Genomics-guided tracing of SARS-CoV-2 targets in human cells identifies Vitamin D and Quercetin as candidate medicinal agents for mitigation of the severity of pandemic COVID-19.

Gennadi V. Glinsky[1]

[1] Institute of Engineering in Medicine

University of California, San Diego

9500 Gilman Dr. MC 0435

La Jolla, CA 92093-0435, USA

Correspondence: gglinskii@ucsd.edu

Web: http://iem.ucsd.edu/people/profiles/guennadi-v-glinskii.html

Running title: Genomics-guided mitigation maps for coronavirus pandemic

Key words: COVID-19; SARS-CoV-2 coronavirus; genomics; mitigation approaches; drugs & medicinal substances repurposing; Vitamin D; Quercetin; Luteolin; Eriodictyol.

Abstract

Genes required for SARS-CoV-2 entry into human cells, *ACE2* and *FURIN*, were employed as baits to build genomics-guided molecular maps of up-stream regulatory elements, their expression and functions in human body, including pathophysiologically-relevant cell types. Repressors and activators of the *ACE2* and *FURIN* genes were identified based on the analyses of gene silencing and overexpression experiments as well as relevant transgenic mouse models. Panels of repressors (*VDR; GATA5; SFTPC; HIF1a*) and activators (*HMGA2; INSIG1*) were then employed to identify existing drugs manifesting gene expression signatures of the potential coronavirus infection mitigation agents. Using this strategy, Vitamin D and Quercetin have been identified as putative COVID-19 mitigation agents. Gene expression profiles of Vitamin D and Quercetin activities and their established safety records as over-the-counter medicinal substances suggest that they may represent viable candidates for further considerations of their potential utility as COVID-19 pandemic mitigation agents. Quercetin has been identified as one of top-scoring candidate therapeutics in the supercomputer SUMMIT drug-docking screen and Gene Set Enrichment Analyses (GSEA) of expression profiling experiments (EPEs), indicating that highly similar structurally Quercetin, Luteolin, and Eriodictyol could serve as scaffolds for development of efficient inhibitors of the SARS-CoV-2 infection. In agreement with this notion, Quercetin alters expression of 98 of 332 (30%) of human genes encoding protein targets of SARS-CoV-2, thus potentially interfering with functions of 23 of 27 (85%) of the SARS-CoV-2 viral proteins in human cells. Similarly, Vitamin D may interfere with functions of 19 of 27 (70%) of the SARS-CoV-2 proteins by altering expression of 84 of 332 (25%) of human genes encoding protein targets of SARS-CoV-2. Considering the potential effects of both Quercetin and Vitamin D, the inference could be made that functions of 25 of 27 (93%) of SARS-CoV-2 proteins in human cells may be altered. GSEA and EPEs identify multiple drugs, smoking, and many disease conditions, including seasonal

and pandemic H1N1, that appear to act as putative coronavirus infection-promoting agents. Discordant patterns of Testosterone versus Estradiol impacts on SARS-CoV-2 targets suggest a plausible molecular explanation of the apparently higher male mortality during coronavirus pandemic. Of major concern is the ACE2 and FURIN expression in many human cells and tissues, including immune cells, suggesting that SARS-CoV-2 coronavirus may infect a broad range of cellular targets in the human body. Infection of immune cells may cause immunosuppression, long-term persistence of the virus, and spread of the virus to secondary targets. Present analyses and numerous observational studies indicate that age-associated Vitamin D deficiency may contribute to high mortality of older adults and elderly. Immediate availability for targeted experimental and clinical interrogations of potential COVID-19 pandemic mitigation agents, namely Vitamin D and Quercetin, as well as of the highly selective ($K_{i,}$ 600 pm) intrinsically-specific FURIN inhibitor (a1-antitrypsin Portland (a1-PDX), is considered an encouraging factor. Observations reported in this contribution are intended to facilitate follow-up targeted experimental studies and, if warranted, randomized clinical trials to identify and validate therapeutically-viable interventions to combat the COVID-19 pandemic.

Introduction

Coronavirus pandemic COVID-19 caused by the newly emerged SARS-CoV-2 virus is rapidly transitioning through the dangerous acute phase of its evolution in the United States. Absence of the vaccine and lack of efficient targeted therapeutic approaches emphasizes the urgent need for identification of candidate pandemic mitigation agents among existing drugs and medicinal substances.

SARS-CoV-2 virus was discovered in December 2019 and shortly thereafter it was isolated and sequenced (Zhou et al., 2020; Zhu et al., 2020). Recent analyses of the structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein revealed the key role of the *ACE2* and *FURIN* genes in facilitating the high-affinity binding of viral particles and their entry into human cells (Walls et al., 2020). The efficient invasion of host cells by the SARS-CoV-2 is further enhanced by the presence of the unexpected furin cleavage site, which is cleaved during the biosynthesis (Walls et al., 2020). This novel feature distinguishes the previously known SARS-CoV and the newly emerged SARS-CoV-2 viruses and possibly contributes to the expansion of the cellular tropism of the SARS-CoV-2 (Walls et al., 2020). The crystal structure and high-resolution cryo-electron microscopy of the SARS-CoV-2 receptor-binding domain (RBD) in complex with human ACE2 revealed specific structural features of the SARS-CoV-2 RBD that appear to enhance its binding affinity to human ACE2 (Shang et al., 2020; Yan et al., 2020). Collectively, these observations firmly established protein products of the human genes *ACE2* and *FURIN* as the principal mediators of the SARS-CoV-2 invasion into human cells acting as the high-affinity receptor (ACE2) and invasion-promoting protease (FURIN), respectively.

In this contribution, genomic screens were performed employing the *ACE2* and *FURIN* genes as baits to build genomics-guided human tissues-tailored maps of up-stream regulatory elements, their expression and functions. To identify the high-priority list of potential candidate

mitigation agents, the validation analyses were performed using gene silencing and overexpression experiments as well as relevant transgenic mouse models with the emphasis on pathophysiologically-relevant cell types. Panels of repressors (*VDR; GATA5; SFTPC*) and activators (*HMGA2; INSIG1*) of the *ACE2* and *FURIN* expression were identified and then employed to identify existing drugs and medicinal substances that could be repurposed to ameliorate the outcomes of the coronavirus infection. Two of the most promising candidate mitigation agents, namely Vitamin D and Quercetin, manifest gene expression-altering activities and have established safety records as over-the-counter medicinal substances that seem sufficient for further assessment and considerations of their potential utility for amelioration of the clinical course of coronavirus pandemic. Collectively, observations reported in this contribution indicate that highly similar structurally Quercetin, Luteolin, and Eriodictyol could serve as scaffolds for development of efficient inhibitors of the SARS-CoV-2 infection. Unexpectedly, present analyses revealed discordant patterns of Testosterone versus Estradiol impacts on SARS-CoV-2 targets with the former manifesting the potential coronavirus infection-promoting activities, which is consistent with the apparently higher male mortality across all age groups during the coronavirus pandemic. Significantly, consistent with findings reported herein numerous observational studies suggest that age-associated Vitamin D insufficiency and/or deficiency may contribute to high mortality of older adults and elderly individuals during the COVID-19 pandemic. Consequently, the Vitamin D supplementation may mitigate the severity of the disease.

Results and Discussion

Enrichr-guided gene set enrichment analyses (GSEA) of genomic features associated with the *ACE2* and *FURIN* genes.

One of the goals of this work was to identify human genes implicated in regulatory cross-talks affecting expression and functions of the *ACE2* and *FURIN* genes to build a model of genomic

regulatory interactions potentially affecting the SCARS-CoV-2 coronavirus infection. To this end, the Enrichr bioinformatics platform was utilized (see Methods; Chen et al., 2013; Kuleshov et al., 2016) at the initial stage of the analyses and the identified records and top-scoring candidate features were further interrogated using targeted evaluation of the publicly-available records of the Gene Expression Omnibus (GEO) database. GSEA were carried out using the *ACE2* and *FURIN* genes as baits applied to a broad spectrum of genomic databases reflecting the current state of knowledge regarding the structural, functional, regulatory, and pathophysiological features that could be statistically linked to these genes. Expression profiling experiments and GSEA revealed ubiquitous patterns of both *ACE2* and *FURIN* genes across human tissues (Supplemental Figure S1) with notable examples of high expression of the *FURIN* gene in the lung (second-ranked tissue in the GTEX database) and testis being identified as the top-ranked *ACE2*-expressing tissue. In addition to the human lung tagged by the *ACE2* expression in the ACRHS4 Human Tissues database search, other noteworthy significantly enriched records are the Peripheral Blood Mononuclear Cells (PBMC), Natural Killer Cells and Macrophages tagged by the *FURIN* expression (Supplemental Figure S1).

GSEA of the virus perturbations' data sets among Gene Expression Omnibus (GEO) records of up-regulated genes identified the SARS-CoV challenge at 96 hrs (GSE47960) as the most significantly enriched record (Supplemental Figure S2) tagged by expression of both *ACE2* and *FURIN* in human airway epithelial cells. These observations suggest that coronavirus infection triggers the increased expression of both *ACE2* and *FURIN* genes 4 days after the initial encounter with host cells (Figure 1; Supplemental Figure S2). These findings were corroborated by the increased *FURIN* expression documented in the PBMC of patients with severe acute respiratory syndrome (Figure 1; Supplemental Figure S1; Reghunathan et al., 2005).

