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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Comprehensive in silico identification of impacts of ACE2 SNPs on COVID-19 susceptibility in different populations

Alireza Paniri a, b, Mohammad Mahdi Hosseini a, Mojtaba Moballeh-Eslam c, Haleh Akhavan-Niaki b, d, *

a Student Research Committee, Babol University of Medical Sciences, Babol, Iran
b Genetics Department, Faculty of Medicine, Babol University of Medical Sciences, Babol, Iran
c Student Research Committee, Sari Branch, Islamic Azad University, Sari, Iran
d Zoonoses Research Center, Pasteur Institute of Iran, Amol, Iran

ARTICLE INFO

Keywords: In silico ACE2 SARS-CoV-2 COVID-19 SNPs Polymorphism miRNAs

ABSTRACT

The COVID-19 pandemic emerges a reminder that wide spectrum discrepancy in response to SARS-CoV-2 infection and antiviral drugs among different populations might be due to their different ACE2 SNPs and/or miRNAs profile. ACE2 is the major component for SARS-CoV-2's cell entry, and disruption of its 3D structure could influence virus-ACE2 interaction. In this study we aimed to investigate the consequence of 16,860 SNPs within ACE2 on its expression as well as protein folding, function, and stability by using several beneficial bioinformatics tools. Only 64 SNPs including 60 intronic, and 4 missense showed different frequencies among different populations. Two missense SNPs including rs149039346 and rs147311723 have been predicted to strongly influence the function and stability of ACE2. rs1514283 creates new acceptor splice site. Also, rs4646175 creates new donor and acceptor splice site. PolymiRTS, and miRSNPs have predicted that rs3746444, rs113808830, and rs3751304 showed a MAF > 0.001, and disrupted mRNA target sites or mRNA function. Finally, rs3746444 hsa-miR-499a-3p, rs113808830 hsa-miR-4532, rs3751304 hsa-miR-6763-3p and hsa-miR-26b-5p were strongly hybridized with ACE2 and might influence its function. Collectively, this study shed some light on fundamental roles of ACE2 SNPs for its interaction with COVID-19, and consequently susceptibility to virus. Therefore, different responses of patients with COVID-19 to ACE2 blocker drugs might be due to their unique ACE2 SNPs. We further discussed the impact of SNPs on miRNAs profile as a factor that may modulate drug response or susceptibility to COVID-19.

1. Introduction

In late December 2019, a novel coronavirus known as SARS-CoV-2 firstly has emerged in Wuhan, China and further spread quickly into the global population. As of October 22, 2020 there were 41,520,997 confirmed cases and 1,136,887 deaths worldwide, including more than 15 million cases and 800,000 deaths in United States, Brazil, India, Russia, Italy, China, and United Kingdom (2020; Selvaraj et al., 2020). Given its high transmissibility and fatality rate, it might be a potential threat to global health, and have created an urgent need to explore its pathogenesis mechanism, develop rapid diagnostic tools, drugs, and investigate cellular processes that might influence individuals' susceptibility to SARS-CoV-2 and drug response. SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus that belongs to Coronaviridae family, and may damage several organs including lung, heart, kidney, gastrointestinal tract, etc (Dariya and Nagaraju, 2020; Ludwig and Zarbock, 2020; Zaim et al., 2020). SARS-CoV-2 enters into target tissues through exploiting two major proteins including transmembrane serine protease 2 (TMPRSS2) and angiotensin converting enzyme 2 (ACE2) on cell surface (Paniri et al., 2020). ACE2 is a zinc metalloproteinase, that plays a fundamental role in regulation of renin-angiotensin system (RAS) through two pathways: ACE2 converts angiotensin II (Ang II) into angiotensin I–7 (Ang 1–7), and cleaves a single residue from angiotensin I (Ang I) to generate Ang 1–9 (Chappel and Ferrario, 2006; Lambert et al., 2008; Xu et al., 2017). ACE2 is highly expressed in cardiovascular system, and kidney (Gheblawi et al., 2020; Groß et al., 2020). ACE2 is

Abbreviations: SNP, polymorphism; miRNAs, micro RNAs; ACE2, angiotensin converting enzyme 2; Ang, angiotensin; MAF, minor allele frequency.

* Corresponding author at: Department of Genetics, Faculty of Medicine, Babol University of Medical Sciences, Babol 4714957534, Iran.
E-mail address: halehakhavan@yahoo.com (H. Akhavan-Niaki).
the main receptor for SARS-CoV-2’s cell interaction and entry into target cells (Gheblawi et al., 2020). Given the importance of 3D structure of proteins in protein-protein interaction (PPI) it’s not surprising that any process resulting in ACE2’s 3D structure change could influence COVID-19’s cell entry. Accumulating evidence has shown that 3D structure of ACE2 might be influenced at both transcriptional and post-transcriptional levels. Single nucleotide polymorphisms (SNPs) are able to impact on protein function, structure, stability, and abundance (Calcagnile et al., 2020; Fadason, 2020). Moreover, noncoding RNAs including microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) could influence the expression of ACE2 at post-transcriptional level (Liu et al., 2015). Notably, different responses to SARS-CoV-2 infection among different populations raise the possibility that different SNPs and miRNAs profile might be responsible for wide spectrum clinical manifestations, and modulate drug response among different populations. Collectively, regarding the fundamental role of ACE2 in SARS-CoV-2’s cell entry, and as a potential therapeutic target for antiviral therapy we exploited various tools to present the in silico analysis of ACE2 to investigate related pathways, expression profile, epigenetic regulation, and SNPs of ACE2.

