

3-2021

Gene Variants of Coagulation Related Proteins that Interact with SARS-CoV-2

David Holcomb
U.S. FDA

Aikaterini Alexaki
U.S. FDA

Nancy Hernandez
U.S. FDA

Ryan Hunt
U.S. FDA

Jacob Kames
U.S. FDA

Below this page find additional works that are available at https://engagedscholarship.csuohio.edu/scibges_facpub

 Part of the [Biology Commons](#)

How does access to this work benefit you? Let us know!

Recommended Citation

Holcomb, David; Alexaki, Aikaterini; Hernandez, Nancy; Hunt, Ryan; Kames, Jacob; Hamasaki-Katagiri, Nobuko; Komar, Anton A.; DiCuccio, Michael; and Kimchi-Sarfaty, Chava, "Gene Variants of Coagulation Related Proteins that Interact with SARS-CoV-2" (2021). *Biological, Geological, and Environmental Faculty Publications*. 248.
https://engagedscholarship.csuohio.edu/scibges_facpub/248

This Article is brought to you for free and open access by the Biological, Geological, and Environmental Sciences Department at EngagedScholarship@CSU. It has been accepted for inclusion in Biological, Geological, and Environmental Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

Authors

David Holcomb, Aikaterini Alexaki, Nancy Hernandez, Ryan Hunt, Jacob Kames, Nobuko Hamasaki-Katagiri, Anton A. Komar, Michael DiCuccio, and Chava Kimchi-Sarfaty

RESEARCH ARTICLE

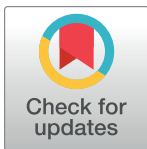
Gene variants of coagulation related proteins that interact with SARS-CoV-2

David Holcomb¹, Aikaterini Alexaki¹, Nancy Hernandez¹, Ryan Hunt¹, Kyle Laurie¹, Jacob Kames¹, Nobuko Hamasaki-Katagiri¹, Anton A. Komar², Michael DiCuccio^{3*}, Chava Kimchi-Sarfaty^{1*}

1 Center for Biologics Evaluation and Research, Office of Tissues and Advanced Therapies, Division of Plasma Protein Therapeutics, Food and Drug Administration, Silver Spring, Maryland, United States of America, **2** Center for Gene Regulation in Health and Disease, Department of Biological, Geological and Environmental Sciences, Cleveland State University, Cleveland, Ohio, United States of America, **3** National Center of Biotechnology Information, National Institutes of Health, Bethesda, Maryland, United States of America

These authors contributed equally to this work.

* dicuccio@ncbi.nlm.nih.gov (MD); Chava.kimchi-sarfaty@fda.hhs.gov (CKS)



OPEN ACCESS

Citation: Holcomb D, Alexaki A, Hernandez N, Hunt R, Laurie K, Kames J, et al. (2021) Gene variants of coagulation related proteins that interact with SARS-CoV-2. *PLoS Comput Biol* 17(3): e1008805. <https://doi.org/10.1371/journal.pcbi.1008805>

Editor: James M. Briggs, University of Houston, UNITED STATES

Received: August 3, 2020

Accepted: February 15, 2021

Published: March 17, 2021

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: This study was in part supported by the National Institutes of Health grant HL151392 (A.A.K.). URL: <https://www.nih.gov/>. This work was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health (M.D.). This work was partly supported by funds from the Hemostasis Branch/Division of Plasma Protein Therapeutics/Office of Tissues and Advanced Therapies/Center for Biologics

Abstract

Thrombosis is a recognized complication of Coronavirus disease of 2019 (COVID-19) and is often associated with poor prognosis. There is a well-recognized link between coagulation and inflammation, however, the extent of thrombotic events associated with COVID-19 warrants further investigation. Poly(A) Binding Protein Cytoplasmic 4 (PABPC4), Serine/Cysteine Proteinase Inhibitor Clade G Member 1 (SERPING1) and Vitamin K epoxide Reductase Complex subunit 1 (VKORC1), which are all proteins linked to coagulation, have been shown to interact with SARS proteins. We computationally examined the interaction of these with SARS-CoV-2 proteins and, in the case of VKORC1, we describe its binding to ORF7a in detail. We examined the occurrence of variants of each of these proteins across populations and interrogated their potential contribution to COVID-19 severity. Potential mechanisms, by which some of these variants may contribute to disease, are proposed. Some of these variants are prevalent in minority groups that are disproportionately affected by severe COVID-19. Therefore, we are proposing that further investigation around these variants may lead to better understanding of disease pathogenesis in minority groups and more informed therapeutic approaches.

Author summary

Increased blood clotting, especially in the lungs, is a common complication of COVID-19. Infectious diseases cause inflammation, which in turn can contribute to increased blood clotting. However, the extent of clot formation that is seen in the lungs of COVID-19 patients suggests that there may be a more direct link. We identified three human proteins that are involved indirectly in the blood clotting cascade and have been shown to interact with proteins of SARS virus, which is closely related to the novel coronavirus. We examined computationally the interaction of these human proteins with the viral proteins. We

Evaluation and Research of the U.S. Food and Drug Administration (D.H., A.A., K.L., J.K., N.H.-K., C.K.-S.). URL: <https://www.fda.gov/about-fda/fda-organization/center-biologics-evaluation-and-research-cber>. This study was partly supported by the U.S. Food and Drug Administration CBER Coronavirus (COVID-19) Supplemental Funding (C. K.S.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

looked for genetic variants of these proteins and examined how they are distributed across populations. We investigated whether variants of these genes could impact severity of COVID-19. Further investigation around these variants may provide clues for the pathogenesis of COVID-19, particularly in minority groups.

Introduction

The Coronavirus disease of 2019 (COVID-19) has been associated with coagulopathy, particularly microclots in the lungs [1–5], that correlates with disease severity [6–9]. There is extensive cross-talk between inflammation and coagulation, and inflammation is presumed to have a role in the observed coagulation phenotype. However, the widespread thrombotic events that are seen in severe COVID-19 patients suggest that there may be a more direct link.

In a study conducted before the onset of the COVID-19 pandemic, the severe acute respiratory syndrome (SARS) coronavirus (CoV)-host interactome was investigated. A few proteins related to the coagulation cascade were experimentally identified to interact with viral proteins (Fig 1). Poly(A) Binding Protein Cytoplasmic 4 (PABPC4) was shown to interact with the nucleocapsid (N) protein. Serine/Cysteine Proteinase Inhibitor Clade G Member 1 (SERPING1 or C1 inhibitor) was shown to interact with nsp14, ORF14, ORF3b, ORF7b, nsp2, nsp8 and nsp13. In addition, Vitamin K epoxide Reductase Complex subunit 1 (VKORC1) was shown to interact with the SARS protein ORF7a. The interactions were initially identified by a high-throughput yeast two-hybrid system and confirmed with LUMIER assay [10].

ORF7a - VKORC1 interaction

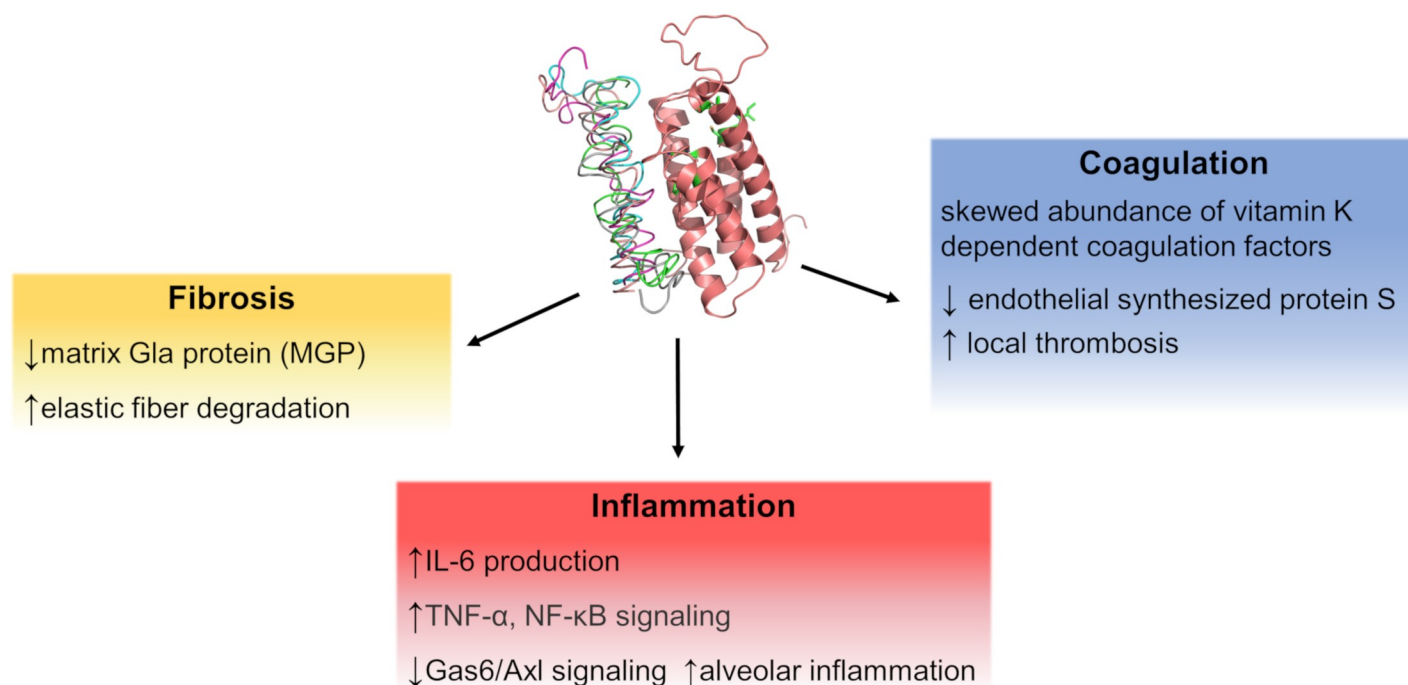


Fig 1. Graphic summary of ORF7a-VKORC1 interaction and possible effects. The interaction between ORF7a and VKORC1 and possible effects of this interaction.

<https://doi.org/10.1371/journal.pcbi.1008805.g001>

PABPC4 localizes primarily to the cytoplasm and binds to the poly(A) tail present at the 3-prime end of mRNA. However, it is also found in the surface of thrombin-activated platelets, and therefore it is known as activated-platelet protein-1 (APP-1) [11,12]. PABPC4 may also be involved in the regulation of protein translation in platelets and megakaryocytes may participate in the binding or stabilization of polyadenylates in platelet dense granules [13]. SERPING1 is a plasma protease involved in the complement, intrinsic coagulation and fibrinolytic pathways. In the coagulation cascade, SERPING1 inactivates plasma kallikrein, factor XIIa and factor XIIIf. The absence of sufficient levels of functional SERPING1 leads to hereditary angioedema (HAE), which is mediated by sustained activation of kallikrein leading to cleavage of high molecular weight kininogen (HMWK), producing bradykinin [14].

ORF7a is a viral protein that has not been well studied. While it counteracts the anti-viral properties of tetherin (BST2) [15,16], allowing for easier dispersal of virions, this protein has been found to be dispensable for viral replication in cell culture [17]. ORF7a may bind to Integrin beta chain-2 (ITGB2), a protein which is necessary for movement and phagocytosis in lymphocytes [18].

VKORC1 is an enzyme critical for coagulation due to its role in converting vitamin K epoxide into active vitamin K [19], the rate-limiting step in the physiological process of vitamin K recycling. Importantly, vitamin K is necessary for the carboxylation of glutamic acid residues to produce Gla residues. Several human proteins have domains with Gla residues, including coagulation factors II, VII, IX, X, and anticoagulant proteins C, S, and Z. VKORC1 is expressed in all tissues, but particularly in the liver, lungs, and female reproductive system. It is generally embedded in the endoplasmic reticulum [20].

Dietary vitamin K deficiency is associated with coagulopathy, specifically bleeding. Vitamin K antagonists are anticoagulant drugs that work by inhibiting the activity of VKORC1, reducing the levels of available active vitamin K and coagulation factors. Of the vitamin K antagonists, warfarin is most commonly used. Some variants in *VKORC1*, particularly those common in African and African American populations, are reported to result in warfarin resistance. Warfarin response is also dependent on dietary factors and liver function [21]. For these reasons, dosing warfarin is complicated, and genotyping of *VKORC1* to determine the presence of known polymorphisms (such as c.1173C>T) is recommended before initiating warfarin treatment.

The impact of viral protein interactions with VKORC1, SERPING1 and PABPC4 on patient outcomes in COVID-19 infection is unknown. While comorbidities, age, and other factors will impact the predisposition to thrombosis or coagulopathy, binding of viral proteins to coagulation related proteins may be partially responsible for the prothrombotic phenotype that is seen in COVID-19 patients.