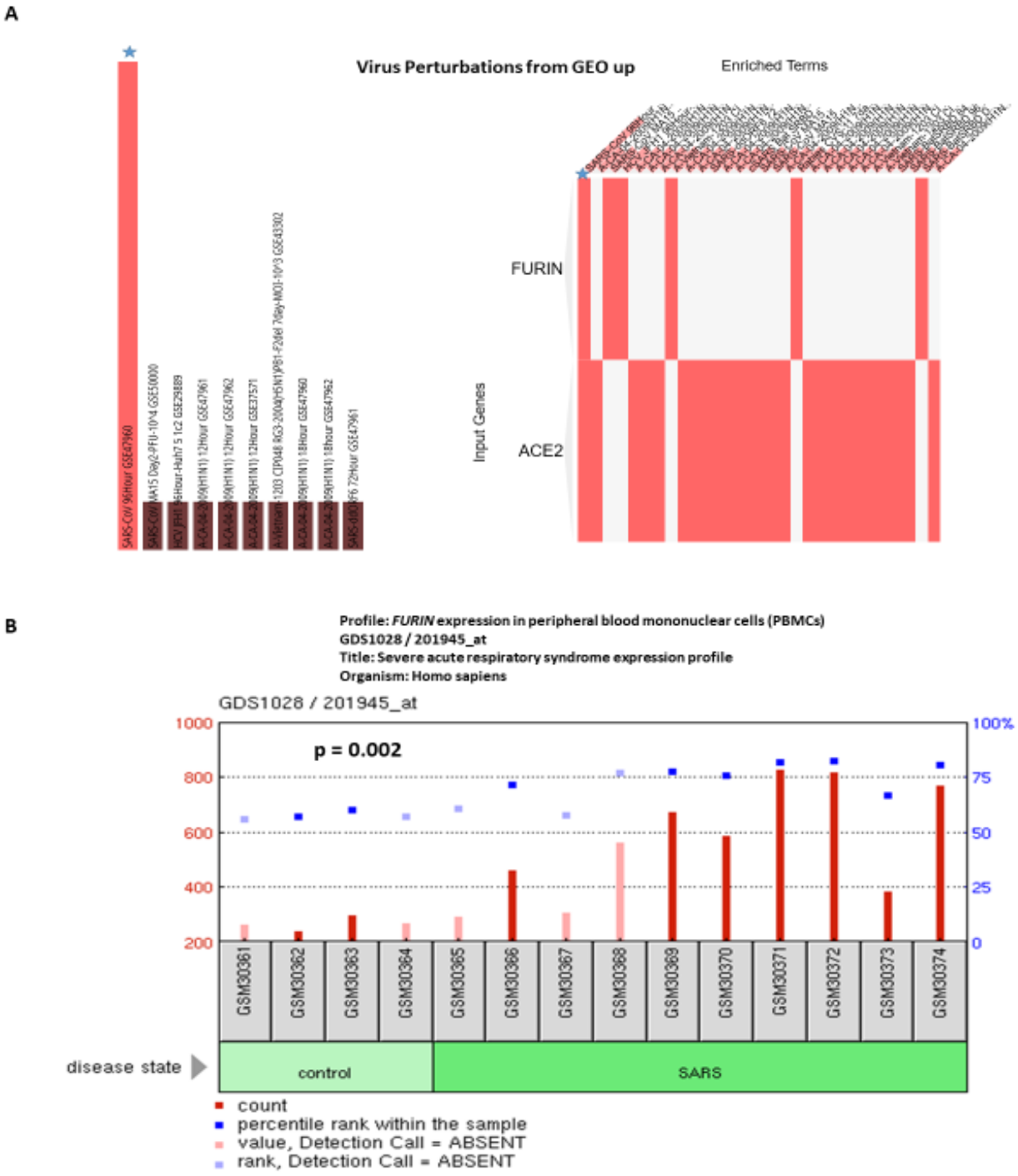


Figure 1. Effects of viral challenges on expression of the *ACE2* and *FURIN* genes.

a. Gene Set Enrichment analyses of the Virus Perturbations from GEO focused on up-regulated genes (Enrichr bioinformatics platform). SARS-CoV p value = 2.24E-04; q value = 0.072).

b. Increased *FURIN* expression in peripheral blood mononuclear cells (PBMC) of patients with Severe Acute Respiratory Syndrome (SARS). P value = 0.002; q value = 0.126.

It would be of interest to investigate whether this potentially infection-promoting effect on expression of the host genes in virus-targeted cells is mediated by the virus-induced release of the biologically-active molecules with the paracrine mode of actions such as interferons, interleukins and cytokines.

GSEA identified numerous significantly enriched records of common human disorders manifesting up-regulation of either *ACE* or *FURIN* genes (Supplemental Figure S3), which is consistent with the clinical observations that individuals with underlying health conditions are more likely to have clinically severe and lethal coronavirus infection.

Of note, both seasonal and pandemic H1N1 influenza virus infection significantly increases the ACE2 expression in human bronchial epithelial cells in vitro (Supplemental Figure S3). Exploration of the DisGeNET database of human disorders highlighted multiple disease states' records manifesting altered expression of either *ACE2* of *FURIN* genes (Supplemental Figure S3). Cigarette smoking appears to significantly increase the *ACE2* expression in human large airway epithelial cells (Supplemental Figure S3), indicating that cigarette smoking should be considered as a potential coronavirus infection-promoting agent.

Gene Ontology (GO) analyses revealed that *ACE2* and *FURIN* genes are associated with the largely non-overlapping records of GO Biological Processes, GO Molecular Functions, and GO Cellular Components (Supplemental Figure S4). The common significantly enriched records are Viral Life Cycle (GO Biological Process 2018); Peptidase activity (acting on L-amino acid peptides) and Endopeptidase activity (GO Molecular Function 2018); Membrane raft (GO Cellular Component 2018); Meprin A complex and Retrotrasposon nucleocapsid (Jensen Compartments).

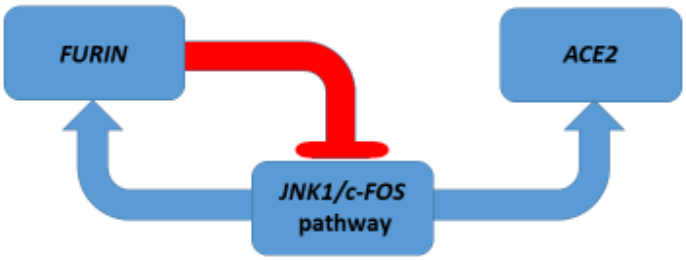
Identifications of the enriched records of transcription factor-binding sites affecting the *ACE2* and *FURIN* expression

GSEA of the enriched records of transcription factors' binding sites (TFBS) using ENCODE TF ChIP-seq 2015 and ChEA 2016 databases revealed predominantly distinct patterns of TFBS associated with the *ACE2* and *FURIN* genes (Supplemental Figure S5). Common TFBS shared by both *ACE2* and *FURIN* genes are *FOS, JUND, EP300* (ENCODE TF ChIP-seq 2015) and *GATA1, GATA2, RUNX1, FOXA1, HNF4A* (ChIP-seq 2015). Consistent with these findings, non-overlapping profiles of significantly enriched records associated with either *ACE2* or *FURIN* genes were observed of pathways (BioPlanet 2019 database), protein-protein interactions (PPI) hub proteins (PPI Hub Proteins database), and drugs affecting *ACE2* and *FURIN* expression (Drug Signatures Database, DSigDB), indicating that regulatory mechanisms governing the expression and activities of the *ACE2* and *FURIN* genes are predominantly discordant (Supplemental Figure S5).

Next, the Gene Expression Omnibus (GEO) database was interrogated to gauge the effects on *ACE2* and *FURIN* expression of transcription factors having TFBS associated with their promoters. There are multiple relevant GEO records reporting the activation effects of the *JNK1/c-FOS* pathway on *ACE2* and *FURIN* expression as well as the activation effects of *FURIN* depletion on expression of the *Fos, Jun, Jund,* and *Junb* genes (Supplemental Figure S5). Conversely, *c-Jun* inhibition (effect of the dominant negative *c-Jun*) nor *c-Jun* depletion (*c-Jun* knockout) has resulted in deceased expression of the *FURIN* gene (Supplemental Figure S5). The summary of these observations is reported in the Figure 2.