### 2. Materials and methods

#### 2.1. Analysis of ACE2 location and its variation within genome

Ensembl genome browser (https://asia.ensembl.org/index.html) that prepares comprehensive information about comparative genomics, evolution, gene sequence, splice variants, sequence variation, transcriptional regulation, regulatory function, protein domains, and disorders for most vertebrate species was used to analyze ACE2 genomic structure and its variations.

#### 2.2. Investigation of the effect of genetic variations on tissue-specific gene expression levels and exon expression by GTEx

Genotype-Tissue Expression (GTEx) available at https://gtexportal.org/home/ is a comprehensive in silico tool which could predict inherited susceptibility to diseases through analyzing the tissue-specific gene expression levels following genetic variation as previously described (6).

#### 2.3. Study of molecular pathways related to ACE2 through KEGG, GO

Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/) is a public web server that identifies biological pathways, diseases, drugs, and chemical substances. GO (gene ontology) (http://geneontology.org/) is a comprehensive computational tool which provides beneficial information including biological processes, molecular functions, cellular components, and molecular pathways for a wide spectrum of genes across all species.

#### 2.4. Databases and characterization of SNPs

National Center for Biotechnology Information (NCBI) website browser (https://www.ncbi.nlm.nih.gov/) and dbSNP were exploited to retrieve sequence and SNPs of ACE2, respectively.

NCBI is an extensive, powerful gene browser that presents a large number of data in the context of genes and proteins characterization, SNPs, mutation types, and clinical variations. ACE2 amino acid composition was obtained from universal protein resource (UniProtKB) database (https://www.uniprot.org/).

#### 2.5. Investigation of functional consequences of SNPs by SIFT

Sorting intolerant from tolerant (SIFT) (https://sift.bii.a-star.edu.sg/) is able to investigate the probable consequence of amino acids substitution (non-synonymous polymorphisms) on protein folding in human genome and non-human organisms as previously described (6). SIFT shows deleterious and benign substitutions with scores ranging from 0 to 0.05 and 0.05 to 1, respectively.

#### 2.6. Analysis of functional effects of SNPs by PolyPhen-2

Polyorphism phenotyping v2 (PolyPhen-2) (http://genetics.bwh.harvard.edu/pph2/), is another analyzer of effects of amino acid substitution on functional and structural proteins as previously described (6). Scores ranging from 0.0 to 0.15, 0.15 to 0.85, and 0.85 to 1.0 are considered benign, possibly damaging, and damaging, respectively.

#### 2.7. Analysis of functional effects of SNPs by PROVEAN

Protein variation effect analyzer (PROVEAN) (http://provean.jcvi.org/index.php) is a in silico database to investigation the impact of amino acid substitutions and on functions and structure of proteins as previously described (6). A score < −2.5 represents a “deleterious” variant while a score > −2.5 corresponds to a “neutral” variant.

#### 2.8. Investigation of functional impacts of SNPs by SNAP2

Analysis of functional impacts of SNPs by Screening for Non-acceptable Polymorphisms (SNAP2) (https://www.rostlab.org/services/snap/) is another free accessible bioinformatics tool that analyzes the impacts of different SNPs on protein function as previously described (6). In heat map, a score equal to −100 (dark blue) and +100 (dark red) depicts that amino acid substitution is completely neutral and pathogenic, respectively.

#### 2.9. Prediction of functional impacts of SNPs using phd-SNP

Predictor of human deleterious single nucleotide polymorphism (phd-SNP) https://snps.biofold.org/phd-snp/phd-snp.html, as a support vector machine, predicts those SNPs that cause disorders by influencing the protein function and structure. This server clusters the impacts of SNPs on the disease and scores vary between 0 and 9. The required inputs are including protein sequence, position, and new residue.

#### 2.10. Prediction of functional effects of SNPs via PANTHER

Protein Analysis Through Evolutionary Relationship (PANTHER) (available at https://www.pantherdb.org/) is a database that uses protein sequences in FASTA format, and can analyze substitution changes. It calculates the protein function alterations resulting from SNPs with Substitution Position-Specific Evolutionary Conservation (subPSEC) scores. A score less than −3 is predicted deleterious while 0 is predicted to be benign.

#### 2.11. Functional analysis of SNPs through Hidden Markov Models (Fathmm)

Functional Analysis of SNPs through Hidden Markov Models (Fathmm) https://fathmm.biocompute.org.uk/, is a powerful tool that calculates SNPs scores, and disorders related to blood, development, ear, nose and throat, endocrine, eye, heart, as well as nervous, genitourinary, immune, and musculoskeletal systems, and etc.

#### 2.12. Analysis of functional effects of SNPs through iPTREE-STAB

iptree-STAB available at http://210.60.98.19/IPTREEr/iptree.htm is a valuable tool that could predict and analyze the protein stability alterations (with the ΔΔG values) upon single nucleotide variation accompanied with high accuracy (%862) and high correlation coefficient (0.70) between predicted and experimental values. Necessary inputs for
Fig. 1. Ensembl gene map of ACE2 and its variants. (A) location of ACE2 on chromosome X and its structure including coding and noncoding region; (B) distribution of ACE2 variants reveals that most variants are intronic.
Fig. 1. (continued).
iPTREE-STAB are including deleted-residue mutation type, introduced-residue mutation type, and the neighbors found inside a symmetrical window centered at the mutated residue.