Through computational modeling, we examined the binding of VKORC1, SERPING1 and PABPC4 to SARS-CoV-2 proteins and generated additional evidence for the binding of ORF7a to VKORC1. To further study the impact of these protein-protein interactions, we analyzed the effect of variants on patient outcomes and on protein function. We analyzed COVID-19 genome-wide association study (GWAS) results to find the most influential variants from these genes and characterize them to find potential causes of effect. Then, we investigated several *VKORC1*, *SERPING1* and *PABPC4* variants that may impact protein function and coagulopathy, and we identified some *VKORC1* variants that may result in warfarin resistance. In particular, we highlight two variants, which are enriched in certain ethnic groups. Better understanding of the contribution of these genes and their variants to COVID-19 pathogenesis may lead to new therapeutic avenues and improved prognosis. This may be of crucial importance for minority groups that are disproportionately affected by severe COVID-19.

Methods

Structural similarities and computational docking of proteins

To assess the binding of SARS-CoV-2 ORF7a and human VKORC1, we used I-TASSER [22–24] to generate homology models for both proteins. However, all ORF7a models showed an interaction between the luminal and transmembrane domains, which would cause clashing between the plasma membrane and the ORF7a luminal domain. For this reason, we used only the transmembrane domains of the models. Then, using the model with the best C-score, we used Zdock [25] to find potential binding sites. From Zdock, we used the protein-protein complex with the N-terminus of the ORF7a fragment closest to the luminal portion of VKORC1, which would introduce the least amount of clashing between the ORF7a luminal domain and the plasma membrane. This complex was used as input to Rosetta Prepack and Rosetta Dock [26–29] to further refine the models by using rigid body perturbations. The top five models were retained.

We examined known SARS-CoV-2 ORF7a variants [30,31]. However, all encountered variants were in the luminal domain and are unlikely to impact the interaction with VKORC1.

In addition, to verify the binding of PABPC4 and SERPING1 with SARS-CoV-2 proteins, we created homology models for each using I-TASSER and Robetta [32,33]. Template structures for all models are given in S1 Table. However, because segments of PABPC4 and SERPING1 have not been crystallized, these regions in the models were of low quality. For this reason, we used Blast and Clustal Omega to create multiple sequence alignments (MSAs) of proteins similar to interacting SARS proteins, and computed the percent of columns of the homologous SARS-CoV-2 protein matching the SARS protein, as well as a loglikelihood score to measure the probability that the SARS-CoV-2 homolog would be included in the MSA (Table 1). In addition, the MSAs were filtered to remove duplicate sequences by performing affinity propagation clustering with the Levenshtein distance matrix formed from the sequences. Only the cluster centers, SARS, and SARS-CoV-2 sequences were used in the MSA. This was done to account for the large number of very similar sequences, generally from different strains of SARS-CoV-2.

Table 1. Sequence homology of selected SARS and SARS-CoV-2 proteins.

Protein	Fraction Matching	Loglikelihood
N	0.888626	-0.04824
ORF7a	0.827869	-0.05688
nsp14	0.884393	-0.19572
ORF7b	0.795455	-0.01103
nsp3	0.75078	-0.46736
nsp2	0.681818	-0.44764
nsp8	0.969697	-0.12897
nsp13	0.948767	-0.1215

MSA fraction matching is the fraction of positions in the SARS-CoV-2 protein matching the homologous SARS protein, when both are aligned in an MSA. Higher number indicates more conserved position and the range is between 0 and 1.

MSA likelihood is the fraction of sequences in an MSA matching SARS-CoV-2 for a given column. Assuming all columns are independent, $\prod_i P(x_i)$ gives the probability of finding the SARS-CoV-2 sequence in the MSA sequences, which ranges between 0 and 1. Taking log of this value gives $\log(\prod_i P(x_i)) = \sum_i (\log P(x_i))$, an additive loglikelihood score which is nonpositive, with lower values indicating more positions in the SARS-CoV-2 sequence that differ from the MSA sequences.

<https://doi.org/10.1371/journal.pcbi.1008805.t001>

We used the ORF7a homology model to query Dali [34] for similar protein structures. The top structures in sequence and structural similarity were the ORF7a proteins for SARS and SARS-CoV-2 (PDBs 1YO4 and 6W37). All human proteins interacting with VKORC1 were taken from BIOGRID, the Biological General Repository for Interaction Datasets [35,36]. In addition, we queried Dali against all other viral protein structures, as modeled in I-TASSER.

Relevant variants from COVID19 HGI GWAS metastudies

All variants from the genomic region containing *VKORC1*, *SERPING1*, and *PABPC4* ± 6000 bp were taken from the ANA2, ANA5, and ANA7 metastudies from COVID19 Host Genetics Initiative [37] and The Severe Covid-19 GWAS Group [38] (Tables 2 and S2). We filtered the resulting variants to keep only those with metastudy p-value below 0.05. The resulting variants were all in non-coding regions, therefore, amino acid and codon features do not apply.

We characterized these variants in terms of splicing, using hexamer scoring tools [39,40], ESEfinder [41,42], ExonScan [43–45], and FAS-ESS [43]. Where ESEfinder, ExonScan, and FAS-ESS found a change in splicing potential between the wild type (WT) and mutant, the change was reported in Table 2 as “Change in splicing”. When the variant occurred in an intron as opposed to a UTR, we further highlighted the value.

Then, we calculated mRNA mean free energy using Kinefold [46], mFold [47–49], and remuRNA [50]. When all three tools were in agreement regarding the direction of the change, the changes in mRNA MFE were converted into Z-scores using mean and standard deviation values computed by randomly sampling WT and mutant sequences. The average of the three Z-scores is reported in Table 2 as “Average change in mRNA MFE (Z-score)”.

We also analyzed miRNA binding changes using miRDB [51,52]. For any variant, there may be multiple affected miRNA species. miRNA binding scores are provided for both the WT and mutant flanking 501 nucleotides in S3 Table, and a summary of miRNA binding changes is provided in Table 2. When all miRNA binding changes were in the same direction, we summarized the effect.

We analyzed conservation using fraction matching in a nucleotide MSA, computed as the fraction of sequences in the MSA matching the wild type sequence in the appropriate column. This value is included in Table 2 as “Fraction matching in MSA”.

Finally, we collected population prevalence data from dbSNP (Tables 3 and S3).

Characterization of synonymous and missense variants of coagulation genes of interest

We found all synonymous (S4 Table) and missense (S5 Table) variants of *VKORC1*, *SERPING1* and *PABPC4* genes [53] from NCBI’s Single Nucleotide Polymorphism Database (dbSNP) [54] and characterized them in terms of (i) population prevalence in the Genome Aggregation Database (gnomAD) [55,56], (ii) the percent of sequences matching the WT at that position in a multiple sequence alignment (MSA) [57], (iii) likelihood of the variant in the column of an MSA, (iv) mRNA MFE computed by both Kinefold and mFold, (v) relative synonymous codon usage (RSCU) and (vi) relative synonymous codon pair usage (RSCPU) [58,59], (vii) rare codon enrichment [60], (viii) and %MinMax codon usage [61]. For nonsynonymous variants, we additionally used amino acid fraction matching in an MSA, likelihood of the variant amino acid in an amino acid MSA, SIFT [62,56], and Polyphen [63,56]. The fraction matching and MSA likelihood measures use sequence homology and may imply selection against the variant. SIFT uses sequence homology as well as physical properties of amino acids, while Polyphen uses multiple sequence and structural features to predict the effect of amino acid substitutions. MFE of mRNA may affect stability of mRNA transcripts, which will affect

Table 2. Possible predicted effect of variants in VKORC1, SERPING1 and PABPC4.

Transcript	Location	Fraction matching in MSA	Change in splicing	Average change in mRNA MFE (Z-score)	miRNA summary
VKORC1					
NM_024006.4:c.-4931C>T	5' UTR	0.431818		2.20025	
NM_024006.4:c.-4851C>T	5' UTR	0.754967		1.06847	
NM_024006.4:c.-2834C>A	5' UTR	0.950943			<u>miRNA gained</u>
NM_024006.4:c.-1639G>A	5' UTR	0.0625	Possible splicing change		
NM_024006.4:c.174-136C>T	Intron	0.020228		0.83549	
NM_024006.4:c.283+124G>C	Intron	0.133333		-1.44871	
NM_024006.4:c.283+837T>C	Intron	0.146727		1.35367	
SERPING1					
NM_000062.2:c.-3537C>G	5' UTR	0.009615			<u>miRNA decrease</u>
NM_000062.2:c.-2415G>A	5' UTR	0.033708	Possible splicing change	0.77565	
NM_000062.2:c.-1675G>A	5' UTR	0.426901			<u>miRNA gained</u>
NM_000062.2:c.52-696C>T	Intron	0.068027	<u>Likely splicing change</u>	0.30003	
NM_000062.2:c.52-130C>T	Intron	0.833333			
NM_000062.2:c.52-130C>T	Intron	0.833333		-0.28155	
NM_000062.2:c.550+794C>A	Intron	0.693694		0.71906	
NM_000062.2:c.685+88G>A	Intron	0.769231		-0.75902	
NM_000062.2:c.685+659C>T	Intron	0.581818		-0.47852	miRNA decrease
NM_000062.2:c.685+659C>T	Intron	0.679245		-0.68662	
NM_000062.2:c.685+1100C>T	Intron	0.772313		2.01226	
NM_000062.2:c.685+1391C>T	Intron	0.841912	<u>Likely splicing change</u>		
NM_000062.2:c.685+1550G>T	Intron	0.926641			
NM_000062.2:c.685+1770C>T	Intron	0.793594		0.84106	
NM_000062.2:c.1029+926G>T	Intron	0.015723			miRNA gained
NM_000062.2:c.1029+1443G>C	Intron	0.595745			miRNA gained
NM_000062.2:c.1029+2110T>C	Intron	0.393617	<u>Likely splicing change</u>		
NM_000062.2:c.1029+2111G>A	Intron	0.687117			miRNA gained
NM_000062.2:c.1030-2243T>G	Intron	0.026616		-1.68665	
NM_000062.2:c.1030-1975G>C	Intron	0.02551		1.92651	
NM_000062.2:c.1030-1436T>C	Intron	0.823529			
NM_000062.2:c.1030-20A>G	Intron	0.25		-0.76751	
NM_000062.2:c.1438G>A	Exon	0.399177	Possible splicing change		
NM_000062.2:c.*1323G>A	3' UTR	0.878788			miRNA lost
NM_000062.2:c.*1521G>T	3' UTR	0.65873			
NM_000062.2:c.*2614A>T	3' UTR	0.016129			miRNA gained
PABPC4					
NM_003819.3:c.-5600T>C	5' UTR	0.48			<u>miRNA lost</u>
NM_003819.3:c.-4432G>A	5' UTR	0.009317			<u>miRNA gained</u>
NM_003819.3:c.-4428A>G	5' UTR	0.221505			
NM_003819.3:c.-3677T>G	5' UTR	0.021645	Possible splicing change	-0.59763	
NM_003819.3:c.-3636G>A	5' UTR	0.022378		2.31727	
NM_003819.3:c.-3198T>C	5' UTR	0.856079	Possible splicing change		<u>miRNA lost</u>
NM_003819.3:c.-2286T>G	5' UTR	0.210526			
NM_003819.3:c.-650C>T	5' UTR	0.829978			<u>miRNA lost</u>
NM_003819.3:c.193+796C>G	Intron	0.666667			
NM_003819.3:c.504-254C>A	Intron	0.247191		0.26322	
NM_003819.3:c.738+85T>C	Intron	0.333333			miRNA lost

(Continued)

Table 2. (Continued)

Transcript	Location	Fraction matching in MSA	Change in splicing	Average change in mRNA MFE (Z-score)	miRNA summary
NM_003819.3:c.877-387C>T	Intron	<u>1</u>			miRNA lost
NM_003819.3:c.972+53A>T	Intron	<u>1</u>			
NM_003819.3:c.972+704C>G	Intron	0.5			
NM_003819.3:c.1333+26C>G	Intron	0.3125	<u>Likely splicing change</u>		
NM_003819.3:c.1621-348C>G	Intron	<u>1</u>			
NM_003819.3:c.*765C>A	3' UTR	<u>1</u>		<u>1.96974</u>	
NM_003819.3:c.*1261C>T	3' UTR	0.771242		-0.91257	miRNA decrease
NM_003819.3:c.*4685A>G	3' UTR	0.054945		-3.4524	miRNA decrease
NM_003819.3:c.*5316C>T	3' UTR	0.696181	Possible splicing change		miRNA lost

Change in splicing is presented when all tools find a change in splicing and all hexamer scores are greater than one standard deviation from the mean, and is marked in red when the variant appears in an intron. mRNA MFE changes are normalized (converted into a Z-score) for KineFold, remuRNA, and mFold, then averaged. When all three mRNA MFE changes are above one standard deviation, we mark the value in underline. miRNA summaries are presented when all miRNA changes agree in direction, and the total change is at least 5. miRNA changes are underlined when the variant appears upstream.

