SARS-CoV-2 mutations altering regulatory properties: Deciphering host's and virus's perspectives

Abul Bashar Mir Md. Khademul Islam [a,*], Md. Abdullah-Al-Kamran Khan [b]

[a] Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka, Bangladesh
[b] Department of Mathematics and Natural Sciences, BRAC University, Dhaka, Bangladesh

ARTICLE INFO

ABSTRACT

Since the first recorded case of the SARS-CoV-2, it has acquired several mutations in its genome while spreading throughout the globe. In this study, we investigated the significance of these mutations by analyzing the host miRNA binding and virus's internal ribosome entry site (IRES). Strikingly, we observed that due to the acquired mutations, five host miRNAs lost their affinity for targeting the viral genome, and another five can target the mutated viral genome. Moreover, functional enrichment analysis suggests that targets of both of these miRNAs might be involved in various host immune signaling pathways. Remarkably, we detected that three particular mutations in the IRES can disrupt its secondary structure which can consequently make the virus less functional. These results could be valuable in exploring the functional importance of the mutations of SARS-CoV-2 and could provide novel insights into the differences observed different parts of the world.

1. Introduction

The recent COVID-19 pandemic has posed a serious threat to the public health sector around the whole globe as it has already spread to 213 countries and territories. As of 9th May (2021), approximately 159 million people got infected with SARS-CoV-2 and more than three millions suffered death from COVID-19 (Worldometer, 2020). The woes of this pandemic are still on the rise as the detailed molecular aspects of the pathogenesis are still intangible.

Since the reporting of the first genome sequence of SARS-CoV-2, a huge number of genomes are being sequenced regularly. Compared to the reference genome, SARS-CoV-2 genome sequences isolated from different parts of the globe show the acquisition of different genomic variations. However, the functional implications of these variations are largely unknown and are not well-correlated with the disease pathobiology or infectivity.

MicroRNAs (miRNAs) are small non-coding RNA (ncRNAs) which can play important roles during a viral infection. miRNAs can provide benefit for the host by exerting antiviral mechanisms through the innate and adaptive immune systems (Trobaugh and Klimstra, 2017); on the other hand, they can also facilitate the viral propagation through

negative modulation of the host immune responses (Głobińska et al., 2014). It was also reported that mutations in the miRNA binding sites in the viral genome can expedite the viral proliferation during an infection (Trobaugh and Klimstra, 2017).

Ribosomal frame-shifting using the internal ribosome entry site (IRES) has previously been reported as important for coronaviruses as they use this site for synthesizing their proteins (Dinman, 2012). Moreover, alterations in the IRES can result in lethal effects on the SARS-CoV propagation and life-cycle (Plant et al., 2013). Kelly et al. suggested that SARS-CoV-2 possesses similar IRES structure like SARS-CoV (Kelly et al., 2020).

In this present study, we sought out the comparative effects of the SARS-CoV-2 genome mutations on the basis of the host miRNA targeting profiles and IRES alterations. We illuminated the putative alterations of the host-miRNA binding profiles which might have arisen due to the acquired mutations in different world-wide isolates. Furthermore, the probable structural alterations in SARS-CoV-2 IRES due to the acquired variations were investigated.

---

2. Materials and methods

2.1. Mining of Single Nucleotide Polymorphisms (SNPs) and their world-wide frequencies

We extracted the SNPs of SARS-CoV-2 which have world-wide frequency ≥ 10 for the miRNA binding analysis. On the other hand, we took all SNPs found in the IRES region of SARS-CoV-2 for analyzing their effects on IRES. We obtained these information from "The 2019 Novel Coronavirus Resource" database along with the SNP annotations as on 30th April 2020 which included information from 8614 "High quality" (As set by GISAID database: Complete and high coverage sequences only, with <1% Ns, <0.05% unique amino acid mutations, and no insertion or deletion unless confirmed by submitter) SARS-CoV-2 genomes (Zhao et al., 2020).