2.13. Prediction of the effects of mutations on splicing signals

Human splicing finder (HSF) is an advantageous software (http://www.umd.be/HSF/) that is able to forecast the disruption of the natural splice sites, and identify splicing motifs in all human sequence as previously described (6). For donor or acceptor splicing site recognition, HSF applies “position weight matrices” algorithm with consensus values (CV) that extend from 0 to 100. CVs higher than 65 are predicted as acceptor or donor splicing site. Moreover, wild type sequence score higher than 65 along with variation score less than –10% predicts that the mutation represents a new splice site. Moreover, a wild type sequence score less than 65 along with variation score higher than...
Fig. 2. (continued).
2.14. Prediction of molecular effects of ACE2 related-SNPs on protein secondary and tertiary structures

Protein homology/analogy recognition engine 2.0 (Phyre2) (http://www.sbg.bio.ic.ac.uk/_phyre2/html/page.cgi?id=index) is a potential tool that predicts the protein secondary structure (a-helices, b-strands and coils), ligand binding sites following amino acid variations (e.g. non-synonymous SNPs (nsSNPs)). GOR IV (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=NPSA/npsa_gor4.html) is another powerful predictor of protein secondary structure tool upon SNPs. PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) is a tool to prediction of secondary and tertiary structure of protein (beta sheets, alpha helixes and coils).

2.15. Analysis of three-dimensional structure of ACE2 by I-TASSER

I-TASSER available at (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) is a strong predictor of protein 3D structure. I-TASSER aims to determine by computational calculations the spatial location of every atom in a protein molecule from the amino acid sequence. The main purpose of I-TASSER is analyzing the iterative protein structure assembly along with prediction of solvent accessibility, normalized B-factor, ligand binding sites, and enzyme commission (EC) numbers and its active sites.

2.16. Prediction of post-translational modifications (PTM) by Modpred

Predictor of post-translational modification (PTM) sites in proteins (Modpred) (http://www.modpred.org/) is an advantageous and comprehensive tool to prediction of various types of PTMs such as...
acetylation, phosphorylation, proteolytic cleavage, methylation, O-linked glycosylation, N-linked glycosylation, and carboxylation as previously described (6).

2.17. Analysis of functional impacts of SNPs on secretory characteristics through phobius

Phobius (http://phobius.sbc.su.se/) is a combined predictor of transmembrane topology and signal peptide following amino acid substitution via analyzing the amino acid sequence. Phobius releases the results in form of a plot by calculating the total probability that a residue may belong to a helix, cytoplasmic or non-cytoplasmic domain.

2.18. Analysis of influence of polymorphisms on miRNAs function and development of severe disease by PolymiRTS, miRSNPs, and miRTarBase

PolymiRTS (available at http://compbio.uthsc.edu/miRSNP/) was used to predict the impacts of polymorphism on miRNA seed regions and its targeting profiling. MiRSNPs (http://bioinfo.bjmu.edu.cn/mirsnp/search/) is another suitable tool for analyzing the effect of polymorphisms on miRNAs profile that identifies more than 414,000 SNPs and their possible affects on miRNA-mRNA binding. MiRSNPs also provides results with a specific miRNA-mRNA binding energy with higher scores representing a more stable miRNA-mRNA binding. Moreover, miRTarBase (http://mirtarbase.cuhk.edu.cn/php/index.php) predicts targets of large number of genes via reporter assay, western blot, qPCR, microarray, next generation sequencing, and pulsed-SILAC (pSILAC).

2.19. Prediction of minimum free energy hybridization of ACE2 and miRNAs RNAhybrid v2.1.2

RNAhybrid (https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid/) is a tool for finding the minimum free energy hybridization of a long and a short RNA. The hybridization is performed in a kind of domain mode, i.e., the short sequence is hybridized to the best fitting part of the long one. The tool was primarily meant as a mean for miRNA target prediction.

3. Results

3.1. ACE2 genetic sequence and protein characterization

Achieved results from Ensembl have demonstrated that ACE2 is located on Xp22.2, and contains 19 exons. It also indicated contigs, regulatory elements such as promoter, enhancer, CCCTC-binding factor (CTCF), and coding and non-coding regions (Fig. 1A). Furthermore, ACE2 structural variants types and their position including copy number variation (CNV), gain, insertion, deletion, duplication, and tandem duplication are shown in Fig. 1B. Also, Gtx provided ACE2 location and eQTL categorizing of all ACE2 variants and their effects on ACE2 expression levels along with normalized effect size (NES) for several tissues including adipose-subcutaneous (ADPSBQ), adipose-visceral (Omentum) (ADPVSC), artery-tibial (ARTTBL), brain-amygdala (BRNAMY), brain-anterior cingulate cortex (BA24) (BRNACC), brain-caudate (basal ganglia) (BRNCDT), etc (Fig. 2A and B). Moreover, it showed that ACE2 is highly expressed in small intestine-terminal lumen, adipose visceral, testis, breast, and adipose subcutaneous tissues while it is weakly expressed in spleen, and pituitary gland. Gtx also have demonstrated that median read count per base for exons 12–16, and 19 in tests and small intestine-terminal lumen is remarkably higher than other tissues (Fig. 2C). Sequences of ACE2 isoforms were retrieved from UniPort. Further, the molecular function and pathways of ACE2 were investigated through KEGG (Fig. 3), and GO (Table S1).