<https://doi.org/10.1371/journal.pcbi.1008805.t002>

transcript abundance and translation. Codon and codon pair usage have been shown to impact translation kinetics [64,65], and their metrics may be useful in assessing the impact of synonymous mutations on protein conformation and function [58]. For all variants, we provide the corresponding identifier in dbSNP ("rs" ID) [54].

We applied filters based on codon usage changes, mRNA MFE changes, and position conservation to identify variants that were potentially impactful on protein expression or conformation, which may affect interactions with SARS-CoV-2 proteins. Then, based on population frequencies, we computed the probability of the presence of at least one filtered variant in each population, and compared with the overall probability.

For a summary of the meaning, use, and range of all scoring tools, see [S6 Table](#).

Results

Computational verification of SARS-CoV-2 viral protein interactions

To study the role of coagulation in COVID-19 pathogenesis, we explored the interactions of VKORC1, SERPING1 and PABPC4 with viral proteins through computational docking. VKORC1 and ORF7a were confirmed to have strong binding affinity. Interactions are generally limited to transmembrane helices as opposed to intervening loops where warfarin is known to bind [66]. The top scoring complexes are shown in [Fig 2](#). Plots of interface energy in Rosetta energy units against interface root mean square error for the RosettaDock results are given in [Fig 3](#). The plots show convergence toward the minimum energy state.

When modeling the entire ORF7a protein, the models showed the luminal and transmembrane domains bound to one another. Due to this interaction, any dockings of ORF7a and VKORC1 implied the luminal domain of ORF7a would clash with the plasma membrane. We believe this interaction between the ORF7a transmembrane and luminal domains is predicted because the transmembrane domain has not been structurally characterized, and because the protein structure prediction software will minimize the folding energy by creating this bond between the two domains. To verify this, we used EVmutation [67] to compute the position co-evolution parameters, which measure dependency between positions in the amino acid sequence and are used to predict amino acid contacts and bonds. We found less coupling between the transmembrane and luminal domains than within each domain, which indicates lower likelihood of interaction between the domains.

Table 3. Population frequencies (gnomAD) of GWAS variants of VKORC1, SERPING1, and PABPC4.

Transcript	Global	African	American	Ashkenazi Jewish	East Asian	European	Other
VKORC1							
NM_024006.4:c.-4931C>T	0.5758	0.5408	0.542	0.514	<u>0.1007</u>	0.63146	0.619
NM_024006.4:c.-4851C>T							
NM_024006.4:c.-2834C>A	0.0049	0.0001	0.001	0	0	0.00762	0.007
NM_024006.4:c.-1639G>A	0.326	<u>0.1009</u>	0.444	0.476	<u>0.8996</u>	0.37236	0.369
NM_024006.4:c.174-136C>T	0.3261	<u>0.1009</u>	0.443	0.476	<u>0.8995</u>	0.37264	0.37
NM_024006.4:c.283+124G>C	0.4163	0.2564	0.442		<u>0.8849</u>		
NM_024006.4:c.283+837T>C	0.6431	0.7907	0.546	0.517	<u>0.1017</u>	0.62682	0.628
NM_014699.3:c.*2082G>C	0.0049	0.0001	0.001	0	0	0.00763	0.007
NM_014699.3:c.*2737G>T	0.0048	0.017	0.002	0	0	0.00005	0
SERPING1							
NM_000062.2:c.-3537C>G	0.0275	0.0078	0.021	0.017	0	0.03846	0.041
NM_000062.2:c.-2415G>A	0.0939	0.0866	0.059	0.141	0.1113	0.09708	0.087
NM_000062.2:c.-1675G>A	0.0937	0.0862	0.059	0.141	0.1105	0.09697	0.087
NM_000062.2:c.52-696C>T	0.3927	0.4767	0.529	0.452	0.7655	0.31673	0.386
NM_000062.2:c.52-130C>T	0.385	0.448	0.525	0.455	0.7668	0.31739	0.385
NM_000062.2:c.52-130C>T	0.385	0.448	0.525	0.455	0.7668	0.31739	0.385
NM_000062.2:c.550+794C>A	0.3936	0.4761	0.531	0.451	0.7697	0.31779	0.388
NM_000062.2:c.685+88G>A	0.2225	0.1006	0.15	0.262	0.1157	0.28733	0.273
NM_000062.2:c.685+1391C>T	0.0248	0.0059	0.022	0.035	0	0.03499	0.038
NM_000062.2:c.685+659C>T	0.3901	0.4743	0.541	0.455	0.769	0.31084	0.381
NM_000062.2:c.685+659C>T	0.3901	0.4743	0.541	0.455	0.769	0.31084	0.381
NM_000062.2:c.685+1100C>T	0.2253	0.1124	0.147	0.262	0.1208	0.28765	0.274
NM_000062.2:c.685+1550G>T	0.2251	0.1127	0.152	0.264	0.1207	0.28734	0.269
NM_000062.2:c.685+1770C>T	0.2216	0.0992	0.15	0.262	0.1184	0.28696	0.27
NM_000062.2:c.1029+926G>T	0.2279	0.1	0.15	0.264	0.1224	0.29523	0.284
NM_000062.2:c.1029+1443G>C	0.2282	0.1004	0.15	0.269	0.1198	0.29577	0.284
NM_000062.2:c.1029+2110T>C	0.612	0.5191	0.469	0.538	0.2347	0.69271	0.624
NM_000062.2:c.1029+2111G>A	0.227	0.1003	0.15	0.264	0.1183	0.29407	0.283
NM_000062.2:c.1030-2243T>G	0.6129	0.5195	0.47	0.541	0.2387	0.69335	0.626
NM_000062.2:c.1030-1975G>C	0.0113	0.0022	0.008	0.024	0	0.01647	0.008
NM_000062.2:c.1030-1436T>C	0.0045	0.0014	0.001	0.003	0	0.00645	0.004
NM_000062.2:c.1030-20A>G	0.6134	0.5197	0.472	0.541	0.2461	0.69353	0.623
NM_000062.2:c.1438G>A	0.2282	0.1007	0.15	0.269	0.1202	0.29561	0.285
NM_000062.2:c.*1323G>A	0.2283	0.1009	0.151	0.269	0.1175	0.29578	0.285
NM_000062.2:c.*1521G>T	0.1496	0.0855	0.166		0.0942		
NM_000062.2:c.*2614A>T	0.6058	0.4936	0.463	0.538	0.2277	0.69504	0.626
PABPC4							
NM_003819.3:c.-5600T>C	0.8127	0.5571	0.918	0.945	0.9909	0.90666	0.891
NM_003819.3:c.-4432G>A	0.0403	<u>0.0093</u>	<u>0.095</u>	0.024	<u>0.3712</u>	0.02506	0.047
NM_003819.3:c.-4428A>G	0.0432	<u>0.0096</u>	<u>0.1</u>	0.024	<u>0.3712</u>	0.02918	0.052
NM_003819.3:c.-3677T>G	0.1792	<u>0.0535</u>	0.122	0.247	0.1111	0.24258	0.219
NM_003819.3:c.-3636G>A	0.0052	0.0027	0.004	0.007	0	0.0068	0.007
NM_003819.3:c.-3198T>C	0.0025	0.001	0.002	0	0	0.00329	0.004
NM_003819.3:c.-2286T>G	0.0136	0.0031	0.014	0.01	0	0.01952	0.015
NM_003819.3:c.-650C>T	0.0079	0.0027	0.002	0	0	0.01172	0.008
NM_003819.3:c.193+796C>G	0.8013	<u>0.509</u>	0.913	0.945	0.9904	0.90714	0.895

(Continued)

Table 3. (Continued)

Transcript	Global	African	American	Ashkenazi Jewish	East Asian	European	Other
NM_003819.3:c.504-254C>A	0.1438	<u>0.0321</u>	0.079	0.2	0.1093	0.19797	0.183
NM_003819.3:c.738+85T>C	0.0573	<u>0.1955</u>	0.012	0.01	<u>0.0013</u>	<u>0.00354</u>	0.014
NM_003819.3:c.877-387C>T	0.113	0.0243	0.065	0.131	0.1086	0.15452	0.146
NM_003819.3:c.972+53A>T	0.0018	0.0065	0	0	0	0	0
NM_003819.3:c.972+704C>G	0.0025	0.001	0.002	0	0	0.00328	0.004
NM_003819.3:c.1333+26C>G	0.0006	0.0001	0	0	0	0.00095	0
NM_003819.3:c.1621-348C>G	0.0003	0.0001	0	0	0	0.00042	0
NM_003819.3:c.*765C>A	0.0403	<u>0.0086</u>	<u>0.096</u>	0.024	0.3656	0.02561	0.049
NM_003819.3:c.*1261C>T	0.0073	0.0256	0	0	0.0006	0.00005	0.003
NM_003819.3:c.*4685A>G	0.7986	<u>0.4999</u>	0.911	0.945	0.991	0.90696	0.894
NM_003819.3:c.*5316C>T	0.007	0.0028	0.008	0	0	0.00955	0.008

Population frequencies are taken from dbSNP. Populations with greater distance from global distribution are underlined.

<https://doi.org/10.1371/journal.pcbi.1008805.t003>

For these reasons, and because the luminal and transmembrane domains of ORF7a are connected by a flexible loop region, we repeated the docking, excluding the ORF7a luminal domain. Docking showed strong interaction with the minimum energy -36.078 Rosetta energy units at the interface between ORF7a and VKORC1. The Rosetta energy in this case measures the impact of interactions and bonds at the interface between ORF7a and VKORC1.

Regarding SERPING1 and PABPC4, due to the lack of structural data for some segments, portions of the models for PABPC4 and SERPING1 were of low quality. Therefore, we continued our analysis by examining sequence homology of SARS-CoV-2 proteins to SARS proteins. Predictably, the homology was high (Table 1), suggesting that homologous SARS-CoV-2 proteins maintain interactions with human proteins as observed for SARS proteins. Specifically, several SARS proteins were found to interact with SERPING1, so it is likely that SARS-CoV-2 proteins interact with SERPING1 too. In addition, PABPC4 was found experimentally to bind to SARS-CoV-2 N protein [68].

Variants that may impact COVID-19 severity

GWAS metastudies on COVID-19 outcomes recently became available [37,38]. We focused on the impact of VKORC1, SERPING1, and PABPC4 gene variants on COVID-19 severity. While over 700 variants from these genes were found in the studies, only 55 variants had a p-value less than 0.05; these are listed in Tables 2 and 3. However, none of them are significantly impactful when controlling for multiple hypothesis testing. Only one variant is a coding variant and may affect protein-protein interactions. However, the non-coding variants may affect translation or splicing, leading to lower availability of protein. We characterized the 55 variants in terms of miRNA binding, splicing, mRNA minimum free energy, and sequence conservation, to understand how they may affect disease outcomes. miRNAs are involved in post-transcriptional regulation by binding to mRNA transcripts, resulting in degradation of the mRNA or less efficient translation. Therefore, higher binding will most likely result in lower expressing protein. Summaries of miRNA changes are given in Tables 2 and S2, and full data is given in S3 Table. Interestingly, for variants which effected a change in miRNA binding potential, most caused a reduction in miRNA binding potential, which may increase protein expression. The mean change between variant and wild type miRNA affinity predictions is -11.72414, and the median is -1.