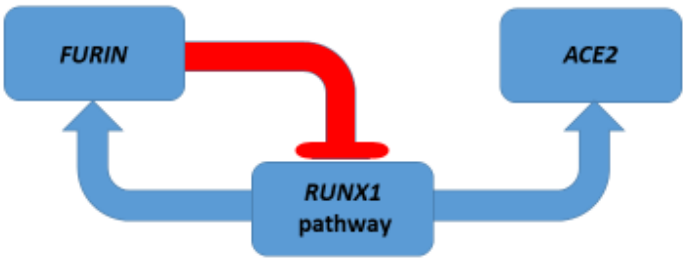
A

JNK1/c-FOS pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the FURIN-mediated repression of the expression of *JUN, JUNB, JUND,* and *c-FOS* genes



B

RUNX1 pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the *FURIN*-mediated repression of the *RUNX1* gene expression



c

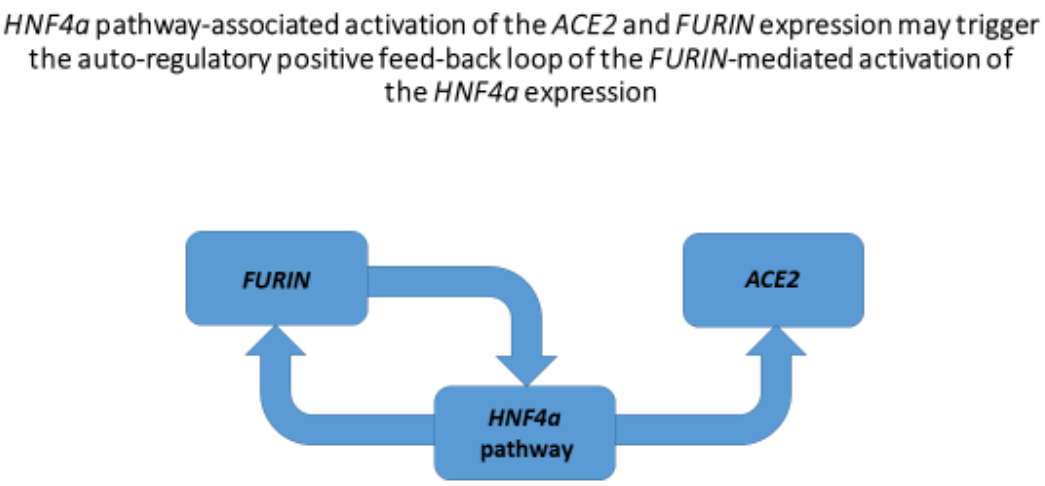


Figure 2. Pathways and genes affecting the newly emerged SARS-CoV-2 virus-related host targets.

a. *JNK1/c-FOS* pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the FURIN-mediated repression of the expression of *JUN, JUNB, JUND,* and *c-FOS* genes.

b. *RUNX1* pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the *FURIN*-mediated repression of the *RUNX1* gene expression

c. *HNF4a* pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory positive feed-back loop of the *FURIN*-mediated activation of the *HNF4a* expression.

Similarly, there are several reports indicating that depletion of either *Hnf4a* or *Runx1* in mouse cells and *RUNX1* in human cells decreases the *ACE2* and *FURIN* expression (Supplemental Figure S6). Conversely, *FURIN* depletion enhances expression of the *Runx1* and *Foxa1* genes in murine T cells (Supplemental Figure S6). In contrast, *FURIN* depletion decreases expression of the *Hnf4a* gene, while *Hnf4a* depletion decreases the *FURIN* gene expression (Supplemental Figure S6). The summary of these observations is reported in the Figure 2.

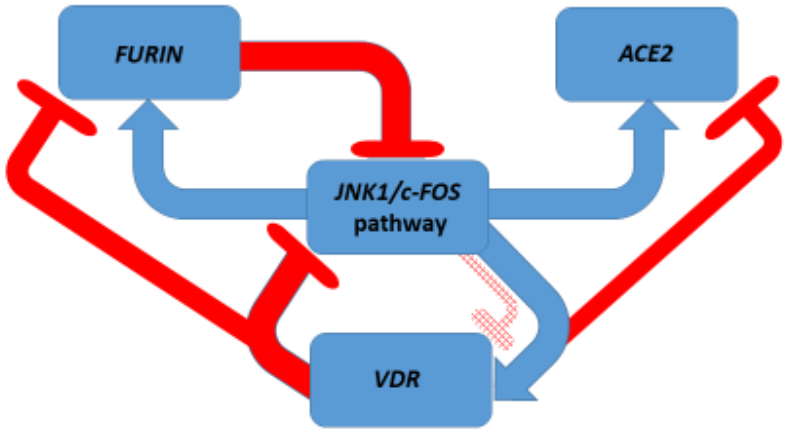
Identification of the *VDR* and *HIF1a* genes as putative repressors of the *ACE2* expression

Next GSEA of genomic databases were performed to identify the potential activators and repressors of the *ACE2* and *FURIN* genes. Analysis of the ARCHS4 transcription factors' co-expression database identified the *VDR* genes that co-expressed with both *ACE2* and *FURIN* genes in human tissues (Supplemental Figure S7). Other significantly enriched records manifest non-overlapping patterns of co-expression with either *ACE2* of *FURIN* genes. The GTEX expression profile of the *VDR* gene in human tissues revealed the ubiquitous pattern of expression and placed the *VDR* expression in human lungs in the top quartile (Supplemental Figure S7). Analysis of gene expression profiling experiments of wild type and vitamin D receptor (Vdr) knockout primary bone marrow-derived macrophages reported by Helming et al. (2005) demonstrate increased expression of the *ACE2* gene in the *VDR* knockout cells (Supplemental Figure S7) implicating the product of the *VDR* gene as the putative repressor of the *ACE2* expression. Consistent with this hypothesis, Vitamin D appears to inhibit the *ACE2* expression in human bronchial smooth muscle cells (Supplemental Figure S7).

Notably, examinations of direct and reciprocal effects of the *VDR* gene and Vitamin D administration on expression of the *JNK1/c-FOS* pathway genes revealed the expression profiles consistent with the potential therapeutic utility of the Vitamin D administration and activation of the *VDR* gene expression (Supplemental Figure S7). Analyses of direct and reciprocal effects of the *VDR* gene and Vitamin D administration on the *HNF4a* expression revealed that *HNF4a* depletion in human and murine cells inhibits the *VDR* gene expression, while the *Vdr* gene depletion increases the *Hnf4a* expression (Supplemental Figure S7). These result are consistent with the hypothesis stating that Vitamin D administration and activation of the *VDR* gene expression may have mitigating effects on the coronavirus infection. The summary of these findings is reported in the Figure 3.

A

JNK1/c-FOS pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the FURIN-mediated repression of the expression of *JUN, JUNB, JUND,* and *c-FOS* genes



B

HNF4a pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory positive feed-back loop of the *FURIN*-mediated activation of the *HNF4a* expression

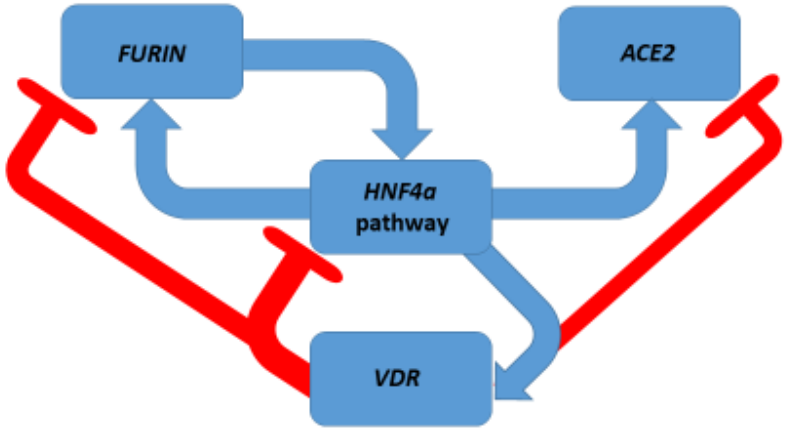


Figure 3. Effects of the *VDR* gene and Vitamin D on pathways and genes affecting the newly emerged SARS-CoV-2 virus-related host targets.

a. *JNK1/c-FOS* pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory negative feed-back loop of the FURIN-mediated repression of the expression of *JUN, JUNB, JUND,* and *c-FOS* genes.

b. *HNF4a* pathway-associated activation of the *ACE2* and *FURIN* expression may trigger the auto-regulatory positive feed-back loop of the *FURIN*-mediated activation of the *HNF4a* expression.

GSEA of the Transcription Factor's Perturbations Followed by Expression database and GEO Gene Perturbations database focused on up-regulated genes identified *HIF1a* and *POU5F1* gene products as putative repressors of the *ACE2* and *FURIN* expression (Supplemental Figure S8). These findings were corroborated by observations that *HIF1a* overexpression in human embryonic kidney cells significantly inhibits the *ACE2* expression (Supplemental Figure S8). Notably, Vitamin D significantly increases expression of the *HIF1a* gene in human bronchial smooth muscle cells (Supplemental Figure S8), suggesting that *VDR* and *HIF1A* genes may cooperate as repressors of the *ACE2* expression.

GSEA identify Estradiol and Quercetin as putative candidate coronavirus infection mitigation agents.