2.2. Viral RNA-host miRNA interaction analysis

We have only taken fragmented viral sequences around our targeted viral reference sequence and SNPs (20 nts upstream and 20 nts downstream of a SNP) for the host miRNA-viral RNA interaction analysis. Subsequently, we used three different tools, namely- RNAhybrid (Krüger and Rehmsmeier, 2006), miRanda (Betel et al., 2008), IntaRNA (Mann et al., 2017) for analyzing the miRNA-RNA interaction. We considered those as high confidence interaction when it has been predicted by every tools that we used and have values- (i) for RNAhybrid: MFE ≤ -35 KJ/mol and *p*-value <0.001, (ii) for miRanda: energy ≤ -15 KJ/mol, (iii) for IntaRNA: $\Delta\Delta G \leq -15$ KJ/mol. We compiled the commonly predicted miRNAs from all of these three tools to reduce the false positives predictions. We did this interaction analysis for every SNPs and their associated wild-types separately. We then figured out the host miRNAs which can bind preferentially around the regions of the target SNPs, either in presence of the wild-type or the mutated allele (Supplementary file 1).

2.3. Retrieval of country-wise COVID-19 statistics

COVID-19 related statistics (Total deaths, Deaths/1 Million population, Case fatality rate) from different countries were extracted from the Worldometer website (Worldometer, 2020) on 24th April 2020 (Supplementary file 2). We clustered the country-wise COVID-19 fatality statistics along with the observed SNPs of the associated country. Furthermore, we integrated these statistics with the miRNA binding profiles around the respective SNPs. This was performed to shed insights on the probable role of altered miRNA binding patterns in the COVID-19 severity.

2.4. Functional enrichment analysis using the target genes of predicted host miRNAs

The experimentally validated targets of host miRNAs that can bind around the SNPs of SARS-CoV-2 were obtained from mirTarBase database (Huang et al., 2019a). Then we performed the functional enrichment analysis using Gitools v1.8.4 (Perez-Llamas and Lopez-Bigas, 2011) utilizing KEGG (Kanehisa and Goto, 2000) and Bioplanet (Huang et al., 2019b) pathway modules. Resulting *p*-values were adjusted for multiple testing using the Benjamin and Hochberg's method of False Discovery Rate (FDR) (Benjamini and Hochberg, 1995). We made two different groups of the targets on the basis of the associated miRNA's binding profile, namely- "Gained" (these miRNAs cannot bind viral RNA if the wild-type alleles are present; but upon the gaining of the mutations of those alleles, they can target the viral genome) and "Lost" (these miRNAs are capable of binding the wild-type alleles of the viral genome, but they lost their targeting capabilities upon the acquisition of the mutations)

2.5. Prediction of the secondary structures of the IRES

The coordinates of the experimentally validated IRES of SARS-CoV-2 was extracted from previously conducted study (Kelly et al., 2020), and we adjusted the coordinates accordingly. While predicting the secondary structures of the IRES, we took 50 additional nucleotides both upstream and downstream from the coordinates of the IRES. We used RNAfold tool (Gruber et al., 2008) (with default configuration) for the prediction of the IRES secondary structures. We modeled the IRES secondary structures with the reported mutations (considering one mutation at a time). As we observed 19 different mutations within the IRES region of SARS-CoV-2 (Supplementary file 1), we predicted the secondary structures for all these 19 mutations. Then we checked for the inconsistent IRES structures which might have resulted due to the incorporation of the mutation.

3. Results and discussion

We have obtained 377 high confidence (mutation frequency ≥ 10) SNPs from the database and used the regions around those for miRNA binding analysis (Data not shown). Ten miRNAs were observed targeting the wild-type and mutated polymorphic sites of the SARS-CoV-2 genome (Supplementary file 1). Five miRNAs were estimated to bind only the regions around the reference alleles, where 5 different other miRNAs can only bind to the regions around the mutant alleles (Fig. 1). We have considered the miRNAs which can bind the mutated sequences as "Gained"; while the miRNAs which failed to bind the mutated sequences but can bind the reference allele were termed as "Lost" miRNAs. Most of these bindings were observed for the targeted regions of the N and ORF1ab gene sequence. Also, miRNAs were also found to bind the targeted regions of ORF3a, ORF7a, and 3′UTR (Fig. 1).

Previously, probable role of differential miRNA targeting pattern on the severity of COVID-19 pathogenicity was reported (Khan et al., 2020).We observed that the countries in which these mutations were much prevalent (Supplementary file 2), a less amount of host miRNAs can bind to the SARS-CoV-2 genome/transcripts because of the acquired mutations, whereas these miRNAs can bind to the associated wild-type SARS-CoV-2 sequence (Fig. 1). We also detected that countries having high fatalities from COVID-19 have more mutations. Probably, these mutations might resist the binding of host miRNAs which can provide a competitive edge for these mutated viruses over the wild-type virus. Still, more experimental evidences should be searched in order to find out the definitive roles of these mutations in escaping from the host antiviral miRNAs.

