DOI: 10.1002/jmv.28149

LETTER TO THE EDITOR

JOURNAL OF MEDICAL VIROLOGY WILEY

Unveiling HERV-K113-ENV as SARS-CoV-2 severity admissible biomarker by mining transcriptome data

Dear Editor,

Discovering the severity biomarker for SARS-CoV-2 can reduce the unnecessary hospital occupancy of COVID-19-positive adult people. Here we report the human endogenous retrovirus K113 envelope (HERV-K113-ENV) transcription as an admissible SARS-CoV-2 severity biomarker by the mining of adult Indian whole-blood transcriptome data. The class I [Gammaretroviruses-HERV-W, ERV-FRD, etc.], class II (Betaretrovirus-HERV-K [HML 1–10], K113, K115, etc.), and class III (HERV-L) HERVs occupy around 8% of the human genome,[1] but it is largely unknown what kind of HERVs are present in the genome of the population in India. When mining the 39 whole-blood transcriptome data associated with SARS-CoV-2 of adult Indians in a single BioProject (PRJNA807370) to detect the prognosis biomarker of COVID-19, it was revealed that these transcriptomes contain transcripts of HERV-K113 (Supporting Information: Data 1), and the absence of HERV-W and HERV-FRD transcripts. Subsequently, when we examined whether there was any difference between the expression levels of HERV-K113 in severe/moderate (needs hospitalization care) and healthy/asymptomatic/mild COVID-19 cases (no hospitalization care required), we noted that there were no statistically significant changes between them (Figure 1A; Supporting Information:Table 1). However, in severe/moderate cases the expression of HERV-K113 seems to be slightly higher (Figure 1A), so we have examined whether the expression of particular genes of HERV-K113 differs. Remarkably, in our analysis, we found a statistically significant difference between the severe/moderate and healthy/asymptomatic/mild COVID-19 cases in terms of the read coverage percentage to the complete genome of HERV-K113 (Figure 1B,C; Supporting Information: Table 1). Further, we studied whether the gag-pro-pol and envelope transcripts of HERV-K113 in Indian transcriptomes could be translated into proteins without mutational defects. Our analysis revealed the presence of multiple stop codons in gag-pro-pol region (Supporting Information:Data 2–3), but most of the transcripts of the HERV-K113-ENV (envelope) gene found in this study were able to translate the entire HERV-K113-ENV protein with 97.3%–98.85% identity without any stop codon via the third frameshift (Supporting Information: Data 4–5 and Table 2). We then focus on HERV-K113-ENV and note that in 70.37% (19/27) healthy/asymptomatic/mild and 29.63% (8/27) severe/moderate COVID-19 cases HERV-K113-ENV is transcriptionally active (Figure 1D,E). On the other hand, 75.00% (9/12) of severe/moderate and 25.00% (3/12) of healthy/asymptomatic/mild COVID-19 cases HERV-K113-ENV is transcriptionally inactive (Figure 1F,G). Interestingly, the role of HERV-K113 in SARS-CoV-2 cases in the Brazilian population has also been reported,[2] further our analysis of Brazilian transcriptomes revealed that 60% (15/25) diseased/discharged cases of HERV-K113-ENV seem to be transcriptionally inactive, but transcriptionally active in 80% (4/5) non-COVID-19 cases. Collectively, the risk of hospitalization for COVID-19-positive humans becoming a severe/moderate condition is largely eliminated until the HERV-K113-ENV component is actively transcribed. Furthermore, the HERV-W-ENV[3] and HERV-FRD[4] expressions have also been found to be biomarkers that can predict the severity of SARS-CoV-2, but it is not widely known what HERV-W-ENV and HERV-FRD are in people in countries such as India. Overall, this study highlights the biomarker's determination for SARS-CoV-2 severity prediction, and HERV-K113 is present and actively transcribed in only 30% of the world's population,[5–7] so they may be expected to be less likely to be hospitalized by SARS-CoV-2, however, future experimental studies are needed to confirm this.

AUTHOR CONTRIBUTIONS

All the authors contributed significantly to this manuscript. Perumal Arumugam Desingu analyzed and wrote the first draft and Kumaresan Nagarajan reviewed the manuscript. All the authors reviewed and approved the final submission.