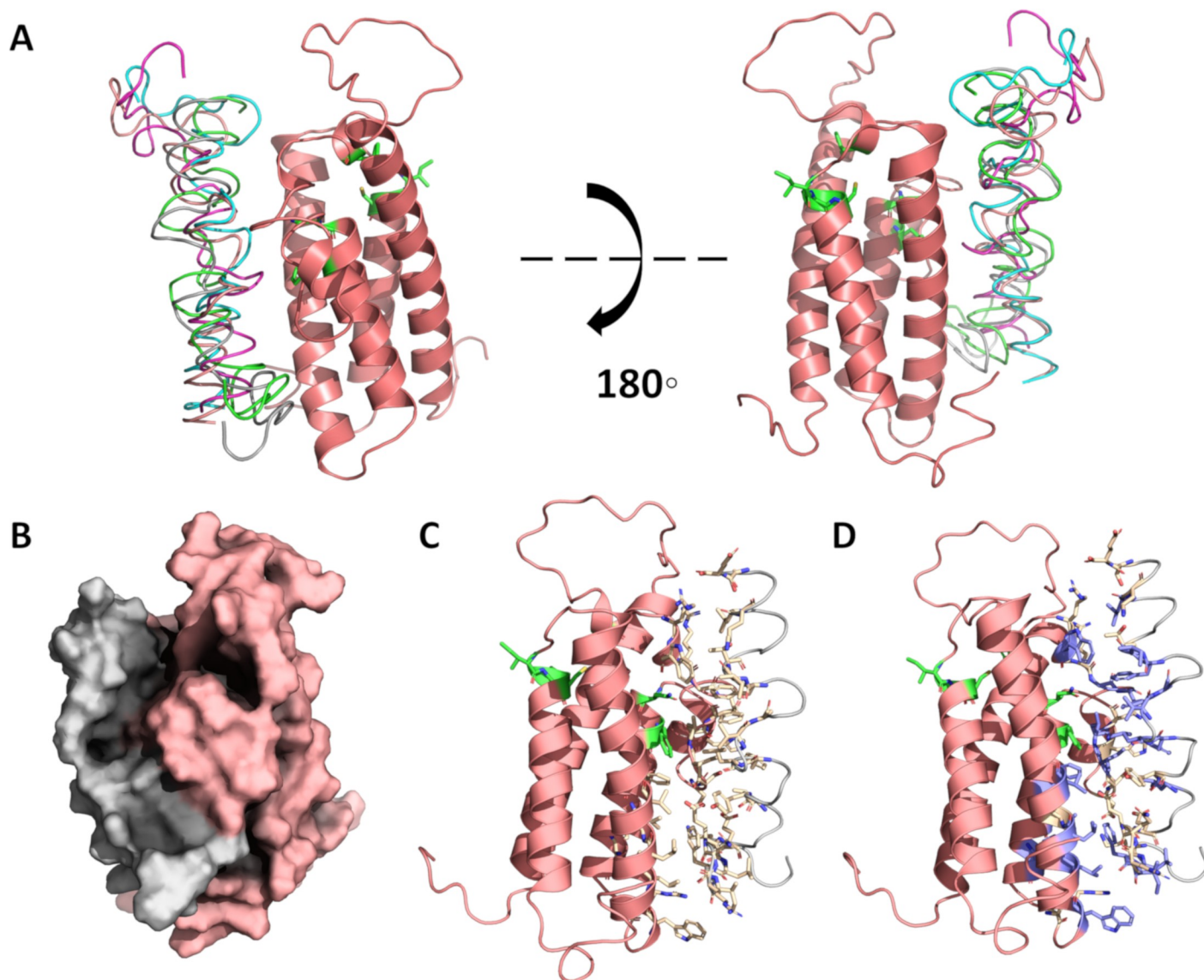


Fig 2. Predicted dock of VKORC1 and ORF7a transmembrane domain. A. Five protein-protein docks depict one main binding site (teal, grey, yellow, green, blue). B. The lowest interface-energy model is shown as a surface representation. C. The lowest interface-energy model, with side chains shown in wheat for amino acids at the interface. D. Another view of the lowest interface-energy model, with side chains shown in wheat at the interface and hydrophobics shown in blue. Amino acids of VKORC1 necessary for vitamin K binding (83F, 80N, 135C, 55F) or warfarin binding (134V, 133I) are given in green.

<https://doi.org/10.1371/journal.pcbi.1008805.g002>

Splicing is involved in the production of mature mRNAs for many genes. Changes in splicing may produce alternative mature mRNAs, preventing accurate translation, and thus resulting in a protein with altered potency or affinity to the virus. While we consider splicing dysregulation as potentially impacting gene expression and disease outcome, it has rarely been shown experimentally. In vitro testing of some of these variants did not reveal differences between the splice forms and WT or substantial differences in expression. For example, Wang et al [69] examined the VKORC1 polymorphisms -1639G>A (rs9923231), 1173C>T (rs9934438), and c.-4931C>T (rs7196161) in various cell lines and did not detect any differences in expression levels. We found several intronic variants in all three genes which resulted in large changes in predicted splicing potential (Table 2). Of these, NM_000062.2:c.52-

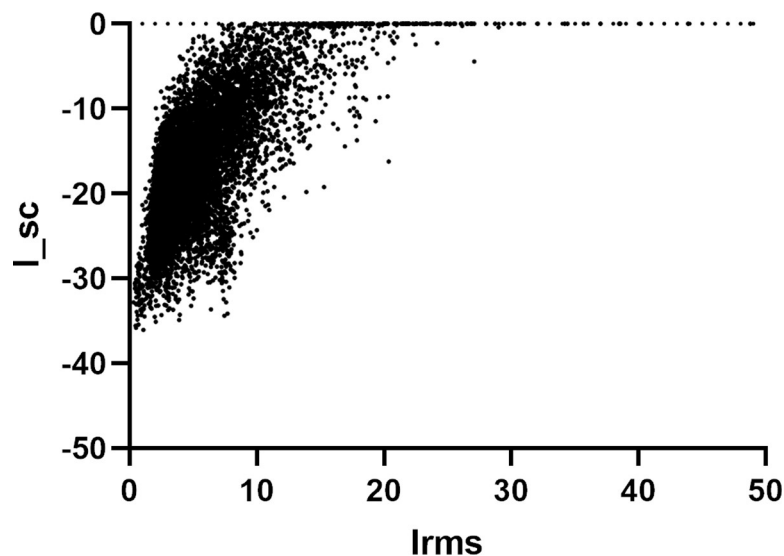


Fig 3. Plots of interface energy (I_{sc}) against interface root mean square error (I_{rms}). Each point represents a complex formed from one of the top 5 ZDock outputs of VKORC1 and ORF7a proteins, using 10,000 decoys. All plots form energy funnels.

<https://doi.org/10.1371/journal.pcbi.1008805.g003>

696C>T is more common in East Asian populations, NM_000062.2:c.1029+2110T>C is more common in European populations, and NM_000062.2:c.685+1391C>T and NM_003819.3:c.1333+26C>G are comparatively rare globally.

Finally, sequence conservation gives an evolutionary view of the significance of any position in a sequence, but it is dependent on the conservation model and the quality of sequence and structural data. Several PABPC4 variants show perfect conservation at the variant position. The full data are given in [S2 Table](#).

We found several upstream variants in VKORC1 that resulted in higher predicted miRNA binding affinity, suggesting lower expression of the protein. Of these, NM_000062.2:c.-1675G>A is relatively common in all populations (9.37% MAF). We also found several upstream variants in PABPC4 that resulted in lower predicted miRNA binding affinity suggesting higher expression of the protein.

mRNA molecules will form secondary structures based on nucleotide arrangement and affinity, which impact their structural stability. We found several variants resulting in large changes in mRNA stability. For example, NM_000062.2:c.685+1100C>T, NM_000062.2:c.1030-1975G>C, and NM_003819.3:c.*765C>A are all strongly predicted to destabilize their respective mRNA transcripts. Higher MFE may suggest higher possibility for mRNA degradation, which leads to decreased availability of transcripts and lower expression. These variants may increase mRNA degradation, reducing protein expression.

In addition, known clinical consequences of these variants are summarized in [Table 4](#).

Prevalence of VKORC1 variants across populations

COVID-19 has spread to the entire world, affecting people with variable genetic and racial backgrounds. Therefore, we explored ORF7a interactions with variants of VKORC1 found across races. There are 160 missense VKORC1 variants in dbSNP and at least 27 which affect warfarin sensitivity [91]. The most common variants are shown in [Table 5](#). The locations of the warfarin sensitive variants are shown in [Fig 4](#). However, many warfarin resistance-causing variants are not listed in dbSNP, and some do not include population frequency information.

Table 4. Variations' clinical impact.

Variation	Clinical Impact based on Literature
VKORC1	
NM_024006.4: c.283+837T>C	South Indians carrying the C nucleotide require lower warfarin dosages relative to WT (T) [70].
NM_024006.4: c.283+124G>C	European Americans carrying the G nucleotide require lower warfarin dosages relative to WT (C) [71].
NM_024006.4: c.174-136C>T	Turkish carrying the T nucleotide require lower warfarin dosages [72]; African Americans and European Americans carrying T nucleotide require lower warfarin dosages relative to WT (C) [73].
NM_024006.4: c.-1639G>A	Chinese carrying the A nucleotide require lower warfarin dosages relative to WT (G) [74].
NM_024006.4: c.-4931C>T	South Indians carrying the T nucleotide require increased warfarin dosages relative to WT (C) [75].
SERPING1	
NM_000062.2: c.52-130C>T	Patients carrying the T nucleotide depicted worsened progression for age-related macular degeneration relative to WT (C) [76]; Chinese and Japanese carrying the T nucleotide lack an association with age-related macular degeneration, seen in Caucasian population studies, although was predicted as pathogenic [77].
NM_000062.2: c.1029+2110T>C	European and Mediterranean patients carrying the C nucleotide did not depict a higher association with hereditary angioedema relative to WT [78].
NM_000062.2: c.1030-1975G>C	The intronic polymorphism 1030+1975G>C has no pathogenic influence on hereditary angioedema although predicted as pathogenic [78].
NM_000062.2: c.1030-20A>G	Association of the G allele with age-related macular degeneration was predicted to decrease the variant splicing form SERPING1, decrease protein expression and potentially limit the regulation of the compliment system [79]. No association was observed for Chinese Han carrying the G nucleotide with age-related macular degeneration [80].
NM_000062.2: c.-2415G>A	Chinese Han patients carrying the A nucleotide did not demonstrate an increased risk of polypoidal choroidal vasculopathy relative to WT (G) [81]. South Korean patients carrying the A nucleotide did not show association with an increased risk of leukemia relative to WT (G) [82]. Caucasians carrying the A nucleotide did not exhibit an increased risk of age-related macular degeneration relative to WT (G) [83].
NM_000062.2: c.52-696C>T	Patients carrying the T nucleotide did not display an increased risk for anterior uveitis relative to WT (C) [84]. Chinese Han carrying the T nucleotide did not display an increased risk for polypoidal choroidal vasculopathy relative to WT (C) [81]. Caucasians carrying the T nucleotide did not display an increased risk for age-related macular degeneration relative to WT (C) [83]. Chinese carrying the T nucleotide did not display an increased risk for diabetic retinopathy relative to WT (C) [85]. European and Mediterranean's carrying the T nucleotide did not display an increased risk for hereditary angioedema relative to WT (C) [78].
NM_000062.2: c.52-130C>T	Chinese carrying the T nucleotide did not display a different association with age-related macular degeneration relative to WT (C) [80]. Caucasians carrying the T nucleotide displayed worsened progression of age-related macular degeneration relative to WT (C) [86]. Patients carrying the T nucleotide depicted worsened symptoms of age-related macular degeneration relative to WT (C) [76]. Chinese carrying the T nucleotide responded poorer to anti-VEGF treatment relative to WT (C) [87].
NM_000062.2: c.685+659C>T	Caucasians carrying the A nucleotide failed to depict a greater association with AMD relative to WT (G) [83]. South Korean patients carrying the A nucleotide did not depict a greater association with leukemia relative to WT (G) [88]. Han Chinese carrying the A nucleotide did not depict a significantly greater association with age-related macular degeneration relative to WT (G) [81].
NM_000062.2: c.685+1100C>T	European and Mediterranean patients carrying the T nucleotide failed to show a greater association with hereditary angioedema relative to WT (C) [78].

(Continued)

Table 4. (Continued)

Variation	Clinical Impact based on Literature
NM_000062.2: c.1029+926G>T	European and Mediterranean patients carrying the T nucleotide failed to show a greater association with hereditary angioedema relative to WT (G) [78]. Chinese Han carrying the T nucleotide did not depict a greater association with polypoidal choroidal vasculopathy relative to WT (G) [81].
NM_000062.2: c.1029+1443G>C	European and Mediterranean patients carrying the C nucleotide failed to show a greater association with hereditary angioedema relative to WT (G) [78].
NM_000062.2: c.1029+2111G>A	European and Mediterranean patients carrying the A nucleotide failed to show a greater association with hereditary angioedema relative to WT (G) [78].
NM_000062.2: c.1438G>A	Patients carrying the A nucleotide did not depict a change in Tacrolimus dosage requirements for transplant operations relative to WT (G) [89]. Chinese Han patients carrying the A nucleotide did not show a higher association with age-related macular degeneration or polypoidal choroidal vasculopathy relative to WT (G) [80].
PABPC4	
NM_003819.3: c.504-254C>A	Increased risk for type 2 diabetes with the 40035928G>T polymorphism based on GWAS studies [90].

