleavage site (Arg-Arg-Ala-Arg) at the junction of the two subunits of the SARS-CoV-2 spike protein, favouring virus entry (Coutard et al., 2020; Hoffmann et al., 2020a; Jaimes et al., 2020; Shang et al., 2020a; Walle et al., 2020). This list is probably not exhaustive, as numerous investigations are still ongoing to properly catalogue the full spectrum of host cell proteases and pathways that have an effect on SARS-CoV-2 cell entry in different cell types.

We postulated that identifying and reporting environmental factors which possibly altered the expression of ACE2 and human proteases involved in the proteolytic processing of the S protein of SARS-CoV-2, might help to better understand the susceptibility to viral infection. We thus exploited large publicly available and freely re-useable gene expression and genomic datasets to support the hypothesis of a potential impact of environmental exposures on the expression of the genes ACE2, TMPRSS2, FURIN, CTSL, CTSV (also known as CTSL2) and CTSB, that in turn could influence virus entry into host cells.

2. Materials and methods

2.1. Genomic and tissue expression resources

Genomic and interactions data were analysed using the UCSC Genome Browser (Lee et al., 2020) and the GeneHancer database (Fishilevich et al., 2017). Genomic data were analysed on the human reference genome GRCh37/hg19. Interactions between GeneHancer regulatory elements (promoters and enhancers) and genes reported in the GeneHancer database were collected from different sources, including expression quantitative trait loci (eQTL) from the genotype-tissue expression (GTEx) project (version v6p), chromosome conformation capture (capture Hi-C) long range interactions, FANTOM eRNA-gene expression correlation or distance based associations (Fishilevich et al., 2017).

2.2. Gene expression datasets and analysis

Gene expression datasets were downloaded through the Gene Expression Omnibus (GEO) database (Barrett et al., 2013). The datasets downloaded and re-analysed are the following: 1) Bronchus from intra-pulmonary airways (the right upper lobe carina), from normal smoking and non-smoking volunteers (including 34 current smokers, 23 nonsmokers and 18 former smokers) (GSE994) (Spira et al., 2004); 2) Bronchial epithelial cells from 21 former smokers and 9 current smokers (GSE112073) (Corbett et al., 2019); 3) Bronchial epithelium of fiberoptic bronchoscopy from 21 never smokers, 52 current smokers and 31 former smokers (GSE7895) (Beane et al., 2007); 4) Small airway epithelium from 38 nonsmokers and 45 smokers obtained by fiberoptic bronchoscopy (GSE11952) (Hübner et al., 2009); 5) Bronchial epithelial cells exposed to cigarette smoke from two brands of cigarettes (n = 4) (GSE10700) (Jorgensen et al., 2008).

Differentially expressed genes were determined using the GEO2R software using p-value correction with false-discovery rate correction (Benjamini & Hochberg) (Barrett et al., 2013). q-values (FDR corrected p-values) were calculated from a genome-wide analysis and fold change between treated and matching control condition (i.e. reported as log2 of the Fold Change). These values were used to prepare the different panels. A q-value < 0.05 was considered statistically significant.

2.3. Comparative toxicogenomics database search and analysis

The Comparative Toxicogenomics Database (CTD) (http://ctdbase.org/) reports manually curated chemical–gene, chemical–phenotype and chemical–exposure interactions for 15,681 chemicals and 46,689 genes (Davis et al., 2019). The database was interrogated on April 14, 2020 using the search function with gene name (i.e. ACE2, CTSB, CTSL, CTSV, FURIN, IQGAP1, MX2, PIR, TMPRSS2 or UNCG45) and human genome (Taxon:9606) as queries. The list of chemical–gene interactions obtained was then manually curated to remove studies involving mixture of chemicals, functional relationships and analyses of protein levels (Davis et al., 2019). Chemicals associated with ACE2, CTSB, CTSL, CTSV, FURIN, IQGAP1, MX2, PIR, TMPRSS2 and UNCG45 gene expression up-regulation or down-regulation were further analysed and intersected. Chemicals associated with the regulation of AhR, either at mRNA or protein level, were also retrieved from the database on April 14, 2020.