GSEA of the Drug Perturbations from GEO database focused on down-regulated genes identified Estradiol and Quercetin among the top significantly enriched records (Supplemental Figure S9). Estradiol appears to affect both *FURIN* and *ACE2* expression, while Quercetin seems to target the *ACE2* expression. Consistently, GSEA of the Ligand Perturbations from GEO focused on down-regulated genes identified five of Estradiol administration records (50%) among top ten significantly enriched ligand perturbations records (Supplemental Figure S9). GSEA of the Drug Perturbations from GEO database focused on up-regulated genes indicated that doxorubicin, imatinib, and bleomycin may act as potential coronavirus infection-promoting agents (data not shown). Collectively, these observations provide the initial evidence supporting the hypothesis that both Estradiol and Quercetin may function as potential candidate coronavirus infection mitigation agents.

Consistent with this hypothesis, interrogation of the GEO records revealed that Quercetin appears to inhibit expression of several potential coronavirus infection-promoting genes: *c-FOS* expression in human and rat cells (Supplemental Figure S9); *Runx1* expression in rat cells (Supplemental Figure S9); *HNF4a* expression in human cells (Supplemental Figure S9).

However, Quercetin administration appears to increase *c-Fos* expression in cultured rat cardiomyocytes (Supplementary Figure S9).

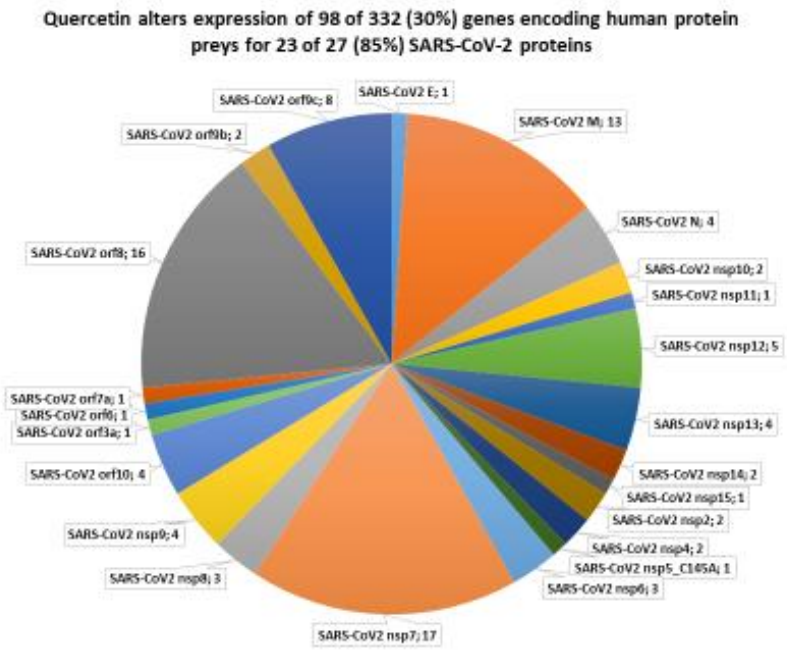
Confirmation of the Estradiol and Quercetin activities as potential candidate coronavirus infection mitigation agents.

Results of GSEA suggest that both Estradiol and Quercetin appear to exhibit biological activities consistent with the activity of medicinal compounds expected to mitigate the coronavirus infection. Next, manual curation of the GEO data sets has been carried out to identify further experimental evidence supporting this hypothesis. Administration of Estradiol appears to inhibit *ACE2* and/or *FURIN* expression in rat, mouse, and human cells (Supplemental Figure S10) and the effects of Estradiol seem to be mediated by the estrogen receptor beta. In agreement with the hypothesis on potential therapeutic utility of the Quercetin, administration of Quercetin has resulted in significantly decreased expression of the *ACE2* gene during differentiation of human intestinal cells (Supplemental Figure S10).

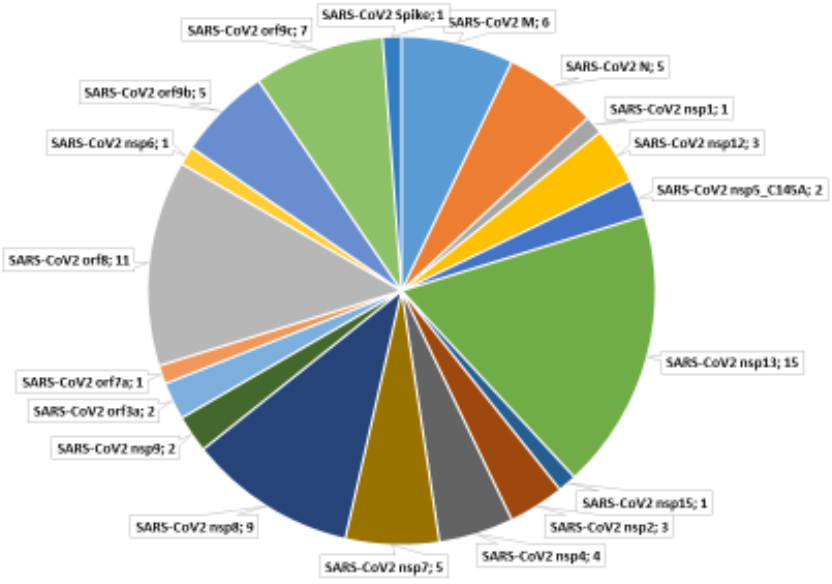
Inference of potential interference of the Quercetin and Vitamin D with functions of SARS-CoV-2 proteins in human cells.

An excellent recent proteomics study of the SARS-CoV-2 interactome in human cells identified 332 high-confidence protein targets of the 27 SARS-CoV-2 viral proteins (Gordon et al., 2020). GSEA of the 332 human genes encoding human prey proteins of the coronavirus SARS-CoV-2 revealed nuclear envelope disassembly, proteins' targeting to mitochondrion, and tRNA and protein transports as major biological processes targeted by the SARS-CoV-2 (Supplemental Figure S14). RNA and GDP binding functions were identified as main biological functions while mitochondrion was identified as one of top-scoring cellular components targeted by the SARS-CoV-2 infection (Supplemental Figure S14).

A


Quercetin alters expression of 98 of 332 (30%) genes encoding human protein preys for 23 of 27 (85%) SARS-CoV-2 proteins
SARS-CoV2 E; 1
SARS-CoV2 M; 13
SARS-CoV2 N; 4
SARS-CoV2 nsp10; 2
SARS-CoV2 nsp11; 1
SARS-CoV2 nsp12; 5
SARS-CoV2 nsp13; 4
SARS-CoV2 nsp14; 2
SARS-CoV2 nsp15; 1
SARS-CoV2 nsp2; 2
SARS-CoV2 nsp4; 2
SARS-CoV2 nsp5_C145A; 1
SARS-CoV2 nsp6; 3
SARS-CoV2 nsp7; 17
SARS-CoV2 nsp8; 3
SARS-CoV2 nsp9; 4
SARS-CoV2 orf10; 4
SARS-CoV2 orf3a; 1
SARS-CoV2 orf6; 1
SARS-CoV2 orf7a; 1
SARS-CoV2 orf8; 16
SARS-CoV2 orf9b; 2
SARS-CoV2 orf9c; 8

B


Vitamin D alters expression of 84 of 332 (25%) genes encoding human prey proteins for 19 of 27 (70%) of SARS-CoV-2 proteins
SARS-CoV2 Spike; 1
SARS-CoV2 M; 6
SARS-CoV2 N; 5
SARS-CoV2 nsp1; 1
SARS-CoV2 nsp12; 3
SARS-CoV2 nsp5_C145A; 2
SARS-CoV2 nsp13; 15
SARS-CoV2 nsp15; 1
SARS-CoV2 nsp2; 3
SARS-CoV2 nsp4; 4
SARS-CoV2 nsp7; 5
SARS-CoV2 nsp8; 9
SARS-CoV2 nsp9; 2
SARS-CoV2 orf3a; 2
SARS-CoV2 orf7a; 1
SARS-CoV2 orf8; 11
SARS-CoV2 nsp6; 1
SARS-CoV2 orf9b; 5
SARS-CoV2 orf9c; 7