3.2. SNPs retrieval within ACE2

NCBI was exploited to achieve all SNPs throughout the ACE2 along with SNPs characteristics including variants type including intrinsic, missense, synonymous, initiator codon variant, inframe variants, non-coding transcript variant, minor allele frequency (MAF), and SNPs frequency among different populations. Obtained results from dbSNP showed that 16,860 SNPS were located within ACE2 including intronic (15964), missense (611), synonymous (259), non-coding transcript variant (19), inframe deletion (5), and insertion (1), and initiator codon variant (1) (Fig. 4). Subsequently, our study was limited into SNPs with MAF > 0.001: therefore 318 SNPs were obtained. Investigation of 318 SNPs (308 intronic, 4 up-stream transcript variants, 2 synonymous, and
Prediction of function consequences of ACE2 missense SNPs on ACE2 structure and function by several powerful tools.

| SNPs | MAF | Substitution | SIFT prediction score | PolyPhen-2 prediction score | SNAP2 prediction score | PhD-SNP prediction score | PANTHER | FATHMM | iPTREE |
|------|-----|-------------|------------------------|----------------------------|------------------------|------------------------|----------|--------|--------|
| 1. rs4646116 | 1.00 | T | 0.9979 | K26R | 0.889 | Tolerated | Probable benign | 1.41 | 176 |
| 2. rs41303171 | 0.9955 | T | 0.9955 | N720D | 0.992 | Tolerated | Probable benign | 1.192 | 32 |
| 3. rs147311723 | 0.9952 | G | 0 | L731F | 0 | Deleterious | Probable disease | 71 | 7 |
| 4. rs149039346 | 0.9987 | A | 0.9987 | S692P | 0.774 | Possibly damaging | Probable damaging | 68 | 2 |

Several powerful databases such as Phyre2, GOR IV and PSIPRED were exploited to prediction of ACE2 SNPs on secondary structure of ACE2. They showed that ACE2 is mainly constituted by alpha helix (44.1%), and random coil (39.63%) (Fig. S1). Moreover, obtained results from Phyre2 have illustrated that rs149039346 (S > P) is located in disorder region with relatively high confidence, and change of serine into proline increases the possibility of disorder (Fig. S2).

### 3.3. Prediction of functional impacts of SNPs on protein function and stability

Several bioinformatics tools including SIFT, PolyPhen-2, PROVEAN, SNAP2, PhD-SNP, PANTHER, FTHMM, and iPTREE were involved to investigation of effect of 4 missense SNPs (rs4646116, rs41303171, rs147311723, and rs149039346) on protein function and stability (Table 1). SIFT, PolyPhen-2, SNAP2, PhD-SNP, FATHMM, and iPTREE predicted that 149039346 is deleterious and destabilizes ACE2. Moreover, 6 tools including SIFT, PolyPhen-2, SNAP2, PANTHER, FTHMM, and iPTREE have reported that rs147311723 alters the function and stability of ACE2. On the other hand, only 3 (PANTHER, FTHMM, and iPTREE) and 1 (iPTREE) software forecasted that rs41303171 and rs4646116 are damaging or destabilize ACE2, respectively. Uniprot also, have shown most consequences of ACE2 SNPs on protein function and phenotype. It predicted that most of ACE2 SNPs belonged partially to deleterious phenotype, and some of them created benign phenotype (Fig. 6).

### 3.4. Analyzing of ACE2 secondary structure upon SNPs

Several powerful databases such as Phyre2, GOR IV and PSIPRED were exploited to prediction of ACE2 SNPs on secondary structure of ACE2. They showed that ACE2 is mainly constituted by alpha helix (44.1%), and random coil (39.63%) (Fig. S1). Moreover, obtained results from Phyre2 have illustrated that rs149039346 (S > P) is located in disorder region with relatively high confidence, and change of serine into proline increases the possibility of disorder (Fig. S2).

### 3.5. Prediction and analysis of tertiary structure of proteins upon ACE2 SNPs

I-TASSER was exploited to analyze the 3D structure of ACE2 following different SNPs. The 3D structure of wild type ACE2 was analyzed and compared with rs149039346 ser692pro, and rs147311723 l731f. I-TASSER predicted that B-factor didn’t show any change between wild type and ACE2 with different SNPs. B-factor is a value to indicate the extent of the inherent thermal mobility of residues/atoms in proteins. Also, results showed that solvent accessibility for rs149039346 and rs147311723 were 3 and 1, respectively and were similar to wild type ACE2. It revealed that possible ligand binding site residues for wild type ACE2 are 151, 167, 168, 169, 228, 229, 230, 231, 232, 233, 234, 373, 381, 412, 413, 416, 522 (Fig. 7A). Moreover it indicated that possible ligand binding residues were changing upon

4 missense by 1000 genome browser have indicated that only the frequency of 64 SNPs was significantly different among different populations (Table S2). Of 64 SNPs, 29 intronic SNPs such as rs233574, rs233575, rs757066, rs1514279, and rs1514280 showed a significant difference in Asians in comparison with other populations. Moreover, 15 intronic (rs1514281, rs1514282, rs1514283, rs4830964, rs73635099, etc) and 2 missense SNPs including rs147311723, and rs149039346 were different between Africans and other populations. It’s noteworthy that 8 intronic SNPs including rs4646148, rs4646188, rs34481900, rs35482260, rs397822493, and rs200217737 and 2 missense SNPs including rs4646116, and rs41303171 presented a different frequency in European population relative to other populations (Fig. 5).
rs149039346 and rs147311723, and were 273, 274, 350, 378, 382, 402, 449, 503, 504, 505, 510, 512, and 515 respectively (Fig. 7B and C). I-TASSER also reported that enzyme commission (EC) numbers and active sites (402, and 515) were similar between wild type ACE2 and rs149039346 and rs147311723 (Fig. 7D).