2.4. Identification of AhR consensus binding motif

Coordinates of promoters were retrieved from the GeneHancer database (Fishilevich et al., 2017). Coordinates were uploaded on the Galaxy interface (https://usegalaxy.org/) and DNA sequences corresponding to these coordinates were retrieved in batch in FASTA format using the function “Extract genomic DNA”. Motif search on these DNA sequences was then performed using the EMBOSS Fuzznuc function with the following criteria: no mismatch allowed to the searched pattern and a search on the Watson and Crick strand. We investigated the AhR/ARNT core sequence 5′-GCGTG-3′ (Jaspur MOTIF Ahr: Arnt_MA0006.1) embedded in the xenobiotic responsive elements (XREs) consisting of T/CNGCGTGA/CG/CA (N = any nucleotide) or a divergent AhR element 5′-CATG(N6)CT/ATG-3′ (Dere et al., 2011; Jackson et al., 2015; Murray et al., 2014).

2.5. ChIP-sequencing analysis of AhR binding sites in human cells

AhR ChIP-seq data were retrieved from the GEO database: GEO GSE90550 and GEO GSE116632 (Neavin et al., 2019; Yang et al., 2018).

GSE90550 reports AhR ChIP-seq experiments performed using an AhR antibody (Santa Cruz, H-211) in triplicate, starting from MCF7 breast cancer cells treated with 10 nM TCDD or vehicle for 45 min (Yang et al., 2018).

GSE116632 reports AhR ChIP-seq experiments performed using an AhR antibody (Cell Signalling Technology, D56H4) in duplicate, starting from a lymphoblastoid cell line (GM17212) treated with 1 μM 3-methylcholanthrene or vehicle for 24 h (Neavin et al., 2019).

Environmental Research 195 (2021) 110317

M. Wankyy et al.

Environmental Research 195 (2021) 110317

M. Wankyy et al.

Environmental Research 195 (2021) 110317

M. Wankyy et al.

Environmental Research 195 (2021) 110317

M. Wankyy et al.

Environmental Research 195 (2021) 110317

M. Wankyy et al.

Environmental Research 195 (2021) 110317

M. Wankyy et al.
3. Results

3.1. Genomic and 3D organisation of the loci encoding the human receptor and proteases involved in SARS-CoV-2 cell entry

We thought that expression levels of ACE2, TMPRSS2, FURIN, CTSB, CTSL and CTSV might be impacted by the expression level of other genes they physically interact with within the nucleus. To get a better understanding of these potential regulatory mechanisms we analysed genomic and chromosome conformation capture data in human cells and listed genes and regulatory sequences that interact with ACE2, TMPRSS2, FURIN, CTSB, CTSL and CTSV.

ACE2 is located on the short arm of chromosome X and it escapes X inactivation in female cells. In the GeneHancer database (Fishilevich et al., 2017), it is reported that the promoter of ACE2 is involved in long-distance interactions with the promoter of PIR, a transcriptional coregulator of the NF-kappa-B pathway (Liu et al., 2013), and an enhancer located in intron 9 of ACE2 (Fig. 1A and Supplementary Table S1). Importantly, higher expression of PIR is recurrently detected in lung biopsies from smokers compared to nonsmokers indicating a regulation of this gene by cigarette smoke. It was also shown that PIR up-regulation may be associated with lung epithelial cell damages and eventually cell death (Gellman et al., 2007).

TMPRSS2 is located on chromosome 21 in the vicinity of MX Dynamin Like GTPase 1 (MX1) and MX Dynamin Like GTPase 2 (MX2) (Fig. 1B and Supplementary Table S1). In the GeneHancer database, TMPRSS2 promoter interacts with the promoter region of MX2. MX2 encodes an interferon-alpha inducible GTPase with antiviral activity against immunodeficiency virus type 1 (HIV1) and vesicular stomatitis virus (Haller et al., 2015; Kane et al., 2013, 2018; Melen et al., 1996; Pavlovic et al., 1990). An alternative promoter of TMPRSS2, upstream of the transcription start site (TSS), mediates interactions with enhancers located in intronic regions of TMPRSS2 (introns 6, 7 and 8) as well as with its canonical promoter (Supplementary Table S1).