Moreover, the functional enrichment analysis using the targets of these observed miRNAs suggests that the loss of miRNA binding can impact the host adaptive and innate immune signaling pathways, and hypoxia responses (Fig. 2A–B). Whereas the targets of the gained miRNAs for the mutations were involved in pathways, such as interleukin signaling, and autophagy (Fig. 2A–B). So, the acquired mutations can lead to the alterations in host miRNA targeting which might result in the altered host immune responses during the viral infection.

Fernández-Miragall et al. reported that mutations in the IRES of Pelargonium flower break virus (PFBV) can diminish the IRES activity, thus reduce the infection potential of the viruses (Fernández-Miragall and Hernández, 2011). Similar phenomenon was also observed in different other studies (Kerr et al., 2016; Shimoike et al., 2006). We scrutinized 19 mutations in the IRES region of SARS-CoV-2 which was previously defined by Kelly et al. (Kelly et al., 2020). We aimed to illustrate the impact of the mutations in the secondary structure of the SARS-CoV-2 IRES. To achieve this goal, we predicted the secondary structures of the mutated and wild-type IRES using RNAfold tool (Gruber et al., 2008). Interestingly, we observed that among the 19 mutations, 3 can significantly alter the secondary structure of the IRES of SARS-CoV-2 (Fig. 2C). These 3 mutations in the IRES region were 13487: C > T, 13506: C > T, and 13507: G > A (Supplementary file 3).