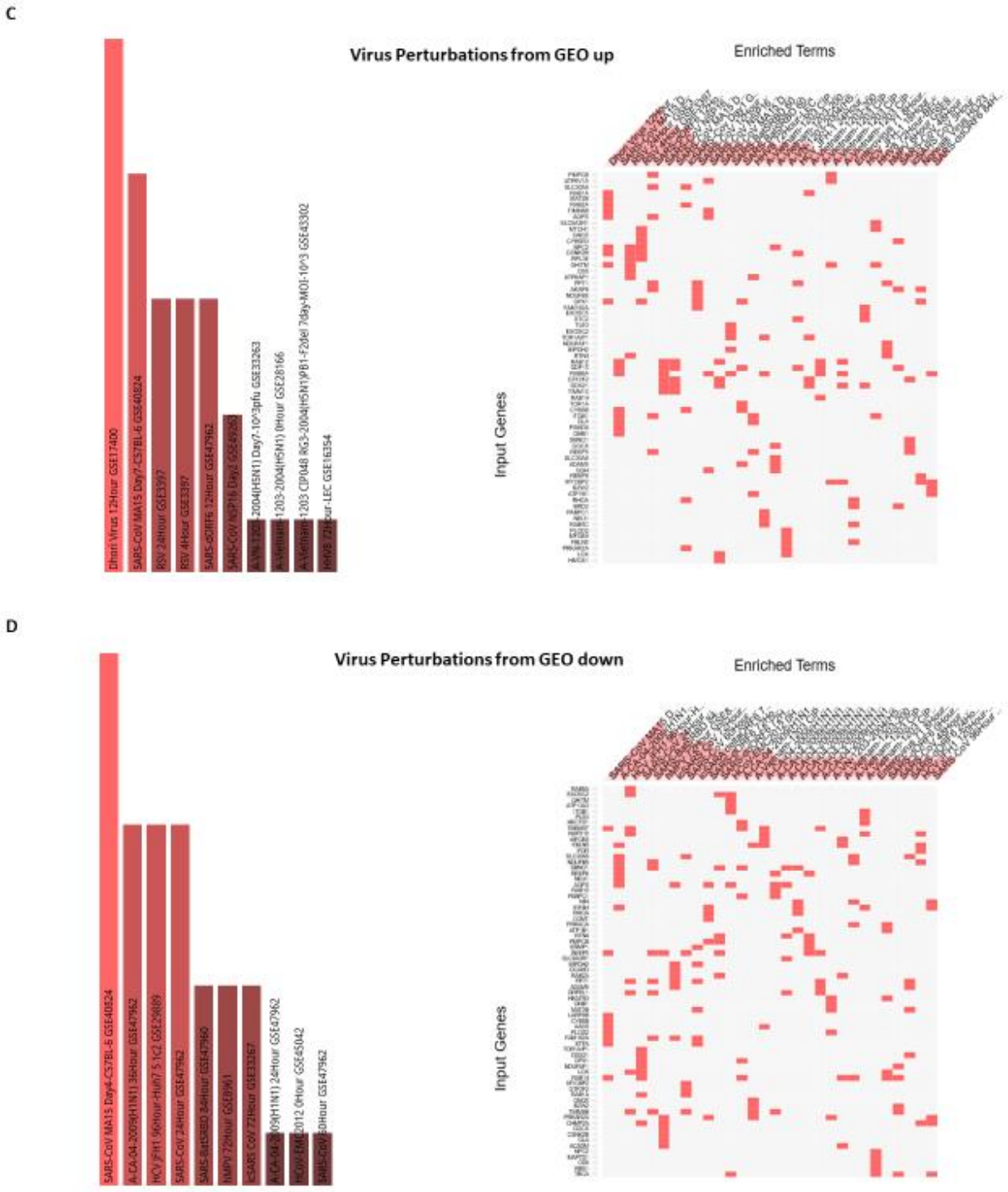


Figure 4. Effects of Quercetin and Vitamin D on expression of genes encoding human prey proteins of the SARS-CoV-2 coronavirus.

(A) Quercetin alters expression of 98 of 332 (30%) of genes encoding human prey proteins for 23 of 27 (85%) of SARS-CoV-2 proteins.
(B) Vitamin D alters expression of 84 of 332 (25%) of genes encoding human prey proteins for 19 of 27 (70%) of SARS-CoV-2 proteins.

Panels C and D report the results of GSEA of the 332 genes in the Virus Perturbations from GEO database for up-regulated (C) and down-regulated (D) genes. See text and Supplemental Figure S14 for details.

Interestingly, Quercetin alters expression of 98 of 332 (30%) of genes encoding protein targets of the SARS-CoV-2 in human cells, thus potentially interfering with activities of the 23 of 27 (85%) of SARS-CoV-2 proteins (Figure 4; Supplemental Figure S14). Similarly, analyses of Vitamin D-regulated genes identified in undifferentiated human THP-1 cells (Seuter et al., 2016; Neme et al., 2017) revealed that Vitamin D alters the expression of 84 of 332 (25%) human genes encoding prey proteins for SARS-CoV-2 coronavirus (Figure 4; Supplemental Figure S14). These observations suggest that Vitamin D may interfere with functions of 19 of 27 (70%) of SARS-CoV-2 proteins (Figure 4; Supplemental Figure S14).

Gene expression-guided inference of the potential effects of a combination of both Quercetin and Vitamin D on functions of SARS-CoV-2 proteins suggests that it may affect the activities of nearly all (25 of 27; 93%) SARS-CoV-2 viral proteins in human cells. It would be of interest and essential to test the validity of these predictions experimentally before the definitive conclusions regarding the potential utility of a combination of Quercetin and Vitamin D as putative COVID-19 mitigation agents can be made.

Notably, Estradiol manifests significant patterns of interference with the expression of genes encoding 332 prey proteins of the SARS-CoV-2 in human cells (Supplemental Figure S14), while Testosterone did not manifest significant associations (data not shown). However, Estradiol administration appears to manifest the cell type-specific effects on c-*FOS* expression (Supplemental Figure S10). For example, it decreases the c-*FOS* expression in endometrium of Macaca mulatta while it increases c-*FOS* expression in the mouse uterus (Supplemental Figure S10).

Collectively, these observations indicate that any definitive conclusions regarding the potential clinical utility of identified herein potential coronavirus infection mitigating agents should be made only after appropriately designed and carefully executed preclinical studies and randomized clinical trials. In contrast to the Estradiol, which exhibit evidence of both putative coronavirus infection-mitigating actions and coronavirus infection-promoting activities, administration of Testosterone appears to manifest more clearly-defined patterns of altered gene expression consistent with Testosterone being identified as the potential coronavirus infection-promoting agent (Supplemental Figure S11).

Potential mechanisms affecting gene expression inferred from transgenic mouse models and observed in pathophysiologically and therapeutically relevant mouse and human cells.

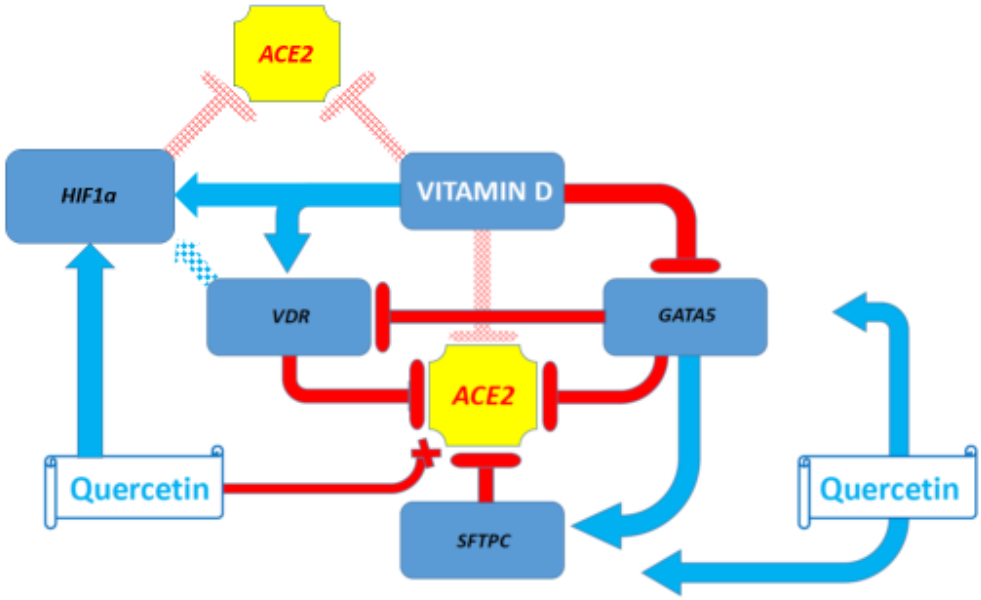
Taking into considerations that the effects of potential coronavirus infection mitigation agents often manifest cell type-specific patterns of gene expression changes, next the manual curation of the GEO gene expression profiles were carried out to identify the relevant host genetic targets and putative mitigation agents. These analyses identified several candidate repressors (*VDR; GATA5; SFTC; HIF1a*) and activators (*INSIG1; HMGA2*) of the *ACE2* and *FURIN* expression (Supplemental Figure S12). Notably, the effects on gene expression of the administration of either Vitamin D or Quercetin appear consistent with their definition as putative coronavirus infection mitigation agents (Supplemental Figure S12). The summary of these observations is presented in the Figure 5. The conclusion regarding the findings of cell type-specific effects on gene expression of putative coronavirus infection mitigating agents remains valid and examples of the potential negative effects of drugs on the ACE2 expression are reported in the Supplemental Figure S12). For example, the *HIF1a* expression is significantly increased in murine alveolar type I cells deficient in sterol-response element-binding proteins inhibitor Insig1 (Supplemental Figure S12). These data indicate that the *INSIG1* gene product,

which appears to function as activator of the *ACE2* expression, may function as the inhibitor of the *HIF1a* expression, thus interfering with the *HIF1a*-mediated *ACE2* repression in specific cell types. Additional examples of the potential positive and negative effects on gene expression inferred from transgenic mouse models are reported in the Supplemental Figure S13.