FURIN (also known as FES upstream region) is located on chromosome 15. In the GeneHancer database, we found that the promoter of FURIN interacts with promoters of the FES Proto-Oncogene Tyrosine Kinase (FES), the Ras GTPase-activating-like protein IQGAP1 (IQGAP1) and Unc-45 Myosin Chaperone A (UNC45A) (Fig. 1C and Supplementary Table S1). FES, IQGAP1 and UNC45A have been described for their role in cell cycle progression and cell migration, and their expression altered in several cancers (Eisa et al., 2019; Johnson et al., 2009; Sodroski et al., 1984).

CTSB is located on chromosome 8 while CTSL and CTSV are located on chromosome 9. In the GeneHancer database, there is no information regarding potential 3D conformational interactions between the
promoters of cathepsins and promoters of other genes (Supplementary Table S1).

Our analysis indicates that promoters of ACE2 and TMPRSS2 interact through long range interactions with promoters of PIR and MX2 genes, which are regulated by environmental cues, such as cigarette smoke and viral infection. We also observed that FURIN promoter interacts with many genes important for cell architecture and cancer signalling. We thus considered the entire list of genes to document a possible relationship between environmental cues and the expression of human genes encoding the receptor and proteases of SARS-CoV-2 spike S glycoprotein.

3.2. A data mining approach shows that expression of genes encoding receptor and proteases involved in SARS-CoV-2 cell entry is modulated by a variety of chemicals

We queried the comparative toxicogenomics database (CTD) to identify studies reporting changes in expression levels of human genes encoding receptor and proteases important for SARS-CoV-2 cell entry upon chemical exposure (Davis et al., 2019). We selected for our analysis any study reporting gene expression in human tissues (or cells) and with negative or positive regulation of mRNA expression. We considered chemical components found in the environment, as well as chemicals used in medicine or industry. Using these criteria, we identified several chemical exposures regulating the expression of receptor ACE2 and human proteases TMPRSS2, FURIN and Cathepsins genes in human cells and tissues (Fig. 2; Supplementary Table S2). Notably, this approach confirms that PIR expression is enhanced by tobacco smoke pollution (Supplementary Table S2).

We unveiled several chemicals found in the environment modulating the expression of at least two genes analysed in our study (Fig. 2). Three pollutants can increase ACE2 expression as well as human proteases of SARS-CoV-2 S protein: phenylmercuric acetate (widely used as a fungicide and a disinfectant), nickel (a metal used in industry) and pentanal (an alkyl aldehyde used in industry). Phenylmercuric acetate is the only chemical reported in the database that enhances the expression of ACE2 and TMPRSS2, the main SARS-CoV-2 receptor/protease couple identified in in vitro assays, suggesting a potential interference with...
SARS-CoV-2 infection. We also noticed that several compounds, including bisphenol A, cyclophosphamide, tobacco smoke and particulate matter increase the expression of \textit{CTSL}, a key regulator of SARS-CoV-2 cell entry by endocytosis. Finally, we also noted that Benzo(a) pyrene, a prototypical polycyclic aromatic hydrocarbon compound, modulates \textit{FURIN} and \textit{TMPRSS2} expression (Fig. 2).

We also identified several chemicals used in medicine as modulator of \textit{ACE2} and human proteases of SARS-CoV-2 expression (Fig. 2). Quite noticeable is the impact of valproic acid, a chemical used for the treatment of mood disorders (Dong et al., 2008) on the expression of 10 out of the 11 genes analysed (Fig. 2). Valproic acid is a histone deacetylase inhibitor that also acts on DNA methylation patterns and hence affects a large number of gene targets, including most gene analysed here. Also intriguing, is the down-regulation of \textit{ACE2}, \textit{CTSL} and \textit{CTSV} expression in cells treated with the anti-cancer tyrosine-kinase inhibitor sunitinib (Fig. 2) (Cabebe and Wakelee, 2006). We also observed that cyclosporine, a modulator of the immune response utilised for the treatment of several auto-immune diseases (e.g. psoriasis, atopic dermatitis) and in organ transplant protocols (Talal, 1988), impacts on the expression level of all the known human proteases involved in the processing of the SARS-CoV-2 spike glycoprotein, \textit{CSTB}, \textit{CTSL}, \textit{CTSV}, \textit{FURIN} and \textit{TMPRSS2} (Fig. 2).