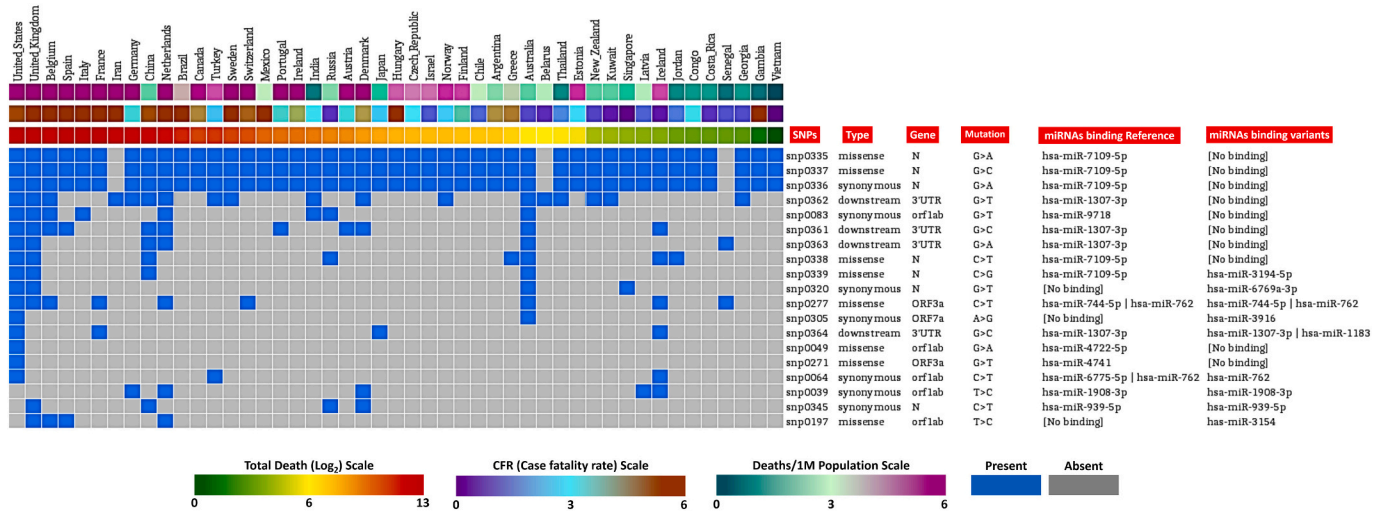
Fig. 1. Clustered country-wise distribution of the analyzed SNPs and their types, along with the case-fatality statistics. miRNAs which can target the viral RNA around these SNPs (in case of wild-type and mutated alleles) are also presented alongside the respective SNPs. Country-wise deaths/1 M population, CFR, and total deaths are also provided in color coded heatmaps. Color-coded clustering of these data was done in Gitools software.
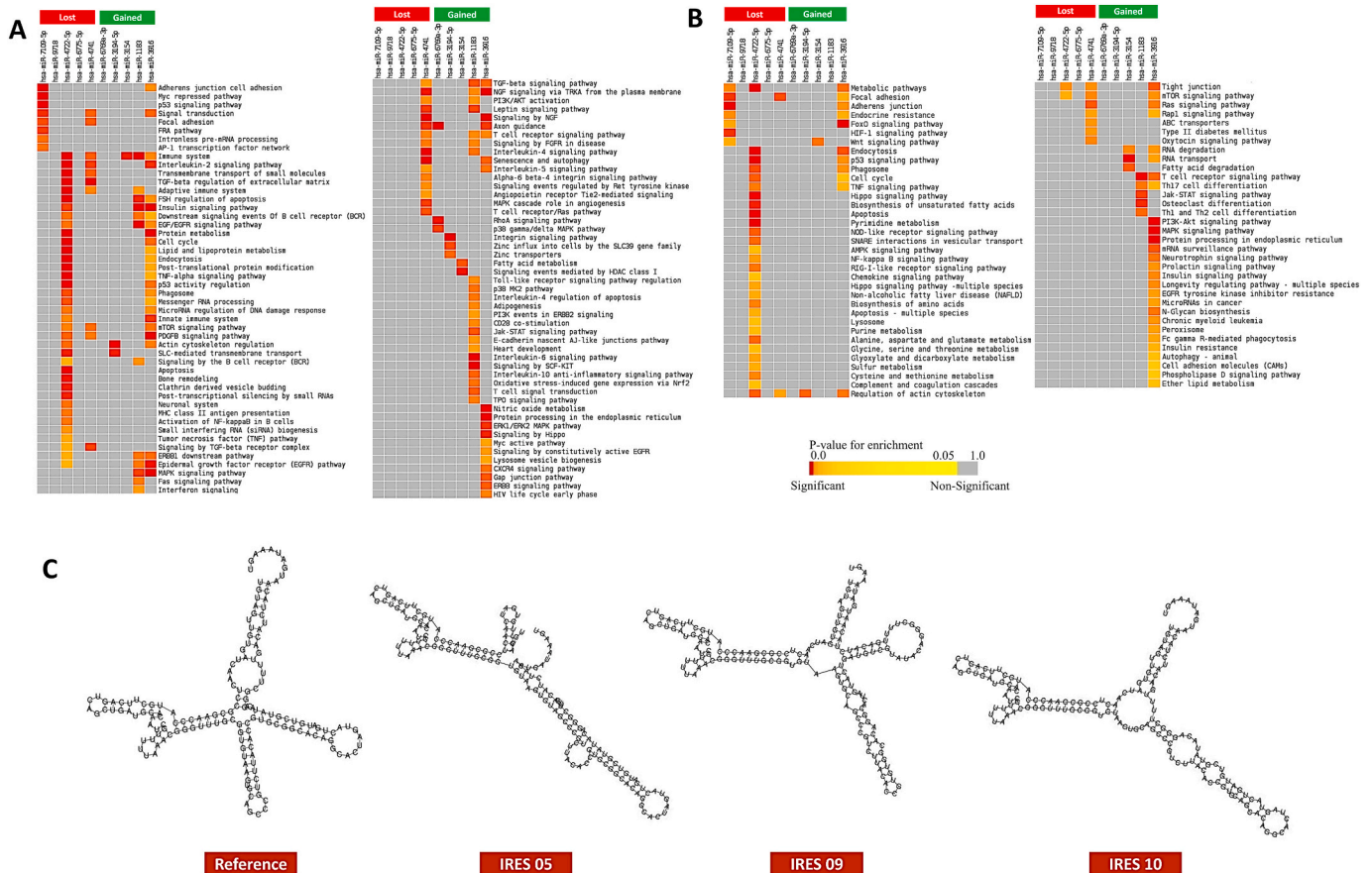

Fig. 2. Functional enrichment analysis using the targets of host miRNAs which can bind differentially around the mutations of SARS-CoV-2 utilizing **A.** KEGG pathway and, **B.** Bioplanet pathway modules. **C.** Secondary structure of the IRES of SARS-CoV-2 for the wild-type alongside the mutated ones. Significance of enrichment in terms of adjusted *p*-value (< 0.05) is represented in color coded *P*-value scale for all heatmaps. Color towards red indicates higher significance and color towards yellow indicates less significance, while grey means non-significant. Only selected significant enriched terms are shown.