3.6. Investigation of ACE2 SNPs on post-translational modifications (PTM) and secretory characteristics of ACE2

Investigation of impacts of ACE2 SNPs on PTM have showed that some of ACE2 residues may undergo modifications including acetylation, farnesylation, N-linked glycosylation, pupylation, ADP-ribosylation, geranylgeranylation, N-terminal acetylation, addition of pyrrolidone carboxylic acid, amidation, GPI-anchor amidation, O-linked glycosylation, sulfation, C-linked glycosylation, hydroxylation, palmitylation, SUMOylation, carboxylation, methylation, phosphorylation, ubiquitination, disulfide linkage, myristoylation, and proteolytic cleavage (Fig. 8). Among all modified residues neither of them showed a MAF > 0.001. Study of changes in secretion of ACE2 upon SNPs including rs4646116, rs41303171, rs147311723, and rs149039346 were performed by phobius which revealed that they didn’t influence ACE2 secretion (Fig. 9).

3.7. Analyzing miRNAs profile following ACE2 SNPs among different populations

PolymiRTS, miRSNPs, and miRTarBase were exploited to investigate possible impacts of ACE2 SNPs on function and biogenesis of miRNAs by influencing their binding to ACE2 transcripts. PolymiRTS have reported 5 SNPs which two of them including rs182366225, and rs142017934 influenced miRNAs target sites and 3 including rs3746444, rs113808830, and rs3751304 located in miRNAs seed causing disruption or creating new target site. Out of 5 predicted SNPs by PolymiRTS only 3 showed a MAF > 0.001 and different frequencies among different populations (Table 3). Among these 3 SNPs rs3746444 disrupted miRNA seeds target, rs113808830 and rs3751304 created miRNA seeds target. Furthermore, miRSNPs predicted two SNPs comprising rs182366225, and rs142017934 that were predicted by PolymiRTS and showed similar frequency among different population (Table 3D). Finally, miRTarBase by Microarray forecasted that hsa-miR-26b-5p strongly targeted ACE2 with high confidency.

3.8. Prediction of minimum free energy hybridization of ACE2 and miRNAs by RNAhybrid v2.1.2

RNAhybrid v2.1.2 was used to analyzing the minimum free energy (mfe) hybridization of ACE2 transcript and 4 miRNAs upon SNPs including rs3746444 hsa-miR-499a-3p, rs113808830 hsa-miR-4532, and rs3751304 hsa-miR-6763-3p and hsa-miR-26b-5p. Results showed that all four miRNAs could hybridize with ACE2 transcript with high stability (negative mfe). It’s noteworthy that the mfe for rs3746444 hsa-miR-499a-3p was significantly higher (~9.6 kcal/mol) representing the highest hybridization stability with ACE2 transcript (Table 4 and Fig. 10). Fig. 11 represents the effects of different SNPs on ACE2 transcript maturation or stability.

| Table 2 | Prediction of splicing site modifications following ACE2 SNPs. |
|---------|---------------------------------------------------------------|
| SNPs    | MAF   | Donor-site | Acceptor-site | Enhancer motif | Silencer motif |
|---------|-------|------------|--------------|----------------|---------------|
| rs714265 | C = 0.6917 | NA | NA | ESE site broken | New ESS site |
| rs1514283 | T = 0.8906 | New acceptor splice site | New ESE site | ESS site broken | New ESS site |
| rs4240157 | C = 0.3179 | NA | NA | NA | ESS site broken |
| rs4646135 | T = 0.9714 | New donor splice site | NA | NA | NA |
| rs4646175 | T = 0.9838 | New acceptor splice site | NA | NA | NA |

NA: not available; ESE: exonic splicing enhancer; ESS: exonic splicing silencer.
4. Discussion

High morbidity and mortality caused by current SARS-CoV-2 created an urgent need for better recognizing those mechanisms that lead to various responses to the virus infection among different populations. As demonstrated in our previous study, **TMPRSS2** SNPs could influence susceptibility to virus infection among different populations by changing **TMPRSS2** expression and function, splicing, PTM, and miRNAs profile (Paniri et al., 2020). Correspondingly, ACE2 is the main protein that allows SARS-CoV-2 cell entry, and thereby organ damaging. It’s noteworthy to mention that it also has a protective role against tissue injury (Ferreira, 2012). Therefore, the fundamental role of ACE2 in virus...
Table 3
Prediction of miRNA profile through SNPs by PolymiRTS and miRSNPs. (A) Prediction of SNPs in miRNAs target site by PolymiRTS; (B) prediction of SNPs in miRNAs seed causing disruption or creating target site by PolymiRTS; (C) prediction of miRNAs profile by miRSNPs; (D) frequencies of rs3746444, rs113808830, and rs3751304 among different populations. NA: not available; D: the derived allele disrupts a conserved miRNA site; C: the derived allele creates a new miRNA site; N: predicted target site with no experimental support.

