

CORONAVIRUS

Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India

Mahesh S. Dhar^{1†}, Robin Marwal^{1†}, Radhakrishnan VS^{1†}, Kalaiarasan Ponnusamy^{1†}, Bani Jolly^{2,3†}, Rahul C. Bhojar^{2,3}, Viren Sardana^{2,3}, Salwa Naushin^{2,3}, Mercy Rophina^{2,3}, Thomas A. Mellan⁴, Swapnil Mishra⁴, Charles Whittaker⁴, Saman Fatimi^{2,3}, Meena Datta¹, Priyanka Singh¹, Uma Sharma¹, Rajat Ujjainiya^{2,3}, Nitin Bhatheja², Mohit Kumar Divakar^{2,3}, Manoj K. Singh¹, Mohamed Imran^{2,3}, Vigneshwar Senthivel^{2,3}, Ranjeet Maurya^{2,3}, Neha Jha², Priyanka Mehta², Vivekanand A^{2,3}, Pooja Sharma^{2,3}, Arvinden VR^{2,3}, Urmila Chaudhary¹, Namita Soni¹, Lipi Thukral^{2,3}, Seth Flaxman⁵, Samir Bhatt^{4,6}, Rajesh Pandey^{2,3}, Debasis Dash^{2,3}, Mohammed Faruq^{2,3}, Hemlata Lall¹, Hema Gogia¹, Preeti Madan¹, Sanket Kulkarni¹, Himanshu Chauhan¹, Shantanu Sengupta^{2,3}, Sandhya Kabra¹, The Indian SARS-CoV-2 Genomics Consortium (INSACOG)[†], Ravindra K. Gupta^{7,8}, Sujeet K. Singh¹, Anurag Agrawal^{2,3*}, Partha Rakshit^{1*}

Delhi, the national capital of India, experienced multiple severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) outbreaks in 2020 and reached population seropositivity of >50% by 2021. During April 2021, the city became overwhelmed by COVID-19 cases and fatalities, as a new variant, B.1.617.2 (Delta), replaced B.1.1.7 (Alpha). A Bayesian model explains the growth advantage of Delta through a combination of increased transmissibility and reduced sensitivity to immune responses generated against earlier variants (median estimates: 1.5-fold greater transmissibility and 20% reduction in sensitivity). Seropositivity of an employee and family cohort increased from 42% to 87.5% between March and July 2021, with 27% reinfections, as judged by increased antibody concentration after a previous decline. The likely high transmissibility and partial evasion of immunity by the Delta variant contributed to an overwhelming surge in Delhi.

After escaping relatively unscathed during the first wave of the COVID-19 pandemic, India witnessed a ferocious second COVID-19 wave starting in March 2021 that accounted for about half of global cases by the first week of May. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had spread widely throughout India during the first wave, with the third national serosurvey in January 2021 finding that 21.4% of adults and 25.3% of 10- to 17-year-old adolescents were seropositive (1). Delhi, the national capital, was not included in the national serosurvey but had undergone multiple periods of high transmission in 2020 (Fig. 1A). In a district-wise stratified serosurvey conducted by the Delhi government in January 2021, overall seropositivity was reported to be 56.1% [95% confidence interval (CI), 55.5 to 56.8%],

ranging from 49.1 to 62.2% across 11 districts (2). This level of seropositivity was expected to confer some protection against future outbreaks.

Despite high seropositivity, Delhi was among the most-affected cities during the second wave. The rise in new cases was exceptionally rapid in April 2021, increasing from ~2000 to 20,000 between 31 March and 16 April. This was accompanied by a rapid rise in hospitalizations and intensive care unit (ICU) admissions (Fig. 1B). In this emergency situation with saturated bed occupancy by 12 April, major private hospitals were declared by the state as COVID-care-only facilities, and senior medical students, including those from branches of alternative medicine, were pressed into service (3). Deaths rose proportionately (Fig. 1C), and the case fatality ratio (CFR), estimated as the scaling factor between time-advanced cases and deaths (Fig. 1D), was stable (mean, SD; 1.9, 0.3%). Population spread of SARS-CoV-2 is underestimated by test-positive cases alone (1, 2). To better understand the degree of spread and the factors leading to the unexpectedly severe outbreak, we used all available data, including testing, sequencing, serosurveys, and serially followed cohorts.

