Japanese Journal of Gastroenterology and Hepatology

Review Article

ISSN 2435-1210 | Volume 5

Putative Regulatory Element Located in the Introns 9 and 17 of the ACE2 Gene May Be Influenced By COVID-19 Risk Variants

Nair A[1*] and Rao AS[1]

[1]School of Basic and Applied Sciences, Dayananda Sagar University, Bengaluru, India

***Corresponding author:**

Ajay Nair,
School of Basic and Applied Sciences,
Dayananda Sagar University,
Bengaluru, India,
E-mail: ajay-sbas@dsu.edu.in

Received: 24 Oct 2020
Accepted: 04 Nov 2020
Published: 06 Nov 2020

Citation:

Nair A. Putative Regulatory Element Located in the Introns 9 and 17 of the ACE2 Gene May Be Influenced By COVID-19 Risk Variants. Japanese Journal of Gstroenterology and Hepatology. 2020; V5(2): 1-4.

Keywords:

ACE2; GWAS; Chromatin states; SNPs; COVID19

1. Abstract

Preliminary genomic data of the SARS-CoV2 virus responsible for COVID-19 showed that, its ability to infect relies on special cell surface spike proteins that have binding affinity to the human protein angiotensin-converting enzyme (ACE2). Among several patients analysed, it usually infects most age groups evenly. However, only a small percent of the confirmed patients progress into the severe phase. It is yet unclear, what accounts for this variation in disease propensity. Genome Wide Association Studies (GWAS) performed to map single-nucleotide polymorphisms (SNPs) associated with susceptibility to diseases show that 90% of GWAS signals reside in non-coding DNA elements (in the form of SNPs), with ~60% mapping to immune-cell enhancers. In this work we probed into the genomic region upstream of the ACE2 gene to identify dynamic chromatinsig natures that can serve as novel biomarkers and may point toward novel pathways for therapeutic intervention. In future, these studies will aithe ground work for the development of rational therapies aimed at restoring homeostatic mechanisms of ACE2 transcription.

2. Introduction

Host-pathogen interactions are highly variable among individuals. Genome Wide Association Studies (GWAS) have allowed unraveling of genetic determinants that account for the aforesaid variability. GWAS performed to map single-nucleotide polymorphisms (SNPs) associated with susceptibility to diseases show that 90% of GWAS signals reside in non-coding DNA elements (in the form of SNPs), with ~60% mapping to immune-cell enhancers [1]. What is peculiar about intergenic SNPs is their ability to get inherited as haplotypes, thereby escaping recombination events. This leads to conservation of disease risk alleles in isolated group of individuals or a population. CRISPR/cas9 mediated editing of such risk variant borne genomic region has shown the implications of such genetic determinants in infectious disease susceptibility [2]. Predisposition to chronic Hepatitis C virus has been attributed to polymorphisms in there promoter of chemokine and cytokine receptor genes [3]. Variation in the promoter elements for such genes has also shown to impact the course of HIV infection [4]. Several variants in the long non coding RNA has shown to confer susceptibility to Kaposi's sarcoma-associated herpes virus and HIV [5].

Similarly, the host genetic background has been shown to influence Severe Acute Respiratory Corona virus (SARS-CoV) that emerged in South Asia in 2002 [6]. However this data is insubstantial due to inconsistency in sample collection. Human amino peptidase N has been identified as the receptor for human corona virus (HCoV) 229E. Close to seven polymorphisms have been found in the intronic and exonic regions of the APN gene that is critical for corona virus binding [7]. Angiotensin Converting Enzyme-2 (ACE2) gene encodes a receptor protein that has shown binding affinity for cell surface spike (S) proteins on SARS-CoV2 virus [8]. Using, RNA sequencing, several coding region variants in ACE2 gene were analysed. However no direct evidence was found supporting the genetic resistance or predisposition to the viral infection [9].

Interferon mediated oligoadenylate synthetase (OAS) pathway is a key immune response triggered against dengue virus infection [10]. Recent study suggests that there might be an association between

polymorphisms in the OAS gene and susceptibility to the infection among patients. Significant association was also noticed between SNPs located in the 3'untranslated region of one of the OAS (OAS1) genes and the disease outcome [11]. Similarly, the pathogenesis and disease outcome of Hepatitis B Virus (HBV) infection also depends on several host genetic factors [12]. In particular, SNPs from the promoter regions of autophagy genes, such as IRGM and ATG161L, have shown protective association with HBV infection [13]. Impact of genetic variants on disease outcome also extends to plant infections. Mung bean Yellow Mosaic Indian Virus (MYMIV) is a Gemini virus, which infects soybean. Large-scale genome wide scanning identified several SNPS, non-synonymous and synonymous around defence related genes. The study generated a large-scale genome resource containing SNPs and INDELS that will allow the characterization of many complex traits [14].