Mutations in these positions might curb the viral translation process and render the virus less virulent. As, the studies regarding the IRES alterations are still in infancy, so more targeted researches should be conducted to elucidate the functional importance of these variations.

4. Conclusion

In our current analysis, we focused on the functional significance of the mutations reported for SARS-CoV-2 on the basis of host miRNA

binding and the putative alterations in the IRES structure. Our preliminary results can suggest that these mutations can alter the binding pattern of host miRNAs which can ultimately result in alternative responses from these binding. Also, we revealed that several mutations in the IRES can disrupt its secondary structure which might suggest why the virus is affecting less in some countries compared to the others. Taking these observations in count, further studies should be conducted to completely understand the functional effects of the acquired mutations in the SARS-CoV-2 genome, and integration of other clinical information with these results can provide more insights on the dynamics of the SARS-CoV-2 infection.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.genrep.2021.101236.

CRediT authorship contribution statement

ABMMKI conceived the project, designed the workflow, and performed the analyses. Both authors wrote the manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

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

Data availability statement

Publicly available data were utilized. Analyses generated data are deposited as supplementary files.

References

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57 (1), 289–300.

Betel, D., et al., 2008. The microRNA.org resource: targets and expression. Nucleic Acids Res. 36 (suppl_1), D149–D153.

Dinman, J.D., 2012. Mechanisms and implications of programmed translational frameshifting. Wiley interdisciplinary reviews. RNA 3 (5), 661–673.

Fernández-Miragall, O., Hernández, C., 2011. An internal ribosome entry site directs translation of the 3′-gene from Pelargonium flower break virus genomic RNA: implications for infectivity. PLoS One 6 (7), e22617.

Głobińska, A., Pawełczyk, M., Kowalski, M.L., 2014. MicroRNAs and the immune response to respiratory virus infections. Expert. Rev. Clin. Immunol. 10 (7), 963–971.

Gruber, A.R., et al., 2008. The Vienna RNA websuite. Nucleic Acids Res. 36 (Web Server issue), W70–W74.

Huang, H.-Y., et al., 2019a. miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. Nucleic Acids Res. 48 (D1), D148–D154.

Huang, R., et al., 2019b. The NCATS BioPlanet – an integrated platform for exploring the universe of cellular signaling pathways for toxicology, systems biology, and chemical genomics. Front. Pharmacol. 10 (445).

Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28 (1), 27–30.

Kelly, J.A., et al., 2020. Structural and functional conservation of the programmed - 1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). J. Biol. Chem. 295 (31), 10741–10748.

Kerr, C.H., et al., 2016. Molecular analysis of the factorless internal ribosome entry site in cricket paralysis virus infection. Sci. Rep. 6 (1), 37319.

Khan, M.A.-A.-K., et al., 2020. Epigenetic regulator miRNA pattern differences among SARS-CoV, SARS-CoV-2, and SARS-CoV-2 world-wide isolates delineated the mystery behind the epic pathogenicity and distinct clinical characteristics of pandemic COVID-19. Front. Genet. 11 (765).

Krüger, J., Rehmsmeier, M., 2006. RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res. 34 (suppl_2), W451–W454.

Mann, M., Wright, P.R., Backofen, R., 2017. IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. Nucleic Acids Res. 45 (W1), W435–W439.

Perez-Llamas, C., Lopez-Bigas, N., 2011. Gitools: analysis and visualisation of genomic data using interactive heat-maps. PLoS One 6(5).

Plant, E.P., et al., 2013. Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production. Viruses 5 (1), 279–294.

Shimoike, T., et al., 2006. Down-regulation of the internal ribosome entry site (IRES)-mediated translation of the hepatitis C virus: critical role of binding of the stem-loop IIId domain of IRES and the viral core protein. Virology 345 (2), 434–445.

Trobaugh, D.W., Klimstra, W.B., 2017. MicroRNA regulation of RNA virus replication and pathogenesis. Trends Mol. Med. 23 (1), 80–93.

Worldometer, 2020. Coronavirus Cases, pp. 1–22.

Zhao, W.M., et al., 2020. The 2019 novel coronavirus resource. Yi Chuan 42 (2), 212–221.