---

**Fig. 3.** Expression level of genes encoding \textit{ACE2} and proteases of SARS-CoV-2 S protein in lung biopsies from smokers and non-smokers. (A) Gene expression in lung biopsies from smokers compared to nonsmokers and former smokers. *, $p < 0.05$; ***, $p < 1e-5$ (vs non or former smokers). (B) Gene expression in bronchial epithelial cells exposed to smoke from a brand of regular cigarette (2R4F) and a brand of typical “light” cigarette. *, $p < 0.05$ (vs mock treated).
In sum, our analysis reveals that >50 chemicals can potentially affect the expression of human receptor ACE2 and human proteases involved in SARS-CoV-2 cell entry, which may modulate the susceptibility to SARS-CoV-2 infection.

3.3. Cigarette smoke does not modulate the expression of human receptor and proteases of SARS-CoV-2 S protein in the respiratory tract

Cigarette smoke is a confounding factor in many lung pathologies and many studies suggest it may impact on COVID-19 prognosis (Leung et al., 2020). We thus further investigated whether cigarette smoke may regulate the expression of the receptor ACE2 and proteases involved in SARS-CoV-2 cell entry that may modulate the susceptibility to infection. The analysis of the CTD database revealed that expression of CTSL and CTSV is increased upon cigarette smoke exposure in bronchial epithelial cells and monocytes and CTSL expression is also increased in 3D-bronchial tissues (Anthérieu et al., 2017; Ishikawa et al., 2019; Wright et al., 2012). We thus further investigated the level of expression of the different genes in the respiratory tract tissues of smokers, non-smokers and former smokers. We first analysed gene expression from intra-pulmonary airways obtained from 34 current, 23 never and 18 former smokers (GSE994) (Spira et al., 2004). We observed significant higher levels of PIR expression in smokers’ biopsies compared to non-smokers and former smokers biopsies, while other genes studied are expressed at similar levels (Fig. 3A). In three additional cohorts (GSE11952, GSE7895 and GSE112073) we also observed similar results (Fig. 3A) (Beane et al., 2007; Corbett et al., 2019; Hübner et al., 2009). These data indicate that among the different genes studied, solely PIR is significantly regulated by cigarette smoke. To further assess these data, we re-analysed gene expression data from bronchial epithelial cells exposed in vitro to cigarette smoke (GSE10700) (Jorgensen et al., 2008). Importantly, in vitro, cigarette smoke also causes an up-regulation of PIR expression in bronchial epithelial cells after 24 h exposure, consistent with a direct effect of cigarette smoke on PIR expression (Fig. 3B). Other genes were not differentially expressed between smoke and control treated cells (Fig. 3B). Our analysis suggests that expression of human receptor and proteases of SARS-CoV-2 spike S glycoprotein are not directly regulated by cigarette smoke at the transcriptional level in lung cells, nor that PIR and ACE2 are co-regulated upon cigarette smoke exposure.

3.4. Many aryl hydrocarbon receptor (AhR) agonists regulate the expression of ACE2 and human proteases of SARS-CoV-2 S protein

We thought to identify the molecular pathways that link chemicals with COVID-19 disease. Intriguingly, we identified several chemicals known as agonist of the aryl hydrocarbon receptor (AhR) (Murray et al., 2014). Among the 53 chemicals regulating the expression of ACE2 and proteases important for SARS-CoV-2 cell entry, 30 were associated with the function and/or regulation of AhR (P (X = 30) = 1.58e-23; enrichment: 9.54) based on the CTD nomenclature. For instance, exposure to particulate matters (PM), well-known agonists of AhR (Castaneda et al., 2018; Harmon et al., 2018), reduces CTSB expression in endothelial cells (Gong et al., 2007) and enhances CTSL expression in dermal fibroblasts (Park et al., 2018) and bronchial Beas-2B cells (Grilli et al., 2018). Piling up evidences indicate that exposure to particulate matters, ranging from PM2.5 and PM10, is associated with greater COVID-19 disease severity and poor clinical outcome (Comunian et al., 2020; Yao et al., 2020; Zoran et al., 2020). We also identified, benzo(a)pyrene, a polycyclic aromatic hydrocarbon, that decreases FURIN expression in HepG2 liver cancer cells (Lizzaraga et al., 2012) and TMPRSS2 expression in human prostate and liver cancer cells (Ihrubá et al., 2011; Kreuzer et al., 2020). Tetrachlorodibenzo-dioxin increases TMPRSS2 expression in MCFC breast cancer cells (Spink et al., 2009). Calcitriol, the physiologically active form of vitamin D, increases TMPRSS2 expression in human breast cancer tissues (Sheng et al., 2016).