2.1. Dark Matter of the Genome

Only 3% of the entire transcriptome is protein-coding, signifying the importance of moving the focus of research on host-pathogen interaction from a protein centric view to one that focuses on understanding the extensive non-coding DNA, its dynamic chromatin states, and its impact on transcriptional regulation [15]. Lately, much attention has been diverted towards characterizing host polymorphisms because most of phenotype associated SNPs reside in the non-coding region of the genome. The expression and functional variations within the lncRNA genes and their role in regulating immune response and inflammation is an area of intense research and likely to discover novel pathways of host-pathogen interaction [16].

In this study, we probed into the genomic region that harbours the ACE2 gene, to gain a preliminary understanding of the mechanisms that govern its transcriptional region [17-19]. This study will help further towards generating a dynamic portrait of enhancer chromatin marks in patient's pre and post COVID-19 flare [20].

3. Methods

The University of California Santa Cruz (UCSC) genome browser web based tool was initially used to display the snapshot of the genomic region around the ACE2 gene region [21]. The genome browser enables the incorporation of several annotation tracks. We also used data from ENCODE-ChIA-PET database was also incorporated in the UCSC genome browser to overlay information on the long and short range chromatin interactions between the genes and non-coding DNA elements. To characterize the potential risk SNPS in terms of their ability to affect chromatin interaction and gene transcriptional regulation , we also used the RegulomedB tool [22].

4. Results

4.1. Putative Regulatory Element in the Intronic Regions of ACE2 Gene Body

The ENCODE project has deposited vast amounts of data that provide unprecedented views of the regulatory and scape of the humangenome. To better understand the transcriptional regulation of ACE2, it is important to appreciate the three-dimensional (3D) chromatin state of the genomic region around ACE2. Heat maps of chromatin folding data from High Through put Chromatin Conformation Capture (Hi-C) experiments integrated into the UCSC genome browser tool (www.genome.ucsc.edu) (Figure 1) below indicate strong regulatory interactions in a region 111,113 bp within chromo some X. The proximity data are displayed as heat maps, with the intensity of the color corresponding to the strength of the interaction. As it is apparent from the figure (Figure 1), ACE2 falls in a region of extreme regulatory interactions suggesting the involvement of elements of the non-coding DNA, such as enhancers, that might play a seminal role in modulating ACE2 expression.

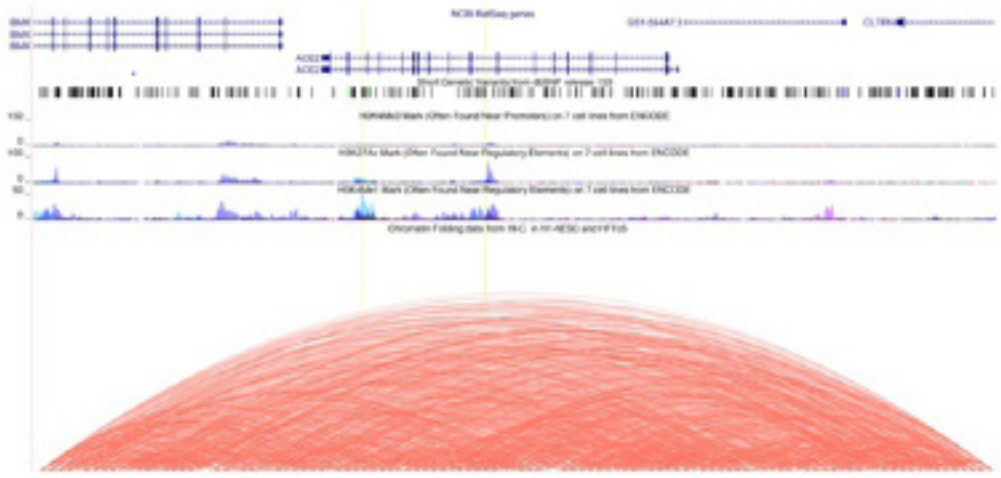

Figure1: 3D regulatory land scape of genomic region flanking *ACE2*.