<https://doi.org/10.1371/journal.pcbi.1008805.t004>

In addition, there are several intronic, upstream and downstream variants which impact warfarin dosage [92]. For example, rs9923231 (c.-1639G>A, NG_011564.1:g.3588G>A), which causes warfarin sensitivity, is very common in East Asian populations (89.95%) and comparatively less common in African populations (10.09%), with intermediate frequency for other populations.

Table 5. Population frequencies of missense and synonymous VKORC1 variants.

	VKORC1 variant	Warfarin Sensitivity	Prevalence								
			Overall	African	Latino	Ashkenazi Jewish	East Asian	Finnish	Non-Finnish European	Other groups	South Asian
Missense	106GT	Resistance	0.00241	0	0.00166	0.03857	0	5.6E-05	0.00069	0.00460	0.00091
	203AG		0.00044	0	0	0	0.00571	0	8.8E-06	0.00016	9.8E-05
	352GC		0.00036	0	0	0	0.00359	0	0	0.00016	0.00078
	202CT		0.00036	6.2E-05	0	0	0	0.00083	0.00059	0.00065	0
	79CG		0.00031	0	0	0	0	0	6.8E-05	0	0.00217
	196GA		0.00020	0.00265	0.00014	0	0	0	0	0	3.3E-05
	427GA		0.00017	0	0.00116	0	0	0	8.8E-06	0.00016	0
	390TG		0.00012	0	0	0	0.00169	0	0	0	0
	157CA		0.00011	0	0	0	0.00139	0	0	0.00017	0
	163TC		1.0E-04	0	0.00067	0	0	0	0	0.00017	0
Synonymous	358CT	Resistance	0.01558	0.19520	0.01186	0.00626	0.00011	0	0.00170	0.01044	0.00065
	36GA	Resistance*	0.01511	0.00210	0.01583	0.02262	5.6E-05	0.07348	0.01171	0.02296	0.00200
	129CT		0.003643	0.00115	0.00093	0.00115	0	0.00504	0.00567	0.00415	0.00282
	54GT		7.7E-05	0	0.00030	0	0.00023	0	2.9E-05	0	3.3E-05
	234AG		3.6E-05	0	0.00012	0	0	0	3.5E-05	0.00016	0
	54GC		2.6E-05	7.1E-05	3.0E-05	0	0	0	2.9E-05	0	3.3E-05
	18GA		1.7E-05	0	9.0E-05	0	0	0	9.8E-06	0	0
	111GA		1.7E-05	0	3.0E-05	0	0.00011	0	9.8E-06	0	0
	72CT		1.7E-05	0	0	0	0	0	3.9E-05	0	0
	186TG		1.6E-05	0	0	0	0	0	3.5E-05	0	0

Warfarin sensitivity is determined by literature review.

<https://doi.org/10.1371/journal.pcbi.1008805.t005>

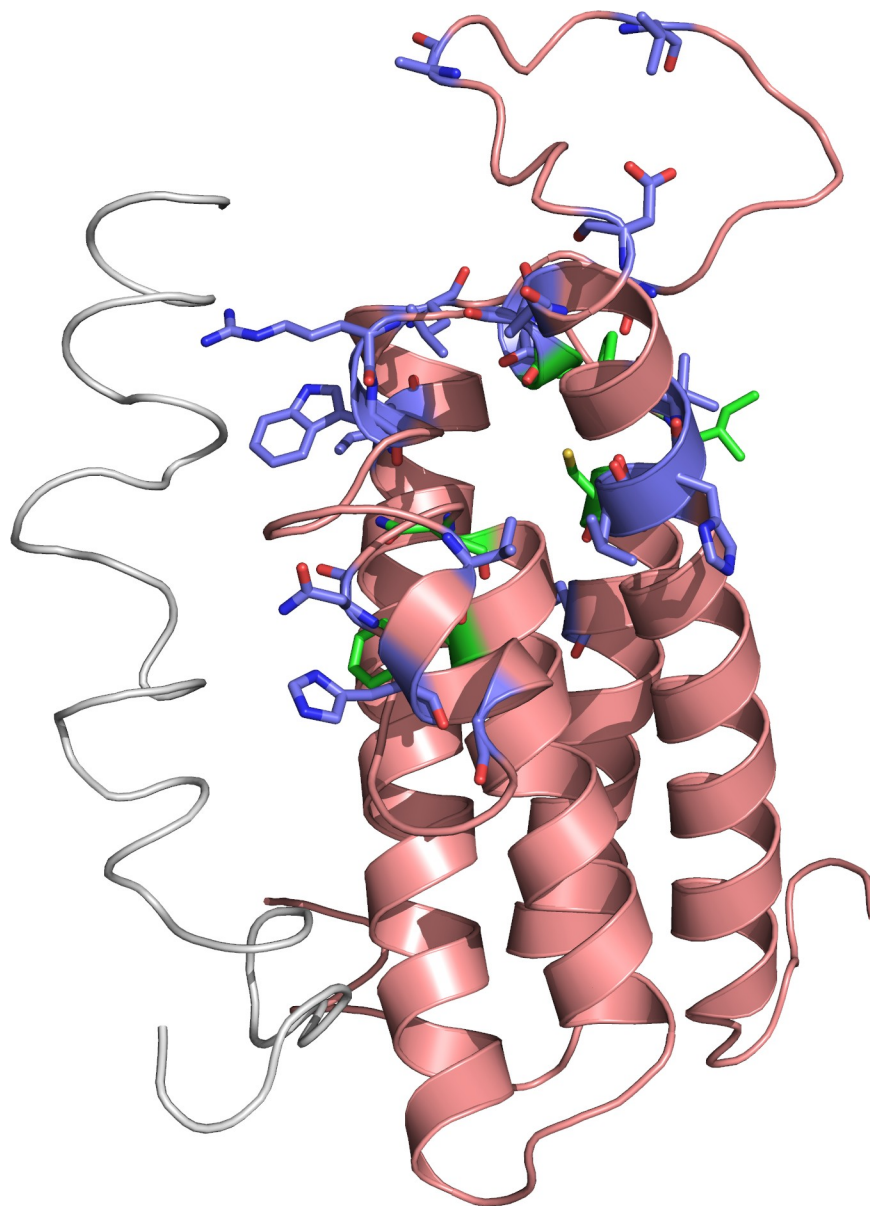


Fig 4. Locations of warfarin dosage affecting nonsynonymous variants in VKORC1. VKORC1 is shown in salmon, while ORF7a is shown in grey. Warfarin dosage affecting nonsynonymous variants are shown in blue. Active site amino acids are shown in green.

<https://doi.org/10.1371/journal.pcbi.1008805.g004>

In the United States, COVID-19 has disproportionately affected African American populations. We sought to investigate whether VKORC1 variants could be implicated in the susceptibility of this population. We found that African and African American populations were much more likely to have at least one synonymous variant that significantly changes codon and codon pair usage in a relatively conserved position. Upon further investigation, we find that this is due to a single synonymous variant, VKORC1:c.358C>T, which is very common in African and African American populations (19.52%) while comparatively rare elsewhere (maximum 1.19% among other populations). This variant is in a relatively conserved position enriched in common codons, with negative changes in relative synonymous codon and codon

pair usage (RSCU, RSCPU). This variant was not predicted to change mRNA MFE, while it showed mixed results for splicing effects: while hexamer splicing scoring tools and ESEfinder showed changes in splicing near the variant, FAS ESS and Exonscan found no changes. Furthermore, this variant is associated with warfarin resistance [93–95], and in linkage disequilibrium with another variant upstream of the coding sequence (CDS), NG_011564.1:g.3350A>G, which is also common in African and African American populations (36.40%) and associated with warfarin resistance.

In addition, we identified one nonsynonymous variant, VKORC1:c.106G>T, which is relatively common in Ashkenazi Jewish populations (3.857%) and rare in other populations (max 0.4599% among other populations). This variant is predicted to be deleterious by both SIFT and Polyphen and associated with warfarin resistance. This variant appears at the end of a transmembrane helix near a loop, and likely impacts loop conformation near the warfarin binding site.

These two variants were interesting, primarily due to their significant population skew. There are many other variants with different prevalence in different populations, but all others are much rarer or much more common across all populations.

We additionally characterized the population prevalence of the variants identified from the GWAS studies, finding great variance in prevalence for some. For example, NM_024006.4:c.283+837T>C is very common in all populations (64.3% MAF globally), but less common in East Asian populations (10.17%).

Furthermore, some nonsynonymous variants were identified from literature to impact drug response or disease status. Associated nucleotide changes are not always given for these variants, so characterizing them has not been possible.

Of the 129 and 127 synonymous and missense VKORC1 variants, respectively, that we considered, 17 synonymous and 29 missense variants were in the VKORC1-ORF7a interface in at least one of our final models. Of these missense variants, only c.355T>A is predicted to be deleterious in both SIFT and Polyphen-2, and is likely to impact protein structure and binding to ORF7a. However, c.157C>T, c.157C>A, c.184G>A, c.229A>T, c.261C>A, c.277T>C, c.280T>C, c.280T>A, c.326G>A, c.344C>G, c.355T>A, c.378C>A, and c.379G>A are predicted to be deleterious by SIFT only, and may also affect binding.

VKORC1 paralog and variants that are impactful on warfarin dosage

VKORC1L1 is a VKORC1 paralog with similar function but reduced warfarin sensitivity [96,97]. We aligned VKORC1 with VKORC1L1 and analyzed the differences between them in the positions of variants, for additional insight into their impact on warfarin sensitivity and possible binding to ORF7a.

In the alignment of VKORC1 and VKORC1L1, seven out of twenty positions for the nonsynonymous variants impacting warfarin dosage are not conserved. This is unsurprising because the non-conserved variants are localized to the loop between transmembrane helices one and two, which is near the warfarin binding site (Fig 4). Swapping this region between VKORC1 and VKORC1L1 causes warfarin resistance in VKORC1 and warfarin sensitivity in VKORC1L1 [96].

In addition, we examined similarities of ORF7a with VKORC1 interacting proteins. Two human proteins are structurally similar to ORF7a and interact with VKORC1: CXADR, a Coxsackievirus and Adenovirus receptor [98], and PCDH1, a Hantavirus receptor [99]. Both proteins are involved in cell-cell adhesion. The structural similarity of ORF7a protein, CXADR, and PCDH1 additionally supports the interaction of ORF7a and VKORC1. The structurally aligned regions are shown in Fig 5. Structural overlap is limited to the beta sheets, with small potential for biomimicry.



Fig 5. Structural alignment of ORF7a, CXADR, and PCDH1 proteins. The alignment is largely confined to the beta sheets.

<https://doi.org/10.1371/journal.pcbi.1008805.g005>

Discussion

COVID-19 illness is characterized by a prothrombotic phenotype that is incompletely understood [1,2,3,4,6]. Developing more effective treatments will require a thorough understanding of the root causes of COVID-19 associated coagulopathy. Although several possible mechanisms have been proposed [100–102] to explain pathologic thrombosis in COVID-19, many aspects remain unexplored. Three proteins, VKORC1, SERPING1 and PABPC4, which influence coagulation have been shown to interact with SARS proteins. We investigated computationally the binding of these proteins to SARS-CoV-2 proteins. Additionally, we identified genetic variants of these proteins and examined their prevalence across populations. We

explored mechanisms by which these variants may impact COVID-19, specifically as it relates to COVID-19 associated coagulopathy. We conclude that each of these proteins may provide a potential link between COVID-19 and coagulation.

VKORC1 is crucial for maintaining active vitamin K levels and hence for the function of several essential coagulation factors. We computationally characterized the interaction between VKORC1 and SARS-CoV-2 ORF7a, guided by previous experimental data showing an interaction between VKORC1 and SARS-CoV ORF7a. Whether the ORF7a-VKORC1 interaction would be inhibitory, neutral or possibly potentiate VKORC1 function is difficult to determine in the absence of experimental testing. However, the importance of this interaction should be considered. Indeed, SARS-CoV-2 interaction with the host vitamin K cycle has again been demonstrated in more recent experimental work. Gordon and colleagues expressed SARS-CoV-2 proteins in HEK cells and found gamma-glutamyl carboxylase to interact with the SARS-CoV-2 M protein, which could similarly impact vitamin K-dependent proteins [103].

The ORF7a-VKORC1 interaction theoretically could lead to insufficient carboxylation of vitamin K dependent coagulation factors. However, COVID-19 associated coagulopathy is not typified by coagulation factor deficiencies as measured using common clinical coagulation laboratory assays [104]. Vitamin K dependent proteins outside of the coagulation cascade also contribute to platelet activation and thrombosis. An example is Gas6, which contributes to platelet activation through tyrosine kinase receptors. However, loss of Gas6 signaling is protective against thrombosis in mice [105]. Clearly therefore other mechanisms linking the VKORC1-ORF7a interaction to COVID-19 pathogenesis and coagulopathy warrant consideration.