These data suggest AhR signalling is an important player in the transcriptional regulation of human receptor and proteases of SARS-CoV-2 S protein, and that AhR signalling may impact on cell infection by SARS-CoV-2.

3.5. Analysis of aryl hydrocarbon receptor (AhR) binding sites in human cells suggests that TMPRSS2 is a direct transcriptional target of chemicals

We investigated whether AhR binds regulatory sequences involved in the regulation of human ACE2 and proteases of SARS-CoV-2 S protein. Upon binding of its ligand, AhR forms a heterodimer with the AhR nuclear translocator (ARNT), forming a DNA binding complex that translocates to the nucleus and binds and activates the transcription of target genes. AhR/ARNT binds the core sequence 5′-GGTGTT-3′ (Jaspar MOTIF Ahr:Arom_MA0006.1) embedded in the xenobiotic responsive elements (XREs) consisting of T/CNGGCGA/CG/CA (N = any nucleotide) or a divergent AhR element 5′-CATG(N6)CT/ATG-3′ (Dere et al., 2011; Jackson et al., 2015; Murray et al., 2014). We thus performed a search for these AhR consensus binding motifs in the promoters of ACE2, CTSL, CTSL, FURIN and TMPRSS2 using coordinates from the GeneHancer database (since no promoter segment is reported in the database for CTSV we did not include this gene in our analysis). We identified consensus binding sites for AhR in the promoters of ACE2, CTSL, CTSL, FURIN and TMPRSS2 (Supplementary Table S3). These data suggest that AhR might directly regulate the expression of these genes. Nonetheless, the frequency of AhR motifs in the promoter of the genes of interest is not significantly higher than the number of expected AhR motif in promoters of the GeneHancer database (Supplementary Table S3).

To support a role of AhR in the regulation of ACE2, CTSL, CTSL, FURIN and TMPRSS2 we thus searched for chromatin immunoprecipitation data for AhR in primary human cells or human tissues. Unfortunately, we could not find maps of AhR binding sites in “normal” cells, but identified two maps of AhR performed in MCF7 breast cancer cells 45 min after addition of the well known AhR pathway activator 2,3,7,8-tetrachlorodibenzo-p-dioxine (TCDD) and in lymphoblastoid cells 24 h after addition of AhR agonist 3-methylcholanthrene (Neavin et al., 2019; Yang et al., 2018). The analysis in MCF7 cells indicates that AhR binding is enhanced at the promoter of TMPRSS2 and CTSL upon TCCD treatment (Fig. 4A–F). This analysis also shows that AhR binding is enhanced upon TCCD exposure to the promoter of MX2 and enhancers of TMPRSS2 and CTSL, and in the vicinity of CTSL (Fig. 4A–F).

The analysis in lymphoblastoid cells indicates that AhR binds to the promoters, or enhancers, of CTSL, MX1, MX2, FURIN, IQGAP1, FES and UNCA5A in untreated condition (Supplementary Figure 1). Upon 3-methylcholanthrene treatment, AhR binding is still detected at these sites (Supplementary Figure 1). Figure fig1

Our analysis revealed that TMPRSS2 and CTSL are direct targets of AhR in human cells upon AhR-signalling activation. These later analyses provide molecular support for a direct link between chemical exposure and SARS-CoV-2 infection.

