Supplementary Information for the article

Comparative genomics of global SARS-CoV-2 quasispecies offers insights into its microevolution and holds implications for pathogenesis and control

Santi M. Mandal[1], Suresh K. Mondal[1], Shriparna Mukherjee[2], Wriddhiman Ghosh[3] and Ranadhir Chakraborty[4,*]

[1] Central Research Facility, Indian Institute of Technology Kharagpur, Kharagpur 721302, West Bengal, India.
[2] Department of Botany, Prasannadeb Women's College, Jalpaiguri, West Bengal, India.
[3] Department of Microbiology, Bose Institute, P-1/12 CIT Scheme VII M, Kolkata 700054, West Bengal, India.
[4] Department of Biotechnology, University of North Bengal, Raja Rammohanpur, Darjeeling 734013, West Bengal, India.

* **Correspondence:** Prof. Ranadhir Chakraborty,
                              Department of Biotechnology,
                              University of North Bengal,
                              Raja Rammohanpur,
                              P.O. - NBU, District - Darjeeling,
                              West Bengal, Pin - 734013, India.
                              Telephone: +91-9434872273; Fax: +913532699001
                              E-mail: rcnbusiliguri@gmail.com

Contents

Supplementary Tables

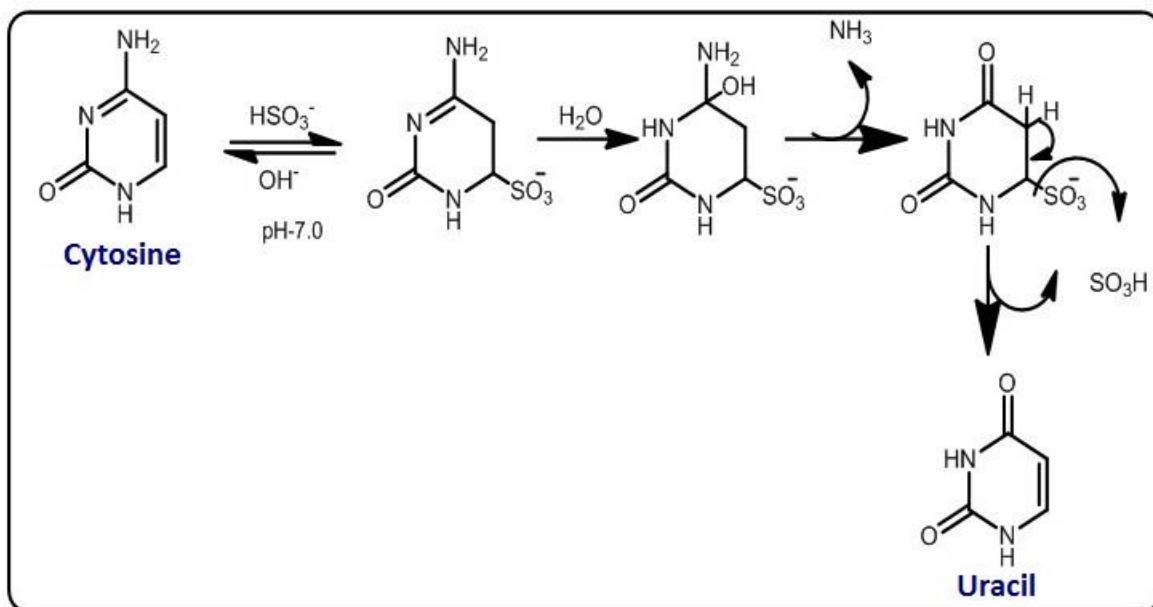
Table S1. Identity of the 46 SARS-CoV-2 genomes included in the India-specific dataset. Since this Table is more than one page long it has been provided as an Excel sheet named Table S1, within the Excel Workbook named Supplementary_Dataset.

Table S2. Gene-wise localization and molecular character of all the point mutations that were identified in one or more of the 46 SARS-CoV-2 genomes included in the India-specific dataset. Since this Table is more than one page long it has been provided as an Excel sheet named Table S2, within the Excel Workbook named Supplementary_Dataset.

Supplementary Figure

Figure S1. Probable chemical reactions causing transitions of C→U and A→G.