### A.

| SNPs         | Variant type | Wobble base pair | Ancestral allele | Allele | miR ID          | Conservation | MiRSite   | Function class | Exp support | Context + score change |
|--------------|--------------|------------------|------------------|--------|----------------|--------------|-----------|------------------|-------------|------------------------|
| rs182366225  | SNP          | Y                | G                | G      | hsa-miR-483-3p | 6            | aactgtGAGTGA | D               | N           | -0.154                 |
|              |              |                  |                  |        | hsa-miR-139-5p | 5            | aACTGTAGtgta | C               | N           | -0.332                 |
|              |              |                  |                  |        | hsa-miR-4639-3p | 6            | aactgtAGTGA | C               | N           | -0.12                  |
|              |              |                  |                  |        | hsa-miR-582-5p | 4            | AACTGTAggtga | C               | N           | -0.15                  |
|              |              |                  |                  |        | hsa-miR-6794-3p | 6            | aactgtAGTGA | C               | N           | -0.218                 |
| rs142017934  | SNP          | Y                | A                | A      | hsa-miR-3609   | 10           | tagcTCACTTTca | D               | N           | -0.081                 |
|              |              |                  |                  |        | hsa-miR-548ah-5p | 10          | tagcTCACTTTca | D               | N           | -0.043                 |

### B.

| SNPs         | miR ID          | Location | miR seed | Allele  | Wobble base pair | MiRSite   | Conservation | Context + score change |
|--------------|-----------------|----------|----------|---------|------------------|-----------|--------------|------------------------|
| rs3746444    | hsa-miR-499a-3p | 15579491 | AC[A/G]UCAC | A/G   | GUGAUGU | 3 | -0.157 |
| rs113808830  | hsa-miR-4532   | 15579620 | [C/T]CCGGGG | C/T | CCUGGGA | 3 | -0.125 |
| rs3751304    | hsa-miR-6763-3p | 15579433 | UCCC[C/T]GG | C/T | CCAGGGAA | 2 | -0.303 |

### C.

| SNPs         | miRNA          | Effect  | Allele | Score | Energy | Conservation |
|--------------|----------------|---------|--------|-------|--------|--------------|
| rs182366225  | hsa-miR-139-5p | Decrease | G      | NA    | -13.85 | NA           |
|              | hsa-miR-582-5p | Enhance | G      | 140.00 | -15.08 | 0.001        |
| rs142017934  | hsa-miR-1246   | Decrease | G      | 143.00 | -13.93 | 0.000        |
|              | hsa-miR-3609   | create  | G      | 142.00 | -12.41 | 0.000        |
|              | hsa-miR-3646   | Enhance | G      | 157.00 | -13.45 | 0.132        |
|              | hsa-miR-548ah-5p | Create | G      | 161.00 | -13.44 | 0.000        |
infection created this hypothesis that different susceptibility to SARS-CoV-2 might be due to different ACE2 SNPs. Gtex have indicated that small intestine-terminal lumen, adipose subcutaneous, adipose-visceral (omentum), artery-coronary, heart (left ventricle), testis, thyroid, kidney cortex, and lung were highly expressing ACE2. In this view several studies showed that ACE2 is expressed in gut, lung, cardiovascular and breast tissues (Cao et al., 2020; Li et al., 2020; South et al., 2020; Zhang et al., 2019). Interestingly, investigation on rat middle cerebral artery occlusion model have illustrated that the ACE2 expression levels in patients with COVID-19 were higher in comparison with the control group which might put them at risk of cerebrovascular disease, and ischemic stroke (Choi et al., 2020). In the light of our finding and other several recent studies we speculate that higher susceptibility of patients with comorbidity including hypertension, coronary artery disease, diabetes mellitus, chronic kidney disease, obesity, smoking, and chronic obstructive pulmonary disease (COPD) to SARS-CoV-2 may be due to higher ACE2 expression in these groups (Guan et al., 2020; Jacobs, 2020; Saheb Sharif-Askari, 2020). Correspondingly, any change in the expression levels of ACE2 upon SNPs, PTM, splicing, and transcription processing could raise the vulnerability of individuals to COVID-19 infection. Furthermore, obtained results from NCBI showed that out of 16,860 SNPs throughout ACE2 only 64 SNPs including 60 intronic and 4 missense showed a MAF > 0.001 along with different frequencies among various populations. Therefore, computing 4 missense SNPs by several beneficial tools such as SIFT, PolyPhen-2, PROVEAN, PhD-SNP, PANTHER, FTHMM, and iPITREE have indicated that rs147311723, and rs149039346 influenced more the function and stability of ACE2 in comparison with rs4646116 and rs41303171. Surprisingly, obtained results from analyzing of SNPs by HSF have demonstrated that 5 SNPs including rs714205, rs1514283, rs4240157, rs4646135, and rs4646175 influenced ACE2 splicing, with rs714205 and rs1514283 creating ESE site broken, and new ESS Site. rs4646135 and rs4646175 led to the formation of new donor splice sites. It’s notable to mention that rs4240157 only caused ESS site broken. Achieved results from HSF raised the possibility that those SNPs disrupted splicing processes might affect ACE2 expression more than missense SNPs, and consequently could influence SARS-CoV-2’s cell entry. Notably, the MAF of those SNPs influencing splicing were remarkably higher than missense SNPs. Correspondingly, 1000 genome project revealed that African and American populations have a different frequency of rs147311723 G allele in comparison with other populations. Also, the African population has illustrated a remarkable different frequency of rs149039346 A allele in comparison with other populations. Regarding rs4646116 Asian and American populations have exhibited different frequency of T allele relative to other populations. Finally, African and Asian populations presented a different frequency of rs41303171 T allele in comparison with other populations. Moreover, investigation of ACE2 secondary