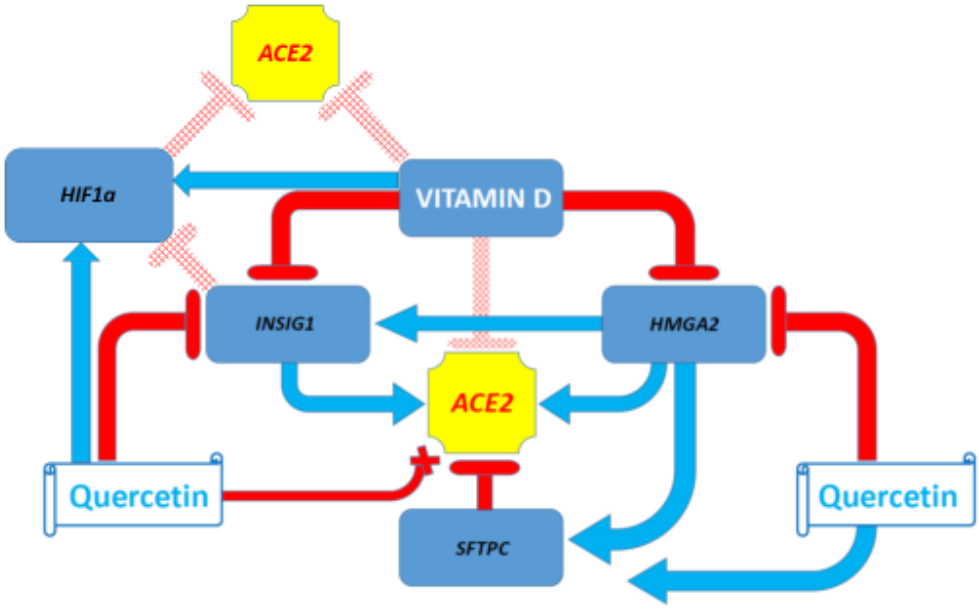
Is Vitamin D deficiency a potential risk factor for increased disease severity in older adults and elderly individuals?

Present analyses suggest that Vitamin D and vitamin D receptor (*VDR*) are putative mitigation factors of the coronavirus infection. Conversely, the Vitamin D deficiency could be a potential aggravating factor for the clinical course of pandemic. Multiple lines of evidence suggest that Vitamin D deficiency, particularly in elderly, might be a negative factor affecting the clinical course of the pandemic. In the United States, approximately 30% of whites and 5% of African Americans have sufficient Vitamin D level (Kennel et al., 2010) and the significant increase of the prevalence of individuals with severe Vitamin D deficiency has been reported (Ginde et al., 2009a). Age-associated decline of the human skin function to produce the Vitamin D in response to the sunlight exposure is likely a contributing factor to the Vitamin D deficiency in older individuals, since it has been reported that elderly people produce 75% less of cutaneous Vitamin D3 than young individuals (Lips, 2001). A meta-analysis of randomized controlled clinical trials indicated that intake of ordinary doses of Vitamin D was associated with significant decrease in total mortality rates (Autier and Gandini, 2007). A prospective study of the 3,408 older adults in the United States demonstrated that a group at high risk of all-cause mortality could be defined by the serum 25-hydroxyvitamin D [25(OH)D] level (Ginde et al., 2009b). A significant, independent, inverse association was observed between the serum 25(OH)D level and all-cause and cardiovascular diseases (CVD) mortality (Ginde et al., 2009b).

A


ACE2
HIF1α
VITAMIN D
VDR
GATA5
ACE2
Quercetin
SFTPC
Quercetin

B


ACE2
HIF1α
VITAMIN D
INSIG1
HMGA2
ACE2
Quercetin
SFTPC
Quercetin

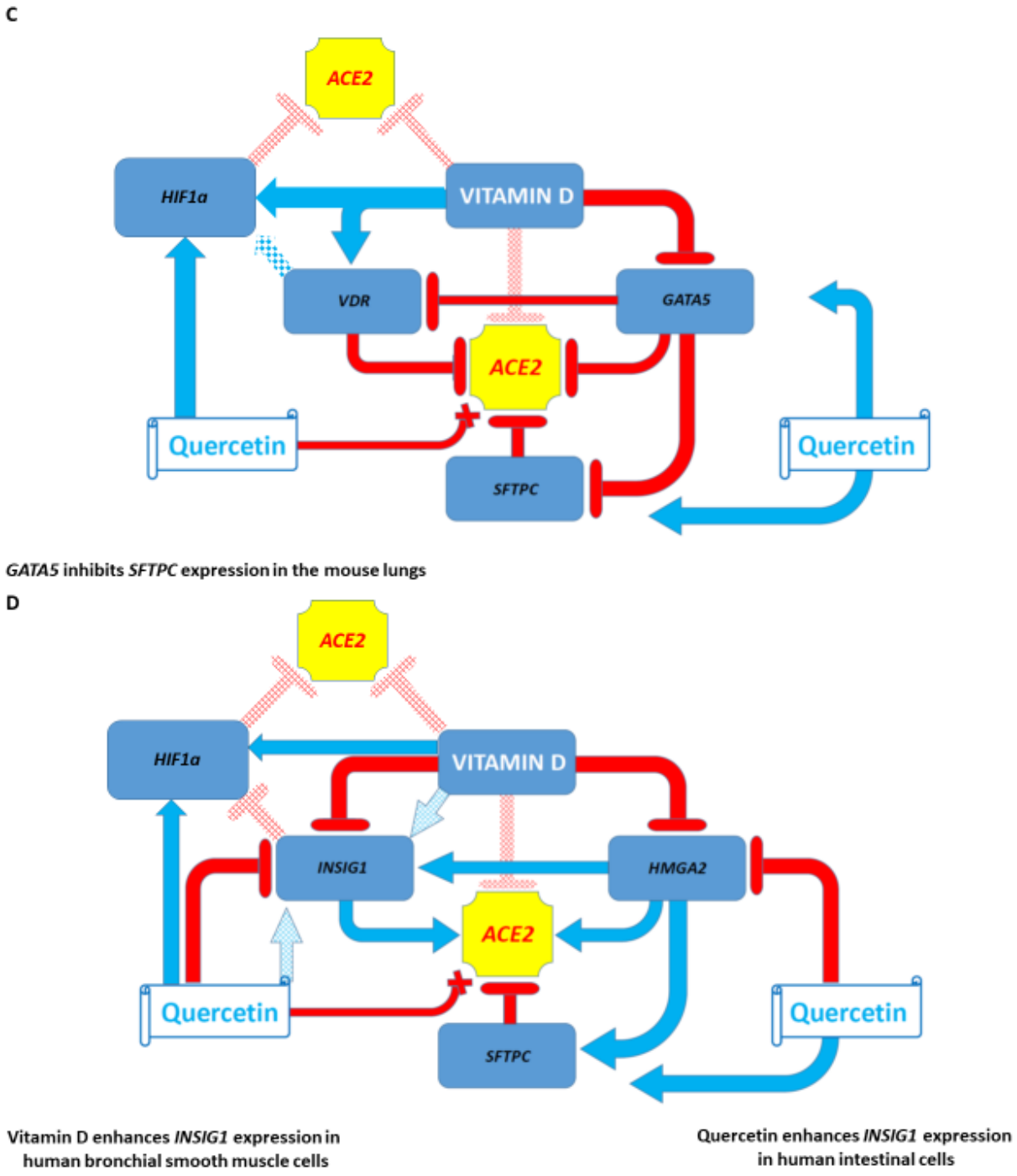


Figure 5. Effects of the *VDR* gene, Vitamin D, and Quercetin on pathways and genes affecting the newly emerged SARS-CoV-2 virus-related host targets.

a. Effects of the *VDR* gene, Vitamin D, and Quercetin on repressors of the *ACE* expression.

b. Effects of the *VDR* gene, Vitamin D, and Quercetin on activators of the *ACE* expression.

c. Effects of the *VDR* gene, Vitamin D, and Quercetin on repressors of the *ACE* expression reflecting *GATA5* inhibitory effects on *SFTPC* expression in the mouse lungs.
d. Effects of the *VDR* gene, Vitamin D, and Quercetin on activators of the ACE expression reflecting the cell type-specific effects of Vitamin D and Quercetin: Vitamin D-induced activation of the *INSIG1* expression in human bronchial smooth muscle cells and Quercetin-induced activation of the *INSIG1* expression in human intestinal cells.

To date, the vast majority of observational studies reported inverse associations between the circulating 25(OH)D concentration and all-cause mortality in generally healthy populations (Heath et al., 2019). In generally healthy adults over 50 years old, significant inverse associations were found between low 25(OH)D levels and all-cause mortality, respiratory and cardiovascular events, as well as markers relating to hip and non-vertebral fractures (Caristia et al., 2019). Therefore, it would be important to determine whether Vitamin D deficiency may be one of the risk factors contributing to the increased disease severity in older adults and elderly individuals during coronavirus pandemic. Significantly, recent studies highlight numerous beneficial clinical effects of the Vitamin D supplementations (Grant et al., 2020; Fabbri et al., 2020) and underscore the significant COVID-19 mitigation potential of the Vitamin D.