4. Discussion

4.1. Chemicals as confounding factors in SARS-CoV-2 infection and disease

Refinement of the COVID-19 course is still ongoing and our report provides evidence that the mRNA expression of the human receptor ACE2 and proteases, which are important for SARS-CoV-2 entry into the cells, is modulated by environmental cues. This information may help better understand the molecular basis of the susceptibility to SARS-CoV-2 and other coronaviruses infection, the course of the disease and help design future epidemiologic studies by considering the exposition to chemicals revealed in this study. Proper molecular and cellular assays will nonetheless be needed to properly assess and validate the
Fig. 4. AhR binding in ACE2, TMPRSS2, FURIN, CTSV and CTSB regions upon TCDD exposure. Schematic representation from the UCSC Genome browser (reference human genome GRC37/hg19) highlighting the regions of (A) ACE2 (chrX:15,380,000–15,630,000), (B) TMPRSS2 (chr21:42,678,000–42,910,000), (C) FURIN (chr15:90,860,000–91,507,000), (D) CTSIL (chr9:90,315,000–90,350,000), (E) CTSV (chr9:99,750,000–99,850,000) and (F) CTSB (chr8:11,600,000–11,780,000). The panels show from the top to bottom: the genomic coordinates on chromosome, the position of the genes and their orientations retrieved from the NCBI RefSeq Select database (one representative transcript per protein-coding gene), the position of the regulatory elements with enhancers represented in grey and promoters represented in red from GeneHancer Double Elite database, the position of the AhR binding sites enhanced upon TCDD exposure highlighted by orange boxes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
correlation described in this study in different cell types, reconstituted tissues, organoids and human samples.

We identify numerous chemicals, either pollutants from the environment or used for the treatment of chronic diseases, as potential regulators of ACE2 and human proteases of SARS-CoV-2 spike glycoprotein in human tissues. Higher levels of TMPRSS2 expression favour human cell infection in vitro, while expression of human ACE2 renders mice susceptible to SARS-CoV-2 infection (Bao et al., 2020; Matsuyama et al., 2020; Monteil et al., 2020; Winkler et al., 2020). These observations suggest that chemicals that promote higher expression levels of human receptor ACE2 and proteases in tissues might increase the risk of infection. This interpretation is obviously simplistic and detailed epidemiologic and clinical studies are needed to refine these data and address their robustness and reproducibility in the human population. Individuals treated for chronic pathologies and disorders, or chronically exposed to chemicals in their life, may present altered expression levels of ACE2 and proteases mediating SARS-CoV-2 cell entry compared to the rest of the population, which may account for the differences in the susceptibility to infection as well as the disease outcome between individuals. For instance, medical issues have emerged regarding the use of modulators of the immune system such as methotrexate, anti-TNF therapies or cyclosporine (Russell et al., 2020). Studies also proposed that particulate matters or high levels of nitrogen dioxide in the air may be favouring more severe forms of COVID-19 (Comunian et al., 2020; Conticini et al., 2020; Ogen, 2020; Yao et al., 2020). In support of these observations, our study identifies several compounds, frequently detected in the airaira, as potential regulators of ACE2 and proteases of SARS-CoV-2 S protein that might modulate the efficiency of viral infection. Importantly, we did not find a clear relationship between ACE2 expression levels and smoking habits in the cohorts we re-analysed. This is in contrast to other studies based on immunofluorescence assays (Brake et al., 2020; Cai et al., 2020) or showing that cigarette smoke is associated with higher levels of ACE2 expression in aged male (>65 years old) but not in female nor in younger male (Leung et al., 2020). Our analysis merely investigated based on publicly available data a potential relationship between environmental exposure (i.e. single exposure) and expression of receptor ACE2 and human proteases important for SARS-CoV-2 cell entry. Further work will be needed to properly assess these cellular data in the human population, and in response to more complex “exposomes”.

In addition, it remains to be investigated how these pollutants interfere with additional steps of SARS-CoV-2 biology, including its replication, its dissemination in the organism and how these environmental and chemical cues influence the course and outcome of COVID-19. For instance, AhR signalling appears as a key player in response to mouse coronavirus infection and to Zika virus infection (Giovannoni et al., 2020; Grunewald et al., 2020). It is thus likely that AhR signalling activity might contribute to the pathology of SARS-CoV-2. Furthermore, in vitro chemical screens identified numerous compounds that impair SARS-CoV-2 replication in cellular models, at different stages of the virus life cycle (Bojkova et al., 2020; Sauvat et al., 2020; Touret et al., 2020).