References used in the the legend of Figure S1

(a)

(b)

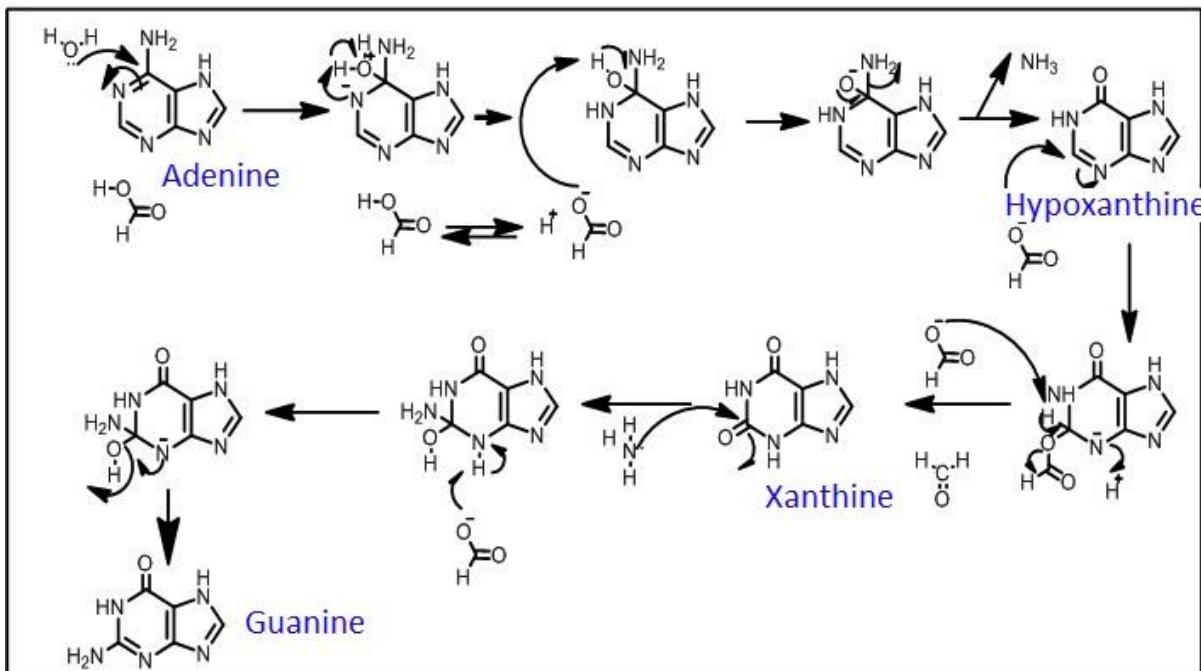



Figure S1. Probable chemical reactions causing transitions of C→U and A→G, to account for the accelerated rate of point mutations in the SARS-CoV-2 genome (besides base changes due to replication errors). **(a)** If the RNA within the virus is subjected to heat treatment (incubated for some time at 95°C) then sulfonated by sodium sulfite (sulfite, including sodium sulfite and sodium metabisulfite are used as disinfectant, antioxidant and preservative agent), at pH 7 and 65°C, the next step of hydrolytic deamination can occur at high pH and room temperature. In the final step, alkali desulfonation (alkaline brines and marinades are frequently employed in Chinese cuisines to alter the texture of meat or seafood) can take place to obtain uracil (for details see Hyatsu, 2008). **(b)** A transition route from A to G by a hydrolytic deamination method using water and formic acid medium (formic acid is a tried and tested biocide, and used in cleaner formulations). The conversion consists of three steps. First of all, the hydrolytic deamination of adenine to form the intermediate compound hypoxanthine. The second step is the formation of xanthine from hypoxanthine by oxidation method, and the third is the amination of xanthine to guanine base using ammonia and water (for details see Tolosa et al., 2019).

References used in the legend of Figure S1

Hayatsu, H., 2008. Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis- A personal account. Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. 84: 321-330. DOI: 10.2183/pjab/84.321.

Tolosa, S., Sanson, J.A., Hidalgo, A., 2019. Theoretical study of adenine to guanine transition assisted by water and formic acid using steered molecular dynamic simulations. Front. Chem. 7: 414. http://doi.org/10.3389/fchem z019.00414.