The UCSC genome browser depicting 111, 113bp within chromo some X. We used the table browser tool in the UCSC Genome browser to cross reference the genomic region around ACE2 gene, with the ENCODE Integrated Regulation super- track, which contains locations of SNP variants, H3K4Me1/3Marks, H3K27Ac Marks, and heat maps of chromatin folding from HI-C data. The data suggest that the aforesaid genomic region is overlapping with strong regulatory interactions between enhancer-enhancer, and enhancer-promoter.

4.2. Regulatory Element in the ACE2 Gene Harbours COVID-19 Risk Variants

In our further mining of this data for the ACE2 locus we identified aputative enhancer in the intronic region that may be affected by SNP variants. Thus, we propose studies to develop evidence implicating this potential novel enhancer in ACE2 transcriptional regulation. ENCODE data demonstrate that this region overlap with enhancer marks (H3K4Me1, H3K27Ac), and chromatin immuno precipitation sequencing (ChIP-seq) clusters, and harbour SNP variants (Figure2) below.

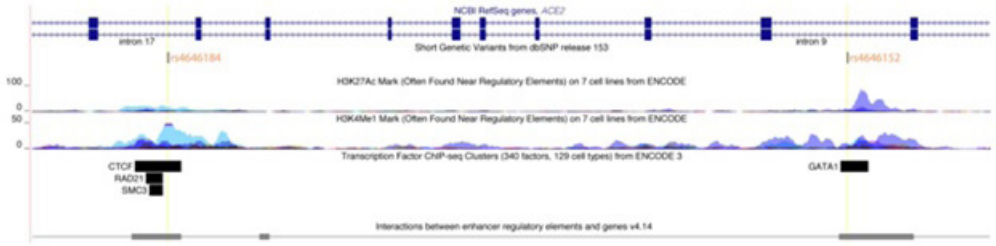

Figure2: SNP variants located imputative enhancer region**.**

Table 1: Description of regulome dB rank and score for putative SNPs situated on the intragenic DNA elements (intron9 and 17) of ACE2 gene. Low rank indicates increasing evidence for variant to be located in a functional region.

SNP ID	Rank	Score	Description (likely to affect)
rs4646152 (A/G)	2b	0.65908	TF binding + any motif + DNase Footprint + DNase peak
rs4646184 (G/A)	4	0.60906	TF binding + DNase peak

We used the table browser tool in the UCSC Genome browser to cross reference the intronic regions (intron 9 and17) with the ENCODE Integrated Regulation super-track, which contains H3K4Me1Marks, H3K27Ac Marks, transcription factor ChIP-seq data and regulatory element-gene interaction data. Potential risk variants (rs4646184, rs4646152) overlap with H3K4Me1and H3K27Ac Marks and are very close to Transcription factor binding sites (the darkness of the segment is proportional to the signal strength).

Among the various DNA binding proteins in the TF cluster, some of the key proteins overlapping the highlighted intronregions, have been displayed. Proteins such as, CTCF, RAD21, SMC3 and GATA1 have been well documented as influencers of chromatin loops. Such loops bring distant regulatory elements in close proximity to gene promoters. Cooperation of the aforesaid and many others DNA binding proteins is key to incrementing three-dimensional loop formation. Usually, methylation of the CCCTC binding site, prevents the CTCF like factors from docking, and allows regions of enhancer and promoter to come in close proximity.

Among several variants harboring the regions of interest (SNPs rs4646152 in intron 9 and rs4646184 in intron 17) was of particular interest because it overlaps enhancer his tone marks, Transcription Factor binding (TF) sites, chromatin looping tracks, and has a low regulo medB rank of 2B and 4 signifying high probability of the variant allele affecting TF binding, and enhancer-gene interaction (www.regulomedB.org).

5. Discussion

We believe that data from resources such as the UCSC genome browser, ENCODE, ChIA-PET, regulome Db etc. has provided sufficient insight into the novel putative enhancer that may be functionally impaired by risk variants carried on the ACE2 risk haplotype. Presence of the risk allele would contribute towards prolonged enhancer- promoter interaction, and the over expression of ACE2 would encourage the propensity towards getting infected with SARS-CoV2 COVID-19 virus. If these experiments verify the aforesaid, the data will further clarify how risk variants in the region of ACE2 influence COVID-19 through modulation of ACE2 expression.

References

1. Farh KK, Marson A, Zhu J et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015; 518: 337-43.
2. Baliou S, Adamaki M, Kyriakopoulos AM et al. CRISPR therapeutic tools for complex genetic disorders and cancer (Review). Int J Oncol.

