Functional and Structural Characterization of COVID-19 Risk-Associated Exonic SNPs: An *In Silico* Analysis

MARCOS JESSÉ ABRAHÃO SILVA[1,*], SEBASTIÃO KAUÃ DE SOUSA BISPO[2], REBECCA LOBATO MARINHO[1], ELIETE COSTA DA CRUZ[3], THIAGO PINTO BRASIL [4], CAROLINY SOARES SILVA[1], CRISTIANE CUNHA FROTA [3], DIANA DA COSTA LOBATO[1], LILIAN CRISTINA SANTOS SINFRÔNIO DA SILVA[2], EVERALDINA CORDEIRO DOS SANTOS[2], KARLA VALÉRIA BATISTA LIMA[2], LUANA NEPOMUCENO GONDIM COSTA LIMA[2]

1 POSTGRADUATE PROGRAM IN PARASITIC BIOLOGY IN THE AMAZON (PPGBPA), UNIVERSITY OF PARÁ STATE (UEPA), BELÉM 66087-670, BRAZIL.
2 BACTERIOLOGY AND MYCOLOGY SECTION (SABMI), EVANDRO CHAGAS INSTITUTE (IEC), ANANINDEUA 67030-000, BRAZIL.
3 DEPARTMENT OF BIOMEDICINE, FEDERAL UNIVERSITY OF CEARÁ (UFC), FORTALEZA 60441-750, BRAZIL.
4 DEPARTMENT OF PATHOLOGY AND LEGAL MEDICINE, FACULTY OF MEDICINE, FEDERAL UNIVERSITY OF CEARÁ (UFC), FORTALEZA 60441-750, BRAZIL.
CORRESPONDING AUTHOR: JESSEABRAHAO10@GMAIL.COM

INTRODUCTION

The individual host susceptibility to coronavirus disease 2019 (COVID-19) can be attributed in part to single nucleotide polymorphisms (SNPs), which may be in exonic sites of the genome. The objective of this work was to analyze, *in silico*, the functional and structural impact of exonic SNPs related in the literature to susceptibility to COVID-19.

METHODS

Literature data were retrieved from PubMed and Science Direct in relation to COVID-19-risk associated SNPs and a separate analysis was performed between synonyms (sSNP) and non-synonymous (nsSNP). To characterize the sSNPs, the following predictions were made: effects on mRNA structure (with RNAfold; CycleFold; Kinefold); splicing effects on mRNA (MaxEnt Scan; Ex Skip); effects on miRNA binding (TargetScan Score). Regarding nsSNPs, the following were performed: functional analysis of protein damage (with SIFT, PolyPhen 2, PhD-SNP, SNPs & GO, Predict SNP 2). After passing pathogenicity criteria, 8 nsSNPs were selected to predict their impacts on stability (CUPSAT), functionality and residual evolution (MutPred, ConSurf).

RESULTS

The functional and structural effects of sSNPs associated with the higher risk of COVID-19 are described in Table 1.

Gene	SNP ID	RNAfold (Energy) in ΔΔG (Kcal/mol)	Kinefold (Pseudoknots)	Cyclefold (Noncanonical base pairings)	MaxEntScan in ΔΔG (Kcal/mol)	EX-SKIP (ESS/ESE/ESS:ESSE ratio)	TargetScan (Score)	Global MAF *
IFIT M3	rs12252	-58.30	-2.6	15.2	-4.33 (equal)	Without mut: 136, 137, 0.99 / with mut: 137, 140, 0.98	-0.37	G= 0.236
TMPRSS2	rs2298659	4.4	-1.9	13.8	-8.42 (equal)	Without mut: 51, 180, 0.28 / with mut: 51, 183, 0.28	-0.39	A= 0.209
IL1RN	rs419598	-1.5	0	16.6	-11.84 (equal)	Without mut: 99, 186, 0.53 / with mut: 96, 182, 0.53	-0.18	C= 0.191
VDR	rs7312336	0	0	15.6	-32.88 (equal)	Without: 65, 140, 0.46 / With mut: 65, 143, 0.45	-0.24	G= 0.276

*Frequency of existing variant in 1000 Genomes combined population.

The functional effects of nsSNPs associated with higher COVID-19 risk are described in Table 2.