A. 

| miRNA   | Target   | Length | Mfe   |
|---------|----------|--------|-------|
| hsa-miR-499a-3p | ACE2 transcript | 122    | −91.6 kcal/mol |
| hsa-miR-26b-5p  | ACE2 transcript | 77     | −67.9 kcal/mol |
| hsa-miR-6763-3p | ACE2 transcript | 65     | −64.0 kcal/mol |
| hsa-miR-4532   | ACE2 transcript | 51     | −53.9 kcal/mol |

Table 4

MiRNAs properties predicted by RNAhybrid v2.1.2.

Fig. 10. ACE2 mRNA hybridization with different miRNAs. (A) hsa-miR-499a-3p; (B) hsa-miR-26b-5p; (C) hsa-miR-6763-3p; (D) hsa-miR-4532.
structure through hyre2, GOR IV, and PSIPRED showed that rs4646116 (K26R), rs41303171 (N720D), rs147311723 (L731F), and rs149039346 (S692P) were located in helix, coil, coil, and helix structure, respectively. Surprisingly, Phyre2 revealed that S692P probably caused disorder with relatively high confidence. Therefore, change in ACE2 structure might influence COVID-19 cell entry, and consequently its replication in host cells especially lung (Liu et al., 2020; Sommerstein et al., 2020). Subsequently, computing of ACE2 PTM via Modpred showed that K26, N720, L731, and S692 didn’t undergo any PTM. Similarly, possible changes in secretory characteristics of ACE2 upon SNPs were investigated through Phobius and showed no significant difference between wild type and missense SNPs. Eventually, PolymiRTS, miRSNP, and miRTarBase were engaged to analyze miRNAs profile change following ACE2 SNPs among different populations. Poly-miRTS, and miRSNP commonly have predicted that three SNPs including rs3746444 (disrupted hsa-miR-499a-3p seeds target), rs113808830 and rs3751304 (created hsa-miR-4532, and hsa-miR-6763-3p seeds target, respectively) influenced miRNAs profile among different populations with a MAF > 0.001. It’s noteworthy that South Asian and American populations have lower and higher frequency for rs3746444, respectively and rs113808830 showed the highest frequency in East Asians in comparison with other populations. Also, East Asian and American populations have demonstrated similar and lowest frequency for rs3751304 while Africans presented the highest frequency. Notably, a meta-analysis based on 52,456 individuals has revealed that microRNA-499 rs3746444 was significantly associated with increased risk of cancer of the respiratory system. Also, they have found a significant association between microRNA-499 rs3746444 and cancer risk in Asians but not Caucasians (Chen et al., 2014). Accordingly, there is some evidence about the fundamental role of rs3746444 in microRNA-499 in pathogenesis of other caners including breast and prostate with significant association in Asian population but not Caucasians (Kabirizadeh et al., 2016; Mi et al., 2018; Tan et al., 2020; Zou et al., 2012). Furthermore, miRTarBase have predicted that hsa-miR-26b-5p strongly targeted ACE2. Correspondingly, increasing body of evidence demonstrated the crucial role of miR-26a in combat with several viral infections including Herpes virus, Influenza A, and H1N1 Influenza A (Gao et al., 2017; Nguyen et al., 2018; Song et al., 2010; Zhang et al., 2020). Data from RNA-seq analysis indicated that miR-26a significantly increased innate anti-viral responses such as type I interferon (IFN), and IFN-stimulated genes expression. Further investigation showed that miR-26a potentially inhibits the replication of porcine

Fig. 11. ACE2 SNPs influencing SARS-CoV-2’s cell entry by disrupting ACE2 transcript processing or stability. Two missense SNPs including rs149039346, and rs147311723 have been predicted to change ACE2 folding. Furthermore, rs714205, rs1514283, rs4240157, rs4646135, and rs4646175 may interfere with ACE2 splicing process, and probably lead to lower ACE2 expression. Moreover, rs3746444 can disrupt hsa-miR-499a-3p seeds target, and rs113808830 and rs3751304 can create seeds target for hsa-miR-4532, and hsa-miR-6763-3p, respectively.
reproductive and respiratory syndrome virus by inducing innate anti-
immune response (Jia et al., 2015). Strikingly, achieved results from
RNAhybrid v2.1.2 have illustrated that these miRNAs are able to bind
ACE2 transcript with high stability. Taken together, RNAhybrid v2.1.2
have raised the possibility that ACE2 expression and its folding are more
influenced by SNPs located in non-coding regions in comparison with
those located in coding regions including missense and nonsense SNPs.
In contrast to our results a recent study has reported that ACE2
expression and polymorphisms didn’t show any significant difference
between Asians and other populations. This discrepancy might be due to
the fact that the fundamental role of SNPs in miRNAs expression and
function as well as the impact of SNPs on ACE2 splicing processes were
ignored in the corresponding study (Chen et al., 2020). Collectively, as
we are at the beginning of the high and wide way of achieving COVID-
19s pathogenesis, further investigation of factors that may influence
SARS-CoV-2s cell entry and replication is necessary.