2018; 53: 443-68.

3. Stanilova SA, Miteva LD, Karakolev ZT, Stefanov CS. Interleukin-10-1082 promoter polymorphism in association with cytokine production and sepsis susceptibility. Intensive Care Med. 2006; 32: 260–6.
4. Martin MP, Dean M, Smith MW, Winkler C, Gerrard B, Michael NL et al . Genetic acceleration of AIDS progression by a promoter variant of CCR5. Science. 1998; 282: 1907–11.
5. Sztuba-Solinska J, Rausch JW, Smith R, Miller JT, Whitby D, Le Grice SFJ et al. Kaposi's sarcoma- associated herpesvirus polyadenylated nuclear RNA: a structural scaffold for nuclear, cytoplasmic and viral proteins. Nucleic Acids Res. 2017; 45: 6805–21.
6. Kobayashi-Ishihara M, Yamagishi M, Hara T, Matsuda Y, Takahashi R, Miyake A et al. HIV-1- encoded antisense RNA suppresses viral replication for a prolonged period. Retrovirology. 2012; 9: 38.
7. Dominguez SR, O'Shea TJ, Oko LM, Holmes KV. Detection of group 1 coronaviruses in bats in North America. Emerg Infect Dis. 2007; 13: 1295–1300.
8. Vijgen, Leen; Keyaerts, Els; Zlateva, Kalina; Van Ranst, Marc. Identification of six new polymorphisms in the human coronavirus 229E receptor gene (aminopeptidase N/CD13). International journal of infectious diseases. 2004; 8: 217-22.
9. Zhang H, Penninger JM, Li Y, Zhong N, Slutsky AS. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target [published online ahead of print, 2020 Mar 3]. Intensive Care Med. 2020; 1–5.
10. Rossa W K Chiu, Nelson L S Tang, David S C Hui, Grace T Y Chung, Stephen S C Chim et al. ACE2Gene Polymorphisms Do Not Affect Outcome of Severe Acute Respiratory Syndrome, Clinical Chemistry. 2004; 50: 1683–6.
11. THAMIZHMANI, Ramanathan and VIJAYACHARI, Paluru. Association of dengue virus infection susceptibility with polymorphisms of 2'-5'-oligoadenylate synthetase genes: a case–control study. Braz J Infect Dis. 2014; 18: 548-550.
12. Chaaithanya IK, Bhattacharya D, Muruganandam N et al. Dengue: a newly emerging viral infection in Andaman and Nicobar Islands. India Epidemiol Infect. 2012.140:1920–4.
13. Knapp S, Yee LJ, Frodsham AJ, et al. Polymorphisms in interferon induced genes and the outcome of hepatitis C virus infection: roles of MxA, OAS-1 and PKR. Genes Immun. 2003; 4: 411–9.
14. THAMIZHMANI, Ramanathan and VIJAYACHARI, Paluru. Association of dengue virus infection susceptibility with polymorphisms of 2'-5'-oligoadenylate synthetase genes: a case–control study. Braz J Infect Dis. 2014; 18: 548-550.
15. Yadav CB, Bhareti P, Muthamilarasan M et al. Genome-wide SNP identification and characterization in two soybean cultivars with contrasting Mungbean Yellow Mosaic India Virus disease resistance traits. PLoS One. 2015; 10: e0123897.
16. Boivin V, Deschamps-Francoeur G, Couture S et al. Simultaneous sequencing of coding and non coding RNA reveals a human transcriptome dominated by a small number of highly expressed non coding genes. RNA. 2018; 24: 950–965.
17. Atianand MK, Fitzgerald KA. Long non-coding RNAs and control of gene expression in the immune system. Trends Mol Med. 2014; 20: 623–631.
18. Cao Y, Li L, Feng Z et al. Comparative genetic analysis of the novel corona virus (2019- nCoV/SARS-CoV-2) receptor ACE2 in different populations. (2020). Cell Discovery. 2020; 6: 11.
19. Jia HP, Look DC, Shi L et al. ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. J Virol. 2005; 79: 14614–21.
20. Soldi M, Mari T, Nicosia L et al. Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct subpopulations of macrophage enhancers. Nucleic Acids Res. 2017; 45: 12195-213.
21. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM et al. The human genome browser at UCSC. Genome Res. 2002; 12: 996-1006.
22. Boyle AP, Hong EL, Hariharan M et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22: 1790-7.