Interestingly, reduced vitamin K levels are associated with worse prognosis in COVID-19 [106]. A deficiency of Vitamin K dependent proteins that protect against pulmonary and vascular elastic damage has been postulated to underlie this relationship. There is also a recognized inverse relationship of vitamin K and the inflammatory response [107,108] and interleukin-6 (IL-6) levels [109]. Inflammation and the immune response are thought to contribute significantly to the pathogenesis of COVID-19 illness [110]. Inflammation is also directly linked to coagulation activation [111]. Inflammation resulting from COVID-19 infection is not unique to adults, as children have also been found to suffer from Kawasaki disease-like symptoms [112].

Finally, the VKORC1—ORF7a interaction may also have an impact on tetherin function. SARS ORF7a is known to inhibit tetherin [113,15], also known as BST-2. Tetherin inhibits virion dispersal [114], and several viruses, including HIV, have auxiliary proteins to counter this effect. The structures of tetherin and VKORC1 are noticeably similar, sharing a coiled-coil architecture: VKORC1 has four consecutive alpha helices [66], while tetherin exists as a homomer of four alpha helices [16,115]. Of note, ORF7a has the highest RSCU of any SARS-CoV-2 protein [116], which may result in more efficient translation and high expression levels compared to other viral proteins, to more effectively counter the effect of tetherin.

While many extrinsic factors can influence the presentation of COVID-19, the impact of host genetic variants on viral protein interaction has not received much attention to date. Modulating the strength of these interactions or the availability of host proteins may reduce the effectiveness of viral protein function, viral replication and ultimately the severity of infection.

It is interesting to consider the impact of genetic variants in VKORC1. Individuals carrying VKORC1 variants may lead to altered protein conformation and differential binding to either warfarin or ORF7a. For instance, the synonymous variant 358C>T is characterized by a large change both in RSCPU and RSCU suggesting that it may be associated with altered

cotranslational folding. Reduced VKORC1 expression or binding to ORF7a, as it may occur in individuals with VKORC1 gene variants, may increase the availability of ORF7a to bind and inhibit tetherin increasing the severity of SARS-CoV-2 infection.

Due to the lack of structural data for some segments, homology models of PABPC4 and SERPING1 could not be constructed with high confidence, precluding the ability to create complexes to model and analyze the interactions between these proteins and viral proteins. While PABPC4 has been found to interact with SARS-CoV-2 N protein experimentally [68], interactions between SERPING1 and SARS-CoV-2 proteins have yet to be directly tested.

SERPING1 encodes C1 esterase inhibitor, a plasma protein that inhibits the C1 complex of classical pathway of complement. C1 esterase inhibitor also is the primary inhibitor of plasma kallikrein, which produces bradykinin from high-molecular-weight kininogens. Viral interactions with SERPING1 may therefore result in excessive levels of complement activation, bradykinin production and angioedema. This impact could be more pronounced in individuals with genetic variants that result in lower expression of SERPING1 or SERPING1 activity. The relationship between the ACE2 (primary receptor for SARS-CoV-2) and the kinin system could exacerbate this impact. ACE2 inactivates des-Arg⁹ bradykinin (DABK) [117,118], an active bradykinin metabolite. Reduced ACE2 activity is associated with enhanced signaling of DABK, angioedema, and neutrophil infiltration in the lungs [119,117]. The combined effect of viral suppression of ACE2 expression and function [120] concurrent with SERPING1 inhibition may result in excessively high levels of bradykinin and pulmonary fluid accumulation.

Altogether, the interaction of VKORC1, SERPING1 and PABPC4 with viral proteins may result in dysregulated coagulation and immune response. Genetic variants in these genes may impact the host-viral protein interaction by altering protein conformation or expression. Because these genetic variants appear at different frequencies in different populations, this may contribute to differential outcomes for COVID-19 patients from various ethnic groups. Indeed, COVID-19 has had an unequal impact on populations across the globe [121,122]. In the United States, as elsewhere, it is clear that demographic subgroups are more susceptible to severe COVID-19 disease. Certainly a large number of non-genetic factors influence clinical outcomes within populations, including age, access to health care, and presence of comorbidities [123,124]. The genetic underpinnings of host-viral protein interaction may also play an underappreciated role in determining the course of COVID-19 illness and COVID-19 associated coagulopathy.

Supporting information

S1 Table. Template crystal structures used for all protein models, from I-TASSER.
(XLSX)

S2 Table. Computed features for all genetic variants of interest in VKORC1, SERPING1, and PABPC4 from GWAS studies.
(XLSX)

S3 Table. Predicted binding score for all relevant miRNA species and the variant (MUT) and wild type (WT) sequences.
(XLSX)

S4 Table. Computed features for all identified synonymous variants identified in proteins that interact with SARS proteins.
(XLSX)

S5 Table. Computed features for all identified missense variants identified in proteins that interact with SARS proteins.

(XLSX)

S6 Table. Description, explanation and range of computed features.

(XLSX)

Author Contributions

Conceptualization: Jacob Kames, Nobuko Hamasaki-Katagiri, Michael DiCuccio, Chava Kimchi-Sarfaty.

Formal analysis: David Holcomb.

Funding acquisition: Anton A. Komar, Chava Kimchi-Sarfaty.

Investigation: David Holcomb, Nancy Hernandez, Ryan Hunt, Kyle Laurie.

Methodology: Chava Kimchi-Sarfaty.

Project administration: Chava Kimchi-Sarfaty.

Software: David Holcomb.

Visualization: Nancy Hernandez.

Writing – original draft: David Holcomb, Aikaterini Alexaki.

Writing – review & editing: Aikaterini Alexaki, Ryan Hunt, Jacob Kames, Nobuko Hamasaki-Katagiri, Anton A. Komar, Michael DiCuccio, Chava Kimchi-Sarfaty.

References

1. Fogarty H, Townsend L, Cheallaigh CN, Bergin C, Martin-Loeches I, Browne P, et al. COVID-19 Coagulopathy in Caucasian patients. *British Journal of Haematology*. 2020; 189(6): 1044–1049.
2. McGonagle D, O'Donnell JS, Sharif K, Emery P, Bridgewood C. Immune mechanisms of pulmonary intravascular coagulopathy in COVID-19 pneumonia. *The Lancet Rheumatology*. 2020; 2(7): E437–E445. [https://doi.org/10.1016/S2665-9913\(20\)30121-1](https://doi.org/10.1016/S2665-9913(20)30121-1) PMID: 32835247
3. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *New England Journal of Medicine*. 2020; 383: 120–128.
4. Nahum J, Morichau-Beauchant T, Daviaud F, et al. Venous Thrombosis Among Critically Ill Patients With Coronavirus Disease 2019 (COVID-19). *Critical Care Medicine*. 2020; 3(5). <https://doi.org/10.1001/jamanetworkopen.2020.10478> PMID: 32469410
5. Katneni UK, Alexaki A, Hunt RC, Schiller T, DiCuccio M, Buehler PW, et al. Coagulopathy and Thrombosis as a Result of Severe COVID-19 Infection: A Microvascular Focus. 2020.
6. Connors J, Levy J. COVID-19 and its implications for thrombosis and anticoagulation. *Blood*. 2020; 135(23): 2033–2040. <https://doi.org/10.1182/blood.2020006000> PMID: 32339221
7. Cui S, Chen S, Li X, Liu S, Wang F. Prevalence of venous thromboembolism in patients with severe novel coronavirus pneumonia. *Journal of Thrombosis and Haemostasis*. 2020; 18(6): 1421–1424. <https://doi.org/10.1111/jth.14830> PMID: 32271988
8. Ng JJ, Choong AMTL. Thromboembolic Events in Patients with SARS-CoV-2. *Journal of Vascular Surgery*. 2020; 72(2): 760–741. <https://doi.org/10.1016/j.jvs.2020.04.488> PMID: 32360371
9. Tavazzi G, Cividari L, Caneva L, Mondogi S, Mojoli F. Thrombotic events in SARS-CoV-2 patients: an urgent call for ultrasound screening. *Intensive Care Medicine*. 2020; 46: 1121–1123. <https://doi.org/10.1007/s00134-020-06040-3> PMID: 32322918
10. Pfefferle S, Schopf J, Kogl M, Friedel CC, Muller MA, Carbajo-Lozoya J, et al. The SARS-coronavirus-host interactome: identification of cyclophilins as target for pan-coronavirus inhibitors. *PLoS Pathogens*. 2011; 7(10): e1002331. <https://doi.org/10.1371/journal.ppat.1002331> PMID: 22046132

11. Feral C, Mattei MG, Pawlak A, Guellaen G. Chromosomal Localization of Three Human poly(A)-binding Protein Genes and Four Related Pseudogenes. *Human Genetics*. 1999; 105(4): 347–353. <https://doi.org/10.1007/s004399900148> PMID: 10543404
12. Houg AK, Maggini L, Clement CY, Reed GL. Identification and Structure of Activated-Platelet protein-1, a Protein With RNA-binding Domain Motifs That Is Expressed by Activated Platelets. *European Journal of Biochemistry*. 2004; 243(1–2): 209–218.
13. PABPC4 poly(A) binding protein cytoplasmic 4 [Homo sapiens (human)]. [Online].; 2020 [cited 2020 07 22]. Available from: <https://www.ncbi.nlm.nih.gov/gene/8761>.
14. Marcelino-Rodriguez I, Callero A, Mendoza-Alvarez A, Perez-Rodriguez E, Barrios-Recio J, Garcia-Robaina JC, et al. Bradykinin-Mediated Angioedema: An Update of the Genetic Causes and the Impact of Genomics. *Frontiers in Genetics*. 2019. <https://doi.org/10.3389/fgene.2019.00900> PMID: 31611908
15. Wang SM, Huang KJ, Wang CT. Severe acute respiratory syndrome coronavirus spike protein counteracts BST2-mediated restriction of virus-like particle release. *Journal of Medical Virology*. 2019; 91(10): 1741–1750. <https://doi.org/10.1002/jmv.25518> PMID: 31199522
16. Schubert HL, Zhai QZ, Sandrin V, Eckert DM, Garcia-Maya M, Saul L, et al. Structural and functional studies on the extracellular domain of BST2/tetherin in reduced and oxidized conformations. *Proceedings of the National Academy of Sciences*. 2010; 107(42): 17951–17956. <https://doi.org/10.1073/pnas.1008206107> PMID: 20880831
17. Yount B, Roberts RS, Sims AC, Deming D, Frieman MB, Sparks J, et al. Severe Acute Respiratory Syndrome Coronavirus Group-Specific Open Reading Frames Encode Nonessential Functions for Replication in Cell Cultures and Mice. *Journal of Virology*. 2005 Nov; 79(23): 14909–14922. <https://doi.org/10.1128/JVI.79.23.14909-14922.2005> PMID: 16282490
18. Hanel K, Willbold D. SARS-CoV accessory protein 7a directly interacts with human LFA-1. *Biological Chemistry*. 2007 Dec; 388(12): 1325–1332. <https://doi.org/10.1515/BC.2007.157> PMID: 18020948
19. Garcia AA, Reitsma PH. VKORC1 and the Vitamin K Cycle. *Vitamins and Hormones*. 2007; 78: 23–33.
20. Jorg S. Principles of Nutrigenetics and Nutrigenomics: Fundamentals of Individualized Nutrition Caterina RD, Martinez JA, Kohlmeier M, editors.: Academic Press; 2020.
21. Testa S, Paoletti O, Giorgi-Pierfranceschi M, Pan A. Switch from oral anticoagulants to parenteral heparin in SARS-CoV-2 hospitalized patients. *Internal and Emergency Medicine*. 2020. <https://doi.org/10.1007/s11739-020-02331-1> PMID: 32297089
22. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein. *Nature Protocols*. 2010; 5(4): 725–738. <https://doi.org/10.1038/nprot.2010.5> PMID: 20360767
23. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research*. 2015; 43(Web Server issue): W174–W181. <https://doi.org/10.1093/nar/gkv342> PMID: 25883148
24. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nature Methods*. 2015; 12: 7–8. <https://doi.org/10.1038/nmeth.3213> PMID: 25549265
25. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK Server: Interactive Docking Prediction of Protein-Protein Complexes and Symmetric Multimers. *Bioinformatics*. 2014; 30(12): 1771–1773. <https://doi.org/10.1093/bioinformatics/btu097> PMID: 24532726
26. Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *Journal of Molecular Biology*. 2003; 331(1): 281–299. [https://doi.org/10.1016/s0022-2836\(03\)00670-3](https://doi.org/10.1016/s0022-2836(03)00670-3) PMID: 12875852
27. Lyskov S, Gray JJ. The RosettaDock server for local protein–protein docking. *Nucleic Acids Research*. 2008;(36 (Web Server issue)): W233–W238. <https://doi.org/10.1093/nar/gkn216> PMID: 18442991
28. Wang C, Schueler-Furman O, Baker D. Improved side-chain modeling for protein-protein docking. *Protein Science*. 2005; 14(5): 1328–1329. <https://doi.org/10.1110/ps.041222905> PMID: 15802647
29. Chaudhury S, Berrondo M, Weitzner BD, Muthu P, Bergman H, Gray JJ. Benchmarking and analysis of protein docking performance in RosettaDock v3.2. *PLoS One*. 2011; 6(8). <https://doi.org/10.1371/journal.pone.0022477> PMID: 21829626
30. Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, et al. Genomic Diversity of Severe Acute Respiratory Syndrome–Coronavirus 2 in Patients With Coronavirus Disease 2019. *Clinical Infectious Diseases*. 2020 August; 71(15): 713–720. <https://doi.org/10.1093/cid/ciaa203> PMID: 32129843
31. Rouchka EC, Chariker JH, Chung D. Variant analysis of 1,040 SARS-CoV-2 genomes. *PLOS ONE*. 2020 November. <https://doi.org/10.1371/journal.pone.0241535> PMID: 33152019