4.2. AhR signalling might contribute to COVID-19 progression

Analysis of two different maps of AhR binding sites in the human genome indicates that TMPRSS2, FURIN and CTSB are direct transcriptional targets of AhR signalling. Upon TCDD exposure, our analysis reveals enhanced binding of AhR at the promoter of TMPRSS2, consistent with a direct relationship between chemical exposure and TMPRSS2 expression. We also observed that promoters of FURIN and CTSB are bound by AhR in lymphocytes, but that AhR agonist 3-MC does not enhance the binding. This suggests that different agonists trigger different levels of AhR signalling and/or that the binding of AhR to promoters of receptor ACE2 and proteases of SARS-CoV-2 S protein is differentially modulated by AhR agonists in different cell types.

Importantly, while ACE2 and TMPRSS2 promoters interact with promoters of PIR and MX2, that are massively up-regulated upon smoking and viral infection respectively, we did not observe a co-regulation of these genes in our analyses. This co-regulation might nonetheless occur under conditions that need to be properly and exhaustively investigated such as mixture of compounds and more complex environmental signals.

In sum, our study provides preliminary evidences that receptor ACE2 and human proteases important for SARS-CoV-2 cell entry are regulated by many chemicals linked with AhR signalling, which should prompt further analysis of the relationship between “exposome” and SARS-CoV-2 pathology.

Author contribution

Manon Watzky, Maud de Dieuleveult, Anne Letessier, Claude Saint-Ruf and Benoît Miotto, Conceptualization, Data acquisition, Formal analysis, Methodology, Drafting the manuscript. Benoit Miotto, Funding acquisition.

Funding

Laboratory of B.M. is partner of Labex “Who Am I?” #ANR-11-LABX-0071 and the Université de Paris IdEx #ANR-18-IDEX-0001 funded by the French Government through its “Investments for the Future” program. This work was funded by Fondation pour la Recherche Médicale (AEJ20151234749), Labex “Who am I?” (ANR-11-LABX-0071 and ANR-11-IDEX-005-02), CNRS, INSERM and University of Paris. MW is supported by a PhD fellowship from the French Government. MdB is supported by a post-doctoral fellowship from Labex “Who Am I?” #ANR-11-LABX-0071.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Dr Morgane Bomsel (INSERM U1016, Paris, France) for discussions during the course of this study. We thank research laboratories and consortiums for making their studies publicly available and freely re-usable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2020.110317.

References

Anthérès, S., Garat, A., Beauval, N., Suez, M., Allorge, D., Garçon, G., Lo-Guidec, J.-M., 2017. Comparison of cellular and transcriptomic effects between electronic cigarette vapor and cigarette smoke in human bronchial epithelial cells. Toxicol. In Vitro 45, 417–425. https://doi.org/10.1016/j.tiv.2016.12.015.
Bao, L., Deng, W., Huang, B., Hao, G., Liu, J., Ren, L., Wei, Q., Yu, P., Xu, Y., Qi, F., Qu, Y., Li, F., Lv, Q., Wang, W., Xue, J., Gong, S., Liu, M., Wang, G., Wang, S., Song, Z., Zhao, Linna, Liu, P., Zhao, Li, Ye, F., Wang, H., Zhou, W., Zhu, N., Zhen, W., Yu, H., Zhang, X., Guo, L., Chen, L., Wang, C., Wang, Y., Wang, X., Xiao, Y., Sun, Q., Liu, H., Zhu, F., Ma, C., Yan, L., Yang, M., Han, J., Xu, W., Tan, W., Peng, X., Jin, Q., Wu, G., Qin, C., 2020. The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. Nature. https://doi.org/10.1038/s41586-020-2312-y.
Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets-update. Nucleic Acids Res. 41, D991–D995. https://doi.org/10.1093/nar/gkt1193.
Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Beane, J., Sebastiani, P., Liu, G., Brody, J.S., Lenburg, M.E., Spira, A., 2007. Reversible

Bocca, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., Münch, C., Bojkova, D., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., Münch, C.,

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., Münch, C.,

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., Münch, C.,

Beane, J., Sebastiani, P., Liu, G., Brody, J.S., Lenburg, M.E., Spira, A., 2007. Reversible

Chen, Y., Xiao, F., Kheradmand, F., Amos, C.I., 2020. Tobacco smoking increases

Conticini, E., Frediani, B., Caro, D., 2020. Can atmospheric pollution be considered a co-

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.

Cai, H., 2020. Sex difference and smoking predisposition in patients with COVID-19.