Limitations

Several shortcomings of present analyses must be considered as limitations of this study. Identified and reported in this study features, records, and traits are top-scoring in the relative ranks of analyzed hundreds or thousands features, records, and traits included in specific databases. However, in several instances when the top-scoring candidate genomic features and/or traits were selected and reported, the statistical significance was achieved only based on the nominal p values. This was primarily due to the relatively small sample size available for the interrogation, which highlights another common limitations of the available records. Where possible, the attempts to address these shortcomings were made by analyzing multiple independent data sets to arrive at the consensus. One of the important limitations is the small number of studies conducted on human cells, in particular, human cells and tissues

pathophisiologically-relevant to the COVID-19 pandemic. Collectively, these considerations underscore the critical need for follow-up targeted experimental studies and, if warranted, randomized clinical trials to identify and validate the utility of the Quercetin and Vitamin D as therapeutically-viable interventions to combat the COVID-19 pandemic.

Conclusion

The main motivation of this work was to identify human genes implicated in regulatory cross-talks affecting expression and functions of the *ACE2* and *FURIN* genes to build a model of genomic regulatory interactions potentially affecting the SCARS-CoV-2 coronavirus infection. A panel of genes acting as activators and/or repressor of the *ACE2* and/or *FURIN* expression then could be employed to search for existing drugs and medicinal substances that, based on their mechanisms of activities, could be defined as the candidate coronavirus infection mitigation agents. After experimental and clinical validation, these existing drugs could be utilized to ameliorate the clinical severity of the pandemic. This knowledge could also be exploited in an ongoing effort to discover novel targeted therapeutics tailored to prevent the SCARS-CoV-2 infection and block the virus entry into human cells. Observations reported in this contribution are in agreement with recent studies describing numerous beneficial clinical effects of the Vitamin D supplementations, emphasizing many detrimental effects of the Vitamin D insufficiency and deficiency, and underscoring the significant COVID-19 mitigation potential of the Vitamin D (Grant et al., 2020; Fabbri et al., 2020).

One of the important findings documented herein is that identified medicinal compounds with potential coronavirus infection-mitigating effects also appear to induce cell type-specific patterns of gene expression alterations. Therefore, based on all observations reported in this contribution, it has been concluded that any definitive recommendations regarding the potential clinical utility of identified herein putative coronavirus infection mitigating agents, namely Vitamin D and Quercetin, should be made only after preclinical studies and randomized clinical trials

have been appropriately designed, carefully executed, and the desired outcomes have been reached.

A super-computer modeling study using the world's most powerful supercomputer, SUMMIT, identified several candidate small molecule drugs which bind to either the isolated SCARS-CoV-2 Viral S-protein at its host receptor region or to the S protein-human ACE2 interface (Smith and Smith, 2020). Interestingly, in this study Quercetin was identified among top 5 scoring ligands for viral S-protein-human ACE2 receptor interface. Thus, Quercetin appears also a potentially promising therapeutic molecule that may directly interfere with the binding of the SCARS-CoV-2 virus to human cells. Previously reported experiments demonstrated that Quercetin appears to inhibit the SARS-CoV virus entry into host cells (Yi et al., 2004). Since the SCARS-CoV-2 virus utilizes for the entry in human cells the same receptor (ACE2) and the accessory protease FURIN as the SARS-CoV coronavirus (Wells et al., 2020), these observations suggest that Quercetin may, indeed, possess antiviral activity against the SARS-CoV-2 as well. Significantly, both Quercetin and Luteolin have been identified among top 5 ligands for the viral S-protein-human ACE2 receptor interface-ligand binding complex (Smith and Smith, 2020), suggesting that these structurally highly similar compounds (Figure 6) could serve as efficient inhibitors of the SARS-CoV-2 infection. Consistent with this hypothesis, it has been reported that both Quercetin and Luteolin significantly inhibit the SARS-CoV virus infection (Yi et al., 2004).

It has been observed that administration of Testosterone appears to manifest clearly-defined patterns of altered gene expression consistent with Testosterone being identified as the potential coronavirus infection-promoting agent, particularly in some cell types that may play a role in the virus entry into human body and the respiratory system (Supplemental Figure S11). This is in contrast to Estradiol, which seems to manifest cell type-specific effects on gene expression consistent with either infection-inhibiting or infection-promoting patterns of gene

expression changes. It would be of interest to determine whether this discordant effects may contribute to the apparently higher mortality among men with coronavirus infection.

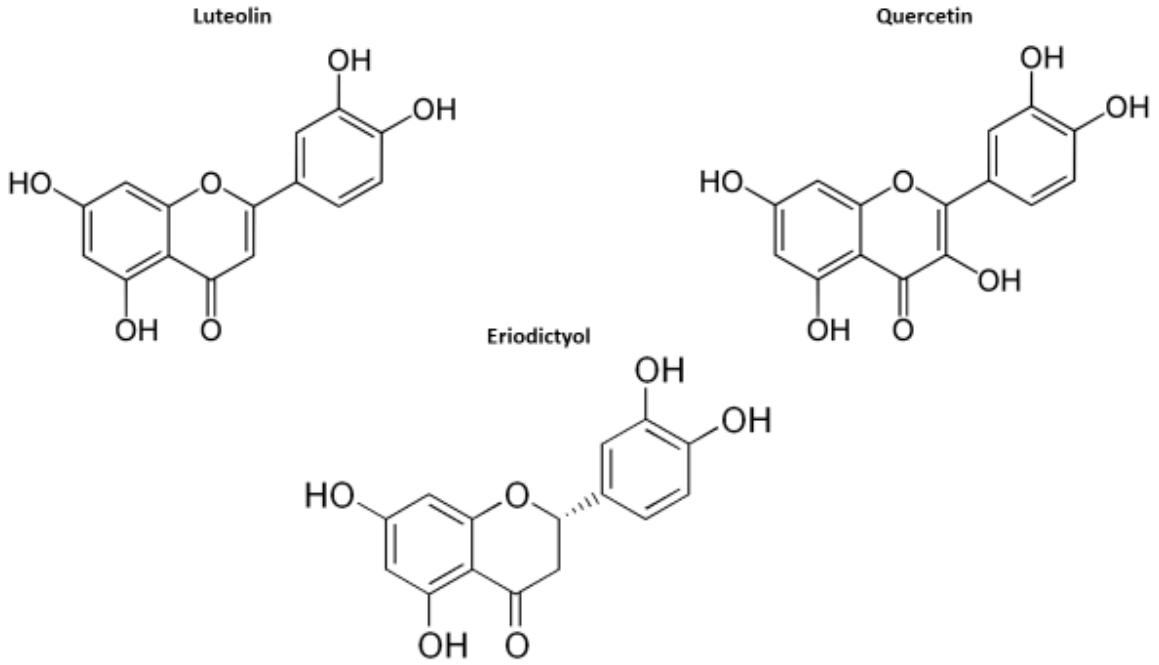


Figure 6. Similarity of chemical structures of the Luteolin, Quercetin, and Eriodictyol. Luteolin, Quercetin, and Eriodictyol have been identified in the recent supercomputer SUMMIT drug-docking screen as top candidate inhibitors of the SARS-CoV-2 spike-protein-human ACE2 receptor interface-ligand binding complex (Smith and Smith, 2020), while Luteolin and Quercetin have been identified as potent inhibitors of the SARS-CoV infection (Yi et al., 2004).

Present analyses highlighted the major uncertainty regarding the outcomes of the current pandemic associated with the potential of the SCARS-CoV-2 virus for the expansion of the cellular tropism (Walls et al., 2020) based on access to genetically vulnerable host cells due to nearly ubiquitous expression of the *ACE2* and *FURIN* genes in the human body. Particularly dangerous seems noted in this contribution the potential ability of the SCARS-CoV-2 virus to infect the immune cells because the infection of immune cells may cause immunosuppression, long-term persistence of the virus in the body, and spread of the virus to secondary targets. Taken together with predominantly cell type-specific patterns of expression of genetic

repressors and activators of the *ACE2* and *FURIN* expression it may complicate the development of universally effective therapeutics. The availability of many genetically-relevant transgenic mouse models, in particular, the *Furin* null mice, should be regarded as a considerable advantage for preclinical development of drug candidates tailored to target the coronavirus infection. Specifically, the potential therapeutic utility of the highly selective ($K_{i,}$ 600 pm) intrinsically-specific FURIN inhibitor (a1-antitrypsin Portland (a1-PDX); Jean et al., 1998) should be tested in the immediate future.

Methods

Data source and analytical protocols

All data analyzed in this study were obtained from the publicly available sources. Gene set enrichment analyses (GSEA) were carried-out using the Enrichr bioinformatics platform, which enables the interrogation of nearly 200,000 gene sets from more than 100 gene set libraries. The Enrichr API (January 2020 through March 2020 releases) (Chen et al., 2013; Kuleshov et al., 2016) was used to test genes linked to the ACE2 and FURIN genes (or other genes of interest) for significant enrichment in numerous functional categories. In all tables and plots (unless stated otherwise), in addition to the nominal p values and adjusted p values, the "combined score" calculated by Enrichr is reported, which is a product of the significance estimate and the magnitude of enrichment (combined score c = log(p) * z, where p is the Fisher's exact test p-value and z is the z-score deviation from the expected rank). Validation of the GSEA findings were carried-out employing the computational retrievals and manual curations of the gene expression profiles of the Gene Expression Omnibus (GEO) database.