Gene	SNP ID	Global MAF *	Amino Acid Change	SIFT	PolyPhen 2	PhD-SNP	SNPs & GO	Predict SNP 2*	FunSeq 2	CADD	DANN	GWAVA	FATHMM
IFIH1	rs1990760	T= 0.356	A946T	T	T	T	T	T	D	T	T	a	T
NOS3	rs1799983	G- 0.823	D298E	T	T	T	T	T	D	T	T	D	T
IL-6R	rs2228145	C=0.293	D358A	T	T	T	T	T	D	T	T	T	T
TYK2	rs34536443	C= 0.001	P1104A	D	D	T	T	D	D	D	D	D	D
EFNA4	rs114301457	T= 0.002	F124L	T	D	D	D	T	T	D	T	D	T
PLSCR1	rs343320	A= 0.029	H262Y	T	D	T	T	D	D	D	D	D	D
IFNA10	rs28368148	G= 0.004	W164C	D	D	D	D	T	D	D	T	D	T
SLC22A31	rs117169628	A= 0.007	P474L	D	D	D	D	T	T	T	T	a	D
P2RX7	rs208294	C= 0.530	Y155N	D	D	D	D	D	T	D	D	a	D
TLR7	rs179008	T= 0.118	Q11P	T	T	D	D	T	T	T	T	a	T
IFNL1	rs30461	G=0.251	N188D	T	T	T	T	T	T	T	T	T	T
DDX58	rs10813831	A= 0.184	R7C	D	D	T	D	D	D	T	D	D	D

a Not found. *Frequency of existing variant in 1000 Genomes combined population. # Predict SNP 2. D= Deleterious. T= Tolerated.

The structural effects of nsSNPs associated with higher COVID-19 risk are described in Table 3.

Gene	SNP ID	Stability	SS Element	Torsion	Predicted ΔΔG (kcal/mol)
IFIH1	rs1990760	Stabilising	Helix	Unfavourable	0.47
NOS3	rs1799983	Stabilising	Other (turns, coils, etc.)	Unfavourable	0.28
IL-6R	rs2228145	Destabilising	Other (turns, coils, etc.)	Favourable	-0.57
TYK2	rs34536443	Stabilising	Helix	Favourable	5.09
EFNA4	rs114301457	Destabilising	Other (turns, coils, etc.)	Favourable	-1.27
PLSCR1	rs343320	Destabilising	Sheet	Favourable	-3.16
IFNA10	rs28368148	Destabilising	Helix	Unfavourable	-2.19
SLC22A31	rs117169628	Stabilising	Helix	Favourable	0.88
P2RX7	rs208294	Destabilising	Sheet	Unfavourable	-2.47
TLR7	rs179008	Destabilising	Helix	Unfavourable	-0.87
IFNL1	rs30461	Destabilising	Other (turns, coils, etc.)	Favourable	-1.64
DDX58	rs10813831	Destabilising	Helix	Unfavourable	-1.51

The structural patterns in relation to the molecular alterations and conservation of the nsSNPs associated with a higher risk of COVID-19 are described in Table 4.

Gene	SNP ID	MutPred		ConSurf Conservation Profile
		PROSITE and ELM Motifs	Molecular Mechanisms	
TYK2	rs34536443	None	None	Highly conserved, exposed, functional residue
EFNA4	rs114301457	ELME000133, ELME000249, PS01299	Altered Transmembrane protein	Highly conserved, buried, structural residue
PLSCR1	rs343320	None	None	Highly conserved, exposed, functional residue
IFNA10	rs28368148	PS00252	Altered Metal binding; Altered Ordered interface; Loss of Allosteric site at W164	Highly conserved, buried, structural residue
SLC22A31	rs117169628	ELME000149, ELME000336, ELME000337, PS00008	Altered Transmembrane protein; Gain of Helix	Highly conserved, exposed, functional residue
P2RX7	rs208294	None	Altered Transmembrane protein; Altered Metal binding; Loss of Strand; Gain of Disulfide linkage at C152; Loss of Proteolytic cleavage at R151; Gain of GPI-anchor amidation at N158	Average conserved, exposed residue.
TLR7	rs179008	None	None	Variable, exposed residue.
DDX58	rs10813831	None	None	Variable, exposed residue.

CONCLUSIONS

A total of 9 exonic SNPs (1 sSNP and 8 nsSNPs) were indicated here as potential candidates for further *in vivo* studies for COVID-19, as they may alter protein stability, interactions, and functional motifs that may be associated with antiviral response pathways.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Thakur, R. & Shankar, J. In silico Analysis Revealed High-risk Single Nucleotide Polymorphisms in Human Pentraxin-3 Gene and their Impact on Innate Immune Response against Microbial Pathogens. Front. Microbiol. 7, (2016).

Ghanavi, J., Farnia, P., Farnia, P. & Velayati, A. Human genetic background in susceptibility to tuberculosis. Int J Mycobacteriol 9, 239 (2020).

Sivashankari, S. & Shanmughavel, P. Comparative genomics - A perspective. Bioinformation 1, 376 (2007).