32. Raman S, Vernon R, Thompson J, Tyka M, Sadreyev R, Pei J, et al. Structure Prediction for CASP8 With All-Atom Refinement Using Rosetta. *Proteins*. 2009; 77 Suppl 9(0 9): 89–99. <https://doi.org/10.1002/prot.22540> PMID: 19701941
33. Song Y, DiMaio F, Wang RYR, Kim D, Miles C, Brunette TJ, et al. High resolution comparative modeling with RosettaCM. *Structure*. 2013; 21(10): 1735–1742. <https://doi.org/10.1016/j.str.2013.08.005> PMID: 24035711
34. Holm L. Benchmarking fold detection by DaliLite v.5. *Bioinformatics*. 2019; 35(24): 5326–5327. <https://doi.org/10.1093/bioinformatics/btz536> PMID: 31263867
35. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*. 2006; 34(Suppl_1): D535–D539. <https://doi.org/10.1093/nar/gkj109> PMID: 16381927
36. Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Research*. 2019; 47(D1): D529–D541. <https://doi.org/10.1093/nar/gky1079> PMID: 30476227
37. The COVID-19 Host Genetics Initiative. [Online].; 2020 [cited 2020 06 25]. Available from: <http://covid19hg.org>.
38. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *New England Journal of Medicine*. 2020. <https://doi.org/10.1056/NEJMoa2020283> PMID: 32558485
39. Ke S, Shang S, Kalachikov SM, Morozova I, Yu L, Russo JJ, et al. Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Research*. 2011; 21(8): 1360–1374. <https://doi.org/10.1101/gr.119628.110> PMID: 21659425
40. Erkelenz S, Theiss S, Otte M, Widera M, Peter JO, Schaal H. Genomic HEXploring allows landscaping of novel potential splicing regulatory elements. *Nucleic Acids Research*. 2014; 42(16): 10681–10697. <https://doi.org/10.1093/nar/gku736> PMID: 25147205
41. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Research*. 2003; 31(13): 3568–3571. <https://doi.org/10.1093/nar/gkg616> PMID: 12824367
42. Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Human Molecular Genetic*. 2006; 15(16): 2490–2508. <https://doi.org/10.1093/hmg/ddl171> PMID: 16825284
43. Wang Z, Rolish ME, Gene Yeo, Tung V, Mawson M, Burge CB. Systematic identification and analysis of exonic splicing silencers. *Cell*. 2004; 119(6): 831–845. <https://doi.org/10.1016/j.cell.2004.11.010> PMID: 15607979
44. Yeo G, Burge CB. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *Journal of Computational Biology*. 2004; 11(2–3): 377–394. <https://doi.org/10.1089/1066527041410418> PMID: 15285897
45. Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science*. 2002; 297(5583): 1007–1013. <https://doi.org/10.1126/science.1073774> PMID: 12114529
46. Xayaphoummine A, Bucher T, Isambert H. Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research*. 2005; 33(Supplemental 2): W605–W610. <https://doi.org/10.1093/nar/gki447> PMID: 15980546
47. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*. 2003; 31(13): 3406–3415. <https://doi.org/10.1093/nar/gkg595> PMID: 12824337
48. Waugh A, Gendron P, Altman R, Brown JW, Case D, Gautheret D, et al. RNAML: A standard syntax for exchanging RNA information. *RNA*. 2002; 8(6): 707–717. <https://doi.org/10.1017/s1355838202028017> PMID: 12088144
49. Zuker M, Jacobson AB. Using reliability information to annotate RNA secondary structures. *RNA*. 1998; 4(6): 669–679. <https://doi.org/10.1017/s1355838298980116> PMID: 9622126
50. Salari R, Kimchi-Sarfaty C, Gottesman MM, Przytycka TM. Sensitive measurement of single-nucleotide polymorphism-induced changes of RNA conformation: application to disease studies. *Nucleic Acids Research*. 2012; 41(1): 44–53. <https://doi.org/10.1093/nar/gks1009> PMID: 23125360
51. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Research*. 2020; 48(D1): D127–D131. <https://doi.org/10.1093/nar/gkz757> PMID: 31504780
52. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biology*. 2019; 20(1): 18. <https://doi.org/10.1186/s13059-019-1629-z> PMID: 30670076

53. Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery*. 2020; 6(14). <https://doi.org/10.1038/s41421-020-0153-3> PMID: 32194980
54. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*. 2001; 29(1): 308–311. <https://doi.org/10.1093/nar/29.1.308> PMID: 11125122
55. Karczewski KJ, al. e. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020; 581: 434–443. <https://doi.org/10.1038/s41586-020-2308-7> PMID: 32461654
56. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biology*. 2016; 17: 122. <https://doi.org/10.1186/s13059-016-0974-4> PMID: 27268795
57. Pei J, Grishin NV. AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics*. 2001; 17(8): 700–712. <https://doi.org/10.1093/bioinformatics/17.8.700> PMID: 11524371
58. Alexaki A, Kames J, Holcomb DD, Athey J, Santana-Quintero LV, Lam PVN, et al. Codon and Codon-Pair Usage Tables (CoCoPUTs): Facilitating Genetic Variation Analyses and Recombinant Gene Design. *Journal of Molecular Biology*. 2019; 431(13): 2434–2441. <https://doi.org/10.1016/j.jmb.2019.04.021> PMID: 31029701
59. Sharp PM, Li WH. The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research*. 1987; 15(3): 1281–1295. <https://doi.org/10.1093/nar/15.3.1281> PMID: 3547335
60. Jacobs WJ, Shakhnovich EI. Evidence of evolutionary selection for cotranslational folding. *Proceedings of the National Academy of Sciences*. 2017; 114(43): 11434–11439. <https://doi.org/10.1073/pnas.1705772114> PMID: 29073068
61. Rodriguez A, Wright G, Emrich S, Clark PL. %MinMax: A versatile tool for calculating and comparing synonymous codon usage and its impact on protein folding. *Protein Science*. 2018; 27(1): 356–362. <https://doi.org/10.1002/pro.3336> PMID: 29090506
62. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*. 2003; 31(13): 3812–3814. <https://doi.org/10.1093/nar/gkg509> PMID: 12824425
63. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nature Methods*. 2010; 7: 248–249. <https://doi.org/10.1038/nmeth0410-248> PMID: 20354512
64. Jacobson GN, Clark PL. Quality over quantity: optimizing co-translational protein folding with non-‘optimal’ synonymous codons. *Current Opinion in Structural Biology*. 2016; 38: 102–110. <https://doi.org/10.1016/j.sbi.2016.06.002> PMID: 27318814
65. Zhang G, Hubalewska M, Ignatova Z. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Natural Structural and Molecular Biology*. 2009; 16: 274–280. <https://doi.org/10.1038/nsmb.1554> PMID: 19198590
66. Chatron N, Khalil RA, Benoit E, Lattard V. Structural Investigation of the Vitamin K Epoxide Reductase (VKORC1) Binding Site with Vitamin K. *Biochemistry*. 2020; 59(13): 1351–1360. <https://doi.org/10.1021/acs.biochem.9b01084> PMID: 32182040
67. Hopf TA, Ingraham JB, Poelwijk FJ, Scharfe CPI, Springer M, Sander C, et al. Mutation effects predicted from sequence co-variation. *Nature Biotechnology*. 2017 January; 35: 128–135. <https://doi.org/10.1038/nbt.3769> PMID: 28092658
68. Gordon DE, et. a. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020; 583: 459–468. <https://doi.org/10.1038/s41586-020-2286-9> PMID: 32353859
69. Wang D, Chen H, Momary KM, Cavallari LH, Johnson JA, Wang Sadee W. et al. Regulatory polymorphism in vitamin K epoxide reductase complex subunit 1 (VKORC1) affects gene expression and warfarin dose requirement. *Blood*. 2008 August; 112(4): 1013–1021. <https://doi.org/10.1182/blood-2008-03-144899> PMID: 18523153
70. Kumar DK, Shewade DG, Parasuraman S, Rajan S, Balachander J, Chandran BVS, et al. Estimation of plasma levels of warfarin and 7-hydroxy warfarin by high performance liquid chromatography in patients receiving warfarin therapy. *Journal of Young Pharmacists*. 2013; 5(1): 13–17. <https://doi.org/10.1016/j.jyp.2013.02.001> PMID: 24023446
71. Li T, Lange LA, Susswein L, Bryant B, Malone R, Lange EM, et al. Polymorphisms in the VKORC1 gene are strongly associated with warfarin dosage requirements in patients receiving anticoagulation. *Journal of Medical Genetics*.; 43(9): 740–744. <https://doi.org/10.1136/jmg.2005.040410> PMID: 16611750
72. Kocael A, Eronat AP, Tuzuner MB, Ekmekci A, Orhan AL, Ikizceli I, et al. Interpretation of the effect of CYP2C9, VKORC1 and CYP4F2 variants on warfarin dosing adjustment in Turkey. *Molecular Biology Reports*. 2019; 46: 1825–1833. <https://doi.org/10.1007/s11033-019-04634-9> PMID: 30712247

73. Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MTM, et al. Warfarin pharmacogenetics: a single VKORC1 polymorphism is predictive of dose across 3 racial groups. *Blood*.; 115(18): 3827–3834. <https://doi.org/10.1182/blood-2009-12-255992> PMID: 20203262
74. Yuan HY, Chen JJ, Lee MTM, Wung JC, Chen YF, Charnig MJ, et al. A novel functional VKORC1 promoter polymorphism is associated with inter-individual and inter-ethnic differences in warfarin sensitivity. *Human Molecular Genetics*.; 14(13): 1745–1751. <https://doi.org/10.1093/hmg/ddi180> PMID: 15888487
75. Kumar DK, Shewade DG, Lorient MA, Beaune P, Balachander J, Chandran BVS, et al. Effect of CYP2C9, VKORC1, CYP4F2 and GGCX genetic variants on warfarin maintenance dose and explicating a new pharmacogenetic algorithm in South Indian population. *European Journal of Clinical Pharmacology*. 2014; 70(1): 47–56. <https://doi.org/10.1007/s00228-013-1581-x> PMID: 24019055
76. Ennis S, Jomary C, Mullins R. Association between the SERPING1 and age-related macular degeneration: a two-stage case–control study. *The Lancet*. 2008; 372(9652): 1828–1834. [https://doi.org/10.1016/S0140-6736\(08\)61348-3](https://doi.org/10.1016/S0140-6736(08)61348-3) PMID: 18842294
77. Liu K, Lai TYY, Ma L, Lai FHP, Young AL, Brelén ME, et al. Ethnic differences in the association of SERPING1 with age-related macular degeneration and polypoidal choroidal vasculopathy. *Scientific Reports*. 2015; 5. <https://doi.org/10.1038/srep09424> PMID: 25800435
78. Vatsiou S, Zamanakou M, Loules G, Psarros F, Parsopoulou F, Csuka D, et al. A novel deep intronic SERPING1 variant as a cause of hereditary angioedema due to C1-inhibitor deficiency. *Allergy International*. 2020; 69(3): 443–449. <https://doi.org/10.1016/j.alit.2019.12.009> PMID: 31959500
79. Kralovicova J, Vorechovsky I. SERPING1 rs2511988 and age-related macular degeneration. *The Lancet*. 2009; 373(9662): 461–462. [https://doi.org/10.1016/S0140-6736\(09\)60168-9](https://doi.org/10.1016/S0140-6736(09)60168-9) PMID: 19200915
80. Lu F, Zhao P, Fan Y, Tang S, Hu J, Liu X, et al. An association study of SERPING1 gene and age-related macular degeneration in a Han Chinese population. *Molecular Vision*. 2010; 16(1): 1–6.
81. Li M, Wen F, Zuo C, Zhang X, Chen H, Huang S, et al. SERPING1 polymorphisms in polypoidal choroidal vasculopathy. *Molecular Vision*. 2010; 16: 231–239. PMID: 20161815
82. Genome-wide association study of childhood acute lymphoblastic leukemia in Korea. *Leukemia Research*. 2010; 34(10): 1271–1274. <https://doi.org/10.1016/j.leukres.2010.02.001> PMID: 20189245
83. Park KH, Ryu E, Tosakulwong N, Wu Y, Edwards AO. Common variation in the SERPING1 gene is not associated with age-related macular degeneration in two independent groups of subjects. *Molecular Vision*. 2009; 15(1): 200–207.
84. Yang Mm, Lai TYY, Tam POS, Chiang SWY, Ng TK, Rong SS, et al. Association of CFH and SERPING1 polymorphisms with anterior uveitis. *British Journal of Ophthalmology*. 2013; 97(11): 1475–1480. <https://doi.org/10.1136/bjophthalmol-2013-303679> PMID: 23966370
85. Broadgate S, Kiire C, Halford S, Chong V. Diabetic macular oedema: under-represented in the genetic analysis of diabetic retinopathy. 2018; 96(A111): 1–51.
86. Lee AY, Kulkarni M, Fang AM, Edelstein S, Osborn MP, Brantley MA. The effect of genetic variants in SERPING1 on the risk of neovascular age-related macular degeneration. *British Journal of Ophthalmology*. 2010; 94(7): 915–917. <https://doi.org/10.1136/bjo.2009.172007> PMID: 20606025
87. Lores-Motta L, Riaz M, Grunin M, Corominas J, van Asten F, Pauper M, et al. Association of Genetic Variants With Response to Anti-Vascular Endothelial Growth Factor Therapy in Age-Related Macular Degeneration. *Journal of the American Medical Association Ophthalmology*. 2018; 136(8): 875–884. <https://doi.org/10.1001/jamaophthalmol.2018.2019> PMID: 29852030
88. Han S, Lee KM, Park SK, Lee JE, Ahn HS, Shin HY, et al. Genome-wide association study of childhood acute lymphoblastic leukemia in Korea. *Leukemia Research*. 2010; 34(10): 1274–1274. <https://doi.org/10.1016/j.leukres.2010.02.001> PMID: 20189245
89. Passey C, Birnbaum AK, Brundage RC, Oetting WS, Israni AK, Jacobson PA. Dosing equation for tacrolimus using genetic variants and clinical factors. *British Journal of Clinical Pharmacology*. 2011; 72(6): 948–957. <https://doi.org/10.1111/j.1365-2125.2011.04039.x> PMID: 21671989
90. Cook JP, Morris AP. Multi-ethnic genome-wide association study identifies novel locus for type 2 diabetes susceptibility. 2016; 24(8): 1175–1180. <https://doi.org/10.1038/ejhg.2016.17> PMID: 27189021
91. Czogalla KJ, Watzka M, Oldenburg J. Structural Modeling Insights into Human VKORC1 Phenotypes. *Nutrients*. 2015; 7(8): 6837–6851. <https://doi.org/10.3390/nu7085313> PMID: 26287237
92. Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, et al. Effect of VKORC1 Haplotypes on Transcriptional Regulation and Warfarin Dose. *New England Journal of Medicine*. 2005; 352(22): 2285–2293. <https://doi.org/10.1056/NEJMoa044503> PMID: 15930419
93. Mitchell C, Gregersen N, Krause A. Novel CYP2C9 and VKORC1 gene variants associated with warfarin dosage variability in the South African black population. *Future Medicine*. 2011; 12(7): 953–963. <https://doi.org/10.2217/pgs.11.36> PMID: 21635147

94. Wu W, Harenerg J. Influence of genetic Polymorphisms of Vitamin K Epoxide Reductase Complex Subunit 1 and Cytochrome P450, Family 2, Subfamily C, Polypeptide 9 on the Dose of Phenprocoumon. Heidelberg University, Institute of Experimental and Clinical Pharmacology and Toxicology; 2009.
95. Anton AI, Cerezo-Manchado JJ, Padilla J, Perez-Andreu V, Corral J, Vicente V. Novel Associations of VKORC1 Variants with Higher Acenocoumarol Requirements. PLOS ONE. 2013; 8(5): e64469. <https://doi.org/10.1371/journal.pone.0064469> PMID: 23691226
96. Lacombe J, Ferron M. VKORC1L1, An Enzyme Mediating the Effect of Vitamin K in Liver and Extrahepatic Tissues. Nutrients. 2018; 10(8): 970. <https://doi.org/10.3390/nu10080970> PMID: 30050002
97. Czogalla KJ, Liphardt K, Honing K, Homung V, Biswas A, Watzka M, et al. VKORC1 and VKORC1L1 have distinctly different oral anticoagulant dose-response characteristics and binding sites. Blood Advances. 2018; 2(6): 691–702. <https://doi.org/10.1182/bloodadvances.2017006775> PMID: 29581108
98. Bergelson JM, Cunningham JA, Droguett G, Kurt-Jones EA, Krithivas A, Hong JS, et al. Isolation of a Common Receptor for Coxsackie B Viruses and Adenoviruses 2 and 5. Science. 1997; 275(5304): 1320–1323. <https://doi.org/10.1126/science.275.5304.1320> PMID: 9036860
99. Jangra RK, Herbert AS, Li R, Jae LT, Kleinfelter LM, Slough MM, et al. Protocadherin-1 is essential for cell entry by New World hantaviruses. Nature. 2018; 563: 559–563. <https://doi.org/10.1038/s41586-018-0702-1> PMID: 30464266
100. Lumbers ER, Delforce SJ, Pringle KG, Smith GR. The Lung, the Heart, the Novel Coronavirus, and the Renin-Angiotensin System; The Need for Clinical Trials. Frontiers in Medicine. 2020; 22(4): 248.
101. Chauhan AJ, Wiffen LJ, Brown TP. COVID-19: a collision of complement, coagulation and inflammatory pathways. Journal of Thrombosis and Haemostasis. 2020. <https://doi.org/10.1111/jth.14981> PMID: 32608159
102. Colling ME, Kanthi Y. COVID-19-associated coagulopathy: An exploration of mechanisms. Vascular Medicine. 2020.
103. Gordon DE, Jang GM, Bouhaddou M, Xy J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020 April; 583: 459–468. <https://doi.org/10.1038/s41586-020-2286-9> PMID: 32353859
104. Al-Samkari H, Leaf RSK, Dzik WH, Carlson JCT, Fogerty AE, Waheed A, et al. COVID-19 and coagulation: bleeding and thrombotic manifestations of SARS-CoV-2 infection. Blood. 2020 July; 136(4): 489–500. <https://doi.org/10.1182/blood.2020006520> PMID: 32492712
105. Angelillo-Scherrer A, Burnier L, Flores N, Savi P, DeMol M, Schaeffer P, et al. Role of Gas6 receptors in platelet signaling during thrombus stabilization and implications for antithrombotic therapy. Journal of Clinical Investigation. 2005 February; 115(2): 237–246. <https://doi.org/10.1172/JCI22079> PMID: 15650770
106. Dofferhoff ASM, Piscoer I, Schurgers LJ, Walk J, van den Ouweland JMW, Hackeng TM, et al. Reduced Vitamin K Status as a Potentially Modifiable Prognostic Risk Factor in COVID-19. Preprints. 2020.
107. Harshmann SG, Shea MK. The Role of Vitamin K in Chronic Aging Diseases: Inflammation, Cardiovascular Disease, and Osteoarthritis. Current Nutrition Reports. 2016; 5(2): 90–98. <https://doi.org/10.1007/s13668-016-0162-x> PMID: 27648390
108. Shea MK, Booth SL, Massaro JM, Jacques PF, D'Agostino RB, Dawson-Highes B, et al. Vitamin K and Vitamin D Status: Associations with Inflammatory Markers in the Framingham Offspring Study. American Journal of Epidemiology. 2008; 167(3): 313–320. <https://doi.org/10.1093/aje/kwm306> PMID: 18006902
109. Reddi K, Henderson B, Meghji S, Wilson M, Poole S, Hopper C, et al. Interleukin 6 Production by Lipopolysaccharide-Stimulated Human Fibroblasts is Potently Inhibited by Naphtoquinone (Vitamin K) Compounds. Cytokine. 1995; 7(3): 287–297. <https://doi.org/10.1006/cyto.1995.0034> PMID: 7640347
110. Moore JB, June CH. Cytokine release syndrome in severe COVID-19. Science. 2020; 368(6490): 473–474. <https://doi.org/10.1126/science.abb8925> PMID: 32303591
111. Levi M, van der Poll T. Inflammation and coagulation. Critical Care Medicine. 2020; 38: S26–S34. <https://doi.org/10.1016/j.ejim.2020.05.018> PMID: 32425364
112. Viner RM, Whittaker E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic. The Lancet. 2020; 395(10239): 1741–1743. [https://doi.org/10.1016/S0140-6736\(20\)31129-6](https://doi.org/10.1016/S0140-6736(20)31129-6) PMID: 32410759
113. Taylor JK, Coleman CM, Postel S, Sisk JM, Bernbaum JG, Venkataraman T, et al. Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering

- p>through a Novel Mechanism of Glycosylation Interference.
- Journal of Virology*
- . 2015; 89(23): 11820–11833.
- <https://doi.org/10.1128/JVI.02274-15>
- PMID: 26378163
114. Evans DT, Serra-Moreno R, Singh RK, Guatelli JC. BST-2/tetherin: a new component of the innate immune response to enveloped viruses. *Trends in Microbiology*. 2011; 18(9): 388–396.
115. Yang H, Wang J, Jia X, McNatt MW, Zang T, Pan B, et al. Structural insight into the mechanisms of enveloped virus tethering by tetherin. *Proceedings of the National Academy of the Sciences*. 2020; 107(43): 18428–18432. <https://doi.org/10.1073/pnas.1011485107> PMID: 20940320
116. Kames J, Holcomb DD, Kimchi O, DiCuccio M, Hamasaki-Katagiri N, Wang T, et al. Sequence analysis of SARS-CoV-2 genome reveals features important for vaccine design. Preprint. 2020.
117. Tolouian R, Vahed SZ, Ghiyasvand S, Tolouian A, Ardalan M. COVID-19 interactions with angiotensin-converting. *Journal of Renal Injury Prevention*. 2020; 9(2): e19.
118. van de Veerdonk FL, Mihai NG, van Dueren M, van der Meer JW, de Mast Q, Bruggemann RJ, et al. Kallikrein-kinin blockade in patients with COVID-19 to prevent acute respiratory distress syndrome. *eLife*. 2020; 9: e57555. <https://doi.org/10.7554/eLife.57555> PMID: 32338605
119. Sodhi CP, Wohlford-Lenane C, Yamagushi Y, Prindle T, Fulton WB, Wang S, et al. Attenuation of pulmonary ACE2 activity impairs inactivation of des-Arg9 bradykinin/BKB1R axis and facilitates LPS-induced neutrophil infiltration. *Lung Cellular and Molecular Physiology*. 2018; 314(1): L17–L31. <https://doi.org/10.1152/ajplung.00498.2016> PMID: 28935640
120. Gutierrez-Chamorro L, Riveira-Munoz E, Barrios C, Palau V, Massanella M, Garcia-Vidal E, et al. SARS-CoV-2 infection suppresses ACE2 function and antiviral immune response in the upper respiratory tract of infected patients. *bioRxiv*. 2020 November.
121. Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of the Sciences*. 2020; 117(18): 9696–9699. <https://doi.org/10.1073/pnas.2004911117> PMID: 32300018
122. United Kingdom Office for National Statistics. [Online].; 20 [cited 20 07 20. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/coronavirusrelateddeathsbyethnicgroupenglandandwales/2march2020to10april2020>.
123. Millett GA, Jones AT, Benkeser D, Baral S, Mercer L, Beyrer C, et al. Assessing differential impacts of COVID-19 on black communities. *Annals of Epidemiology*. 2020; 47: 37–44. <https://doi.org/10.1016/j.annepidem.2020.05.003> PMID: 32419766
124. Stokes EK, Zambrano LD, Anderson KN, Marder EP, Raz KM, Felix SEB, et al. Coronavirus Disease 2019 Case Surveillance—United States, January 22–May 30, 2020. *MMWR Morbidity and Mortality Weekly Report*. 2020; 69(24): 759–765. <https://doi.org/10.15585/mmwr.mm6924e2> PMID: 32555134