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.genrep.2020.100979.

Declaration of competing interest

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

Acknowledgements

None.

Funding

This research did not receive any specific grant from funding
agencies in the public, commercial, or not-for-profit sectors.

References

Calcanoglith, M., Forger, P., Iannelli, A., Bucci, C., Alifano, M., Alifano, P., 2020. ACE2
Polymorphisms and Individual Susceptibility to SARS-CoV-2 Infection: Insights From
an In Silico Study. bioRxiv.

Cao, Y., Li, L., Feng, Z., Sun, H., Huang, P., Sun, X., Wang, W., 2020. Comparative
genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2
in different populations. Cell Discovery 6 (1), 1–4.

Chappel, M., Ferrario, C., 2006. ACE and ACE2: their role to balance the expression of
angiotensin II and angiotensin-(1–7). Kidney Int. 70 (1), 8–10.

Chen, C., Yang, S., Chaugai, S., Wang, Y., Wang, D.W., 2014. Meta-analysis of Hsa-mir-
499 polymorphism (rs3746444) for cancer risk: evidence from 31 case-control
studies. BMC Medical Genetics 15 (1), 126.

Chen, Y., Shan, K., Qian, W., 2020. Asians do not exhibit elevated expression or unique
expression and polymorphisms didn’t show any significant difference
between Asians and other populations. This discrepancy might be due to
the fact that the fundamental role of SNPs in miRNAs expression and
function as well as the impact of SNPs on ACE2 splicing processes were
ignored in the corresponding study (Chen et al., 2020). Collectively, as
we are at the beginning of the high and wide way of achieving COVID-
19s pathogenesis, further investigation of factors that may influence
SARS-CoV-2s cell entry and replication is necessary.

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.genrep.2020.100979.

Declaration of competing interest

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

Acknowledgements

None.

Funding

This research did not receive any specific grant from funding
agencies in the public, commercial, or not-for-profit sectors.

References

Calcanoglith, M., Forger, P., Iannelli, A., Bucci, C., Alifano, M., Alifano, P., 2020. ACE2
Polymorphisms and Individual Susceptibility to SARS-CoV-2 Infection: Insights From
an In Silico Study. bioRxiv.

Cao, Y., Li, L., Feng, Z., Sun, H., Huang, P., Sun, X., Wang, W., 2020. Comparative
genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2
in different populations. Cell Discovery 6 (1), 1–4.

Chappel, M., Ferrario, C., 2006. ACE and ACE2: their role to balance the expression of
angiotensin II and angiotensin-(1–7). Kidney Int. 70 (1), 8–10.

Chen, C., Yang, S., Chaugai, S., Wang, Y., Wang, D.W., 2014. Meta-analysis of Hsa-mir-
499 polymorphism (rs3746444) for cancer risk: evidence from 31 case-control
studies. BMC Medical Genetics 15 (1), 126.

Chen, Y., Shan, K., Qian, W., 2020. Asians do not exhibit elevated expression or unique
expression and polymorphisms didn’t show any significant difference
between Asians and other populations. This discrepancy might be due to
the fact that the fundamental role of SNPs in miRNAs expression and
function as well as the impact of SNPs on ACE2 splicing processes were
ignored in the corresponding study (Chen et al., 2020). Collectively, as
we are at the beginning of the high and wide way of achieving COVID-
19s pathogenesis, further investigation of factors that may influence
SARS-CoV-2s cell entry and replication is necessary.

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.genrep.2020.100979.

Declaration of competing interest

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

Acknowledgements

None.

Funding

This research did not receive any specific grant from funding
agencies in the public, commercial, or not-for-profit sectors.

References

Calcanoglith, M., Forger, P., Iannelli, A., Bucci, C., Alifano, M., Alifano, P., 2020. ACE2
Polymorphisms and Individual Susceptibility to SARS-CoV-2 Infection: Insights From
an In Silico Study. bioRxiv.

Cao, Y., Li, L., Feng, Z., Sun, H., Huang, P., Sun, X., Wang, W., 2020. Comparative
genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2
in different populations. Cell Discovery 6 (1), 1–4.

Chappel, M., Ferrario, C., 2006. ACE and ACE2: their role to balance the expression of
angiotensin II and angiotensin-(1–7). Kidney Int. 70 (1), 8–10.

Chen, C., Yang, S., Chaugai, S., Wang, Y., Wang, D.W., 2014. Meta-analysis of Hsa-mir-
499 polymorphism (rs3746444) for cancer risk: evidence from 31 case-control
studies. BMC Medical Genetics 15 (1), 126.

Chen, Y., Shan, K., Qian, W., 2020. Asians do not exhibit elevated expression or unique
expression and polymorphisms didn’t show any significant difference
between Asians and other populations. This discrepancy might be due to
the fact that the fundamental role of SNPs in miRNAs expression and
function as well as the impact of SNPs on ACE2 splicing processes were
ignored in the corresponding study (Chen et al., 2020). Collectively, as
we are at the beginning of the high and wide way of achieving COVID-
19s pathogenesis, further investigation of factors that may influence
SARS-CoV-2s cell entry and replication is necessary.

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.genrep.2020.100979.