ACKNOWLEDGMENTS

P. A. D. is a DST-INSPIRE faculty supported by research funding from the Department of Science and Technology, India (DST/INSPIRE/04/2016/001067), and Science and Engineering Research Board, Department of Science and Technology, India (CRG/2018/002192).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Perumal Arumugam Desingu[1]
Kumaresan Nagarajan[2]

[1]*Department of Microbiology and Cell Biology, Division of Biological Science, Indian Institute of Science (IISc), Bengaluru, India*
[2]*Department of Veterinary Pathology, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University (TANUVAS), Chennai, India*

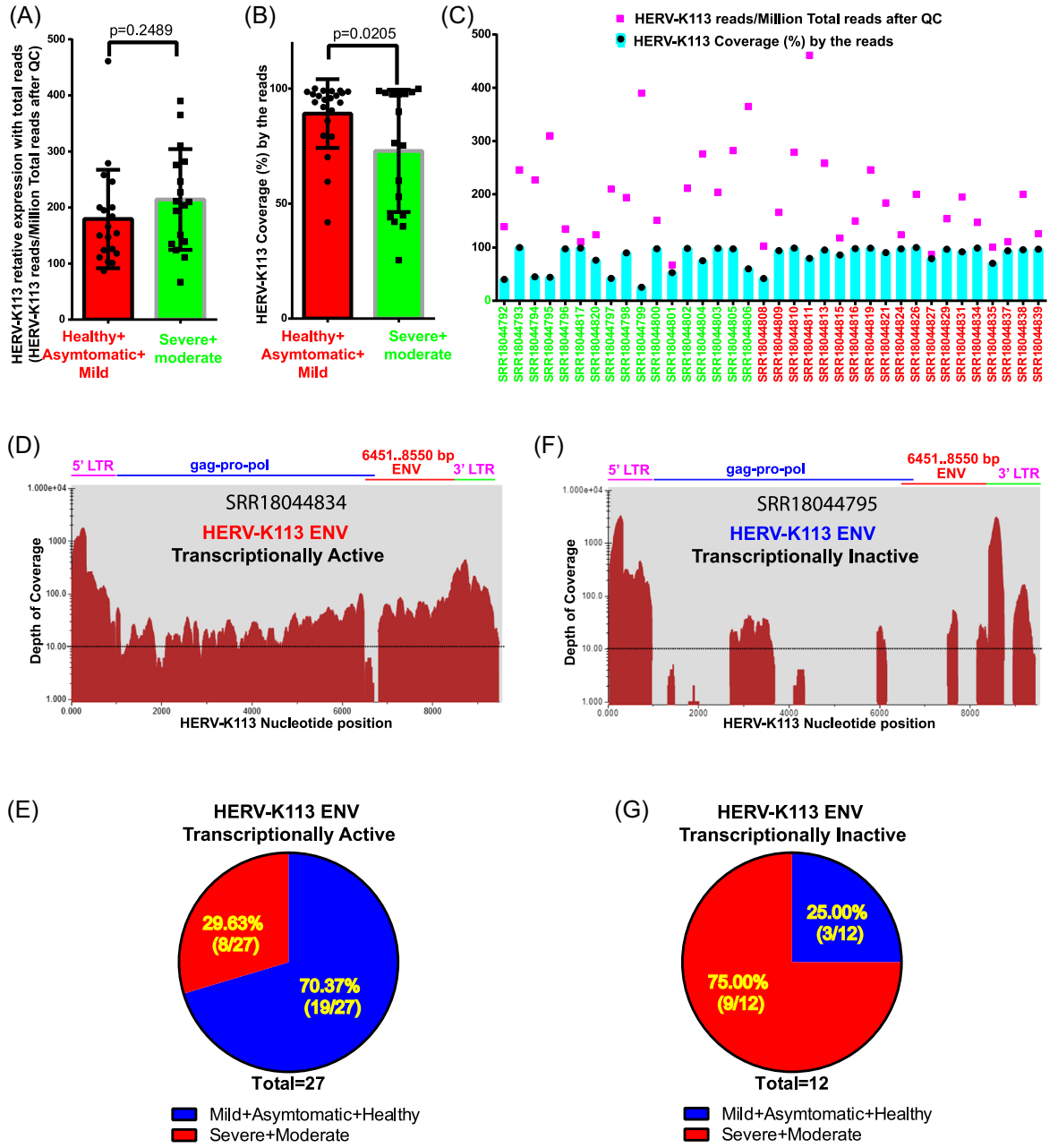

FIGURE 1 HERV-K113-ENV as SARS-CoV-2 severity biomarker. (A) Relative expression of HERV-K113 expression with total reads, HERV-K113 specific reads/million total reads after quality control, found no significant difference between healthy/asymptomatic/mild cases and severe/moderate cases. However, in severe/moderate cases there seem to be slightly higher reads per million (RPM) after quality control. RPM after quality control for each whole-blood transcriptome is presented in Supporting Information: Table 1. (B) HERV-K113 complete genome coverage percentage is significantly lower in severe/moderate cases than in healthy/asymptomatic/mild cases. HERV-K113 genome coverage (%) of the specific reads is presented in Supporting Information: Table 1. (C) Compare HERV-K113 specific reads/million total reads after quality control and HERV-K113 genome coverage (%) of the specific reads in each SRA Run file. The SRA Run files with low genome coverage (%) have been shown to have high HERV-K113 specific reads/million total reads after quality control. From these, transcriptome-derived HERV-K113 specific reads indicate that genes in regions where the virus' genome cannot be covered are not actively transcribed. (D) Representative map depicting the depth of HERV-K113 genome coverage by reads for the activity transcribed HERV-K113-ENV gene. The reads in the transcriptome cover almost the entire gag-pro-pol and envelope genes, indicating that these genes have been actively transcribed. A slight coverage gap in the envelope gene area indicates a deletion-type envelope. It is well known that 292 bp "deletion" HERV-K113 type is present in pol-ENV boundaries. (E) Transcriptionally active HERV-K113-ENV gene is found in 70.37% (19/27) healthy/asymptomatic/mild and 29.63% (8/27) severe/moderate COVID-19 cases. (F) Representative map depicting the depth of HERV-K113 genome coverage by reads