Statistical Analyses of the Publicly Available Datasets

All statistical analyses of the publicly available genomic datasets, including error rate estimates, background and technical noise measurements and filtering, feature peak calling, feature selection, assignments of genomic coordinates to the corresponding builds of the reference

human genome, and data visualization, were performed exactly as reported in the original publications (Glinsky, 2015-2020; Glinsky and Barakat, 2019; Glinsky et al., 2019; Guffanti et al., 2018) and associated references linked to the corresponding data visualization tracks (http://genome.ucsc.edu/). Any modifications or new elements of statistical analyses are described in the corresponding sections of the Results. Statistical significance of the Pearson correlation coefficients was determined using GraphPad Prism version 6.00 software. Both nominal and Bonferroni adjusted p values were estimated. The statistical significance between the mean values was estimated using the Student T-test. The significance of the differences in the numbers of events between the groups was calculated using two-sided Fisher's exact and Chi-square test, and the significance of the overlap between the events was determined using the hypergeometric distribution test (Tavazoie et al., 1999).

Supplemental Information

Supplemental information includes Supplemental Figures S1-S14. Supplemental information is available as separate files online.

Author Contributions

This is a single author contribution. All elements of this work, including the conception of ideas, formulation, and development of concepts, execution of experiments, analysis of data, and writing of the paper, were performed by the author.

Acknowledgements

This work was made possible by the open public access policies of major grant funding agencies and international genomic databases and the willingness of many investigators worldwide to share their primary research data. This work was supported in part by the OncoScar, Inc.

References

Autier P, Gandini S. Vitamin D supplementation and total mortality: a meta-analysis of randomized controlled trials. Arch Intern Med. 2007; 167: 1730-1737.

Caristia S, Filigheddu N, Barone-Adesi F, Sarro A, Testa T, Magnani C, Aimaretti G, Faggiano F, Marzullo P. Vitamin D as a biomarker of ill health among the over-50s: A systematic review of cohort studies. Nutrients. 2019; 11. pii: E2384. doi: 10.3390/nu11102384.

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14, 128. doi: 10.1186/1471-2105-14-128.

Fabbri A, et al. Editorial – Vitamin D status: a key modulator of innate immunity and natural defense from acute viral respiratory infections. European Review for Medical and Pharmacological Sciences. 2020. 24: 4048-4052.

Ginde AA, Liu MC, Camargo CA Jr. Demographic differences and trends of vitamin D insufficiency in the US population, 1988-2004. Arch Intern Med. 2009a; 169: 626-632.

Ginde AA, Scragg R, Schwartz RS, Camargo CA Jr. Prospective study of serum 25-hydroxyvitamin d level, cardiovascular disease mortality, and all-cause mortality in older U.S. Adults. J Am Geriatr Soc. 2009b; 57: 1595-1603.

Glinsky GV. 2015. Transposable elements and DNA methylation create in embryonic stem cells human-specific regulatory sequences associated with distal enhancers and non-coding RNAs. Genome Biol Evol 7: 1432-1454.

Glinsky GV. 2016. Mechanistically distinct pathways of divergent regulatory DNA creation contribute to evolution of human-specific genomic regulatory networks driving phenotypic divergence of Homo sapiens. Genome Biol Evol 8:2774-88.

Glinsky GV. 2016. Activation of endogenous human stem cell-associated retroviruses (SCARs) and therapy-resistant phenotypes of malignant tumors. Cancer Lett 376:347-359.

Glinsky GV. 2016. Single cell genomics reveals activation signatures of endogenous SCAR's networks in aneuploid human embryos and clinically intractable malignant tumors. Cancer Lett 381:176-93.

Glinsky GV. 2017. Human-specific features of pluripotency regulatory networks link NANOG with fetal and adult brain development. BioRxiv. https://www.biorxiv.org/content/early/2017/06/19/022913; doi: https://doi.org/10.1101/022913.

Glinsky GV. 2018. Contribution of transposable elements and distal enhancers to evolution of human-specific features of interphase chromatin architecture in embryonic stem cells. Chromosome Res. 2018. 26: 61-84.

Glinsky G, Durruthy-Durruthy J, Wossidlo M, Grow EJ, Weirather JL, Au KF, Wysocka J, Sebastiano V. 2018. Single cell expression analysis of primate-specific retroviruses-derived HPAT lincRNAs in viable human blastocysts identifies embryonic cells co-expressing genetic markers of multiple lineages. Heliyon 4: e00667. doi: 10.1016/j.heliyon.2018.e00667. eCollection 2018 Jun. PMID: 30003161.

Glinsky GV, Barakat TS. 2019. The evolution of Great Apes has shaped the functional enhancers' landscape in human embryonic stem cells. 37:101456. PMID: 31100635. DOI: 10.1016/j.scr.2019.101456

Glinsky GV. 2020. A catalogue of 59,732 human-specific regulatory sequences reveals unique to human regulatory patterns associated with virus-interacting proteins, pluripotency and brain development. DNA and Cell Biology, 2020: 39: 126-143. doi: 10.1089/dna.2019.4988.

Gordon DE, et al. A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug Repurposing. 2020. doi: https://doi.org/10.1101/2020.03.22.002386

Grant WR, Lahore H, McDonnell SL, et al. Evidence that Vitamin D Supplementation Could Reduce Risk of Influenza and COVID-19 Infections and Deaths. Nutrients, 2020: 12: 988; doi:10.3390/nu12040988

Guffanti G, Bartlett A, Klengel T, Klengel C, Hunter R, Glinsky G, Macciardi F. 2018. Novel bioinformatics approach identifies transcriptional profiles of lineage-specific transposable elements at distinct loci in the human dorsolateral prefrontal cortex. Mol Biol Evol. 35: 2435-2453. PMID: 30053206. PMCID: PMC6188555. DOI: 10.1093/molbev/msy143.

Heath, AK, Kim IY, Hodge AM, English DR, and Muller, DC. Vitamin D status and mortality: A systematic review of observational studies. Int. J. Environ. Res. Public Health 2019, 16, 383; doi:10.3390/ijerph16030383.

Helming L, Böse J, Ehrchen J, Schiebe S, Frahm T, Geffers R, Probst-Kepper M, Balling R, Lengeling A. 1alpha,25-Dihydroxyvitamin D3 is a potent suppressor of interferon gamma-mediated macrophage activation. Blood. 2005; 106: 4351-8.

Kennel KA, Drake MT, Hurley DL. Vitamin D deficiency in adults: when to test and how to treat. Mayo Clin Proc. 2010; 85: 752-7.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research. 2016; gkw377.

Lips P. Vitamin D deficiency and secondary hyperparathyroidism in the elderly: consequences for bone loss and fractures and therapeutic implications. Endocr Rev. 2001; 22: 477-501.

Neme, A., Seuter, S., and Carlberg, C. (2017). Selective regulation of biological processes by vitamin D based on the spatio-temporal cistrome of its receptor. Biochim. Biophys. Acta 1860, 952–961. doi: 10.1016/j.bbagrm.2017.07.002

Reghunathan R, Jayapal M, Hsu LY, Chng HH, Tai D, Leung BP, Melendez AJ. Expression profile of immune response genes in patients with Severe Acute Respiratory Syndrome. BMC Immunol. 2005; 6:2.

Seuter, S., Neme, A., and Carlberg, C. (2016). Epigenome-wide effects of vitamin D and their impact on the transcriptome of human monocytes involve CTCF. Nucleic Acids Res. 44, 4090–4104. doi: 10.1093/nar/gkv1519

Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F. Structural basis of receptor recognition by SARS-CoV-2. Nature. 2020. doi: 10.1038/s41586-020-2179-y. Online ahead of print. PMID: 32225175

Smith M, Smith JC. Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface. ChemRxiv. 2020. Preprint. https://doi.org/10.26434/chemrxiv.11871402.v3

Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., and Church, GM. Systematic determination of genetic network architecture. Nat Genet 1999; 22, 281–285.

Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020. pii: S0092-8674(20)30262-2. doi: 10.1016/j.cell.2020.02.058. [Epub ahead of print]

Yi L, Li Z, Yuan K, Qu X, Chen J, Wang G, Zhang H, Luo H, Zhu L, Jiang P, Chen L, Shen Y, Luo M, Zuo G, Hu J, Duan D, Nie Y, Shi X, Wang W, Han Y, Li T, Liu Y, Ding M, Deng H, Xu X. Small molecules blocking the entry of severe acute respiratory syndrome coronavirus into host cells. J Virol. 2004; 78: 11334-9.

Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science. 2020; 367: 1444-1448. doi: 10.1126/science.abb2762. Epub 2020 Mar 4. PMID: 32132184

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. Published online February 3, 2020. https://doi.org/10.1038/s41586-020-2012-7.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., et al. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 382, 727–733.