for the transcriptionally inactive HERV-K113-ENV gene. The SRA Run files with low genome coverage (%) have been shown to have high HERV-K113 specific reads/million total reads after quality control. Despite the high HERV-K113 specific reads/million total reads after quality control in the SRA Run file transcriptome, it is clear that these genes are not actively transcription since HERV-K113 specific reads are unable to cover most areas of the envelope gene. (G) Transcriptionally inactive HERV-K113-ENV gene is found in 75.00% (9/12) of severe/moderate and 25.00% (3/12) of healthy/asymptomatic/mild COVID-19 cases. HERV-K113-ENV, human endogenous retrovirus K113 envelope.

Correspondence

Perumal Arumugam Desingu, Department of Microbiology and Cell Biology, Division of Biological Science, Indian Institute of Science (IISc), Bangaluru 560012, India.

Email: perumald@iisc.ac.in and padesingu@gmail.com

ORCID

Perumal Arumugam Desingu http://orcid.org/0000-0003-4395-4472

REFERENCES

1. Curty G, Marston JL, de Mulder Rougvie M, Leal FE, Nixon DF, Soares MA. Human endogenous retrovirus K in cancer: a potential biomarker and immunotherapeutic target. *Viruses*. 2020;12. doi:10.3390/v12070726
2. Temerozo JR, Fintelman-Rodrigues N, Dos Santos MC, et al. Human endogenous retrovirus K in the respiratory tract is associated with COVID-19 physiopathology. *Microbiome*. 2022;10:65. doi:10.1186/s40168-022-01260-9
3. Balestrieri E, Minutolo A, Petrone V, et al. Evidence of the pathogenic HERV-W envelope expression in T lymphocytes in association with the respiratory outcome of COVID-19 patients. *EBioMedicine*. 2021;66:103341. doi:10.1016/j.ebiom.2021.103341
4. Kitsou K, Kotanidou A, Paraskevis D, et al. Upregulation of human endogenous retroviruses in bronchoalveolar lavage fluid of COVID-19 patients. *Microbiol Spec*. 2021;9:e0126021. doi:10.1128/Spectrum.01260-21
5. Beimforde N, Hanke K, Ammar I, Kurth R, Bannert N. Molecular cloning and functional characterization of the human endogenous retrovirus K113. *Virology*. 2008;371:216-225. doi:10.1016/j.virol.2007.09.036
6. Moyes DL, Martin A, Sawcer S, et al. The distribution of the endogenous retroviruses HERV-K113 and HERV-K115 in health and disease. *Genomics*. 2005;86:337-341. doi:10.1016/j.ygeno.2005.06.004
7. Zwolińska K, Knysz B, Gąsiorowski J, et al. Frequency of human endogenous retroviral sequences (HERV) K113 and K115 in the Polish population, and their effect on HIV infection. *PLoS One*. 2013;8:e77820. doi:10.1371/journal.pone.0077820

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.