In the absence of finely resolved or serial data from national and state surveys, we focused on data for Delhi participants of a national serosurvey of the Council of Scientific and Industrial Research (CSIR, India) employees and their family members (Fig. 2A and table S1). Samples were initially collected

from the end of July to mid-September 2020 (phase I). Subsequently, second and third surveys were done in January and February (phase II) and the end of May to early July 2021 (phase III), bracketing the time period of interest. The cohort details and serosurvey methodology have been previously published (4).

Infection was determined by anti-nucleocapsid assay, which is not affected by immunization with spike protein-based vaccines. The presence of neutralizing antibodies to wild-type SARS-CoV-2 spike protein was estimated by a surrogate viral neutralization test (sVNT, Genscript). Previous results from the full cohort have been comparable to government serosurveys, but Delhi cohort values have been slightly lower. This may be due to an overrepresentation of members with the ability to reduce exposure and avoid public transport (Fig. 2A). Within these limitations, the Delhi cohort showed a rise in seropositivity from 14.7% in phase I (95% CI, 12.6 to 17.0%) to 42.1% in phase II (95% CI, 39.0 to 45.2%). About one-third of neutralizing antibody-positive subjects at phase I became neutralizing antibody-negative by phase II, with most showing declining inhibition on sVNT assays (Fig. 2B) (5). Phase III seropositivity increased to 87.5% (95% CI, 85.0 to 89.7%) among unvaccinated subjects. New infections between March and June 2021 are thus likely to have vastly exceeded known cases. Among 91 previously infected subjects with serial measurements at three time points, including phase III (T3), 25 (27.5%; 95% CI, 18.4 to 37.5%) had a pattern of declining antibody concentration between T1 and T2, followed by a sharp rise at T3, indicative of reinfection (Fig. 2C). Confirmation of reinfection by either reverse transcription polymerase chain reaction (RT-PCR) ($n = 8$) or symptomatic illness ($n = 2$) was available for 10 of the 25 subjects. No severe illness or hospitalization was reported in reinfections.

Time periods of increased transmission were associated with declining RT-PCR cycle threshold (Ct) values (Fig. 3A and fig. S1), attributable to a higher proportion of recently infected individuals with high viral loads (6). However, the Ct decline was far greater in April 2021 (dCt, SE: -4.06, 0.27; $P < 0.001$) than seen previously (November 2020 dCt: ~-1.5). Comparing April 2021 with November 2020, high viral load samples (Ct < 20) doubled in clinical samples ($P < 0.001$) and nearly doubled in campus surveillance testing data, where most positives were from recently infected individuals [15% ($n = 297$) versus 9% ($n = 358$); $P = 0.02$].

Genome analysis trends, in representative samples drawn from the general population over the same period, showed seeding and expansion of B.1.1.7 (Alpha), B.1.617.1 (Kappa), and B.1.617.2 (Delta) lineages in 2021, with

¹National Centre for Disease Control, Delhi, India. ²CSIR-Institute of Genomics and Integrative Biology, New Delhi, India.

³Academy for Scientific and Innovative Research, Ghaziabad, India.

⁴Medical Research Council (MRC) Centre for Global Infectious Disease Analysis, Jameel Institute, School of Public Health, Imperial College London, London, UK. ⁵Department of Mathematics, Imperial College London, London, UK.

⁶Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.

⁷Department of Medicine, Cambridge Institute of Therapeutic Immunology and Infectious Disease (CITIID), University of Cambridge, Cambridge, UK. ⁸Africa Health Research Institute, KwaZulu-Natal, South Africa.

*Corresponding author. Email: partho_rakshit@yahoo.com (P.R.); a.agrawal@igib.in (A.A.)

†These authors contributed equally to this work.

‡The Indian SARS-CoV-2 Genomics Consortium (INSACOG) collaborators and affiliations are listed in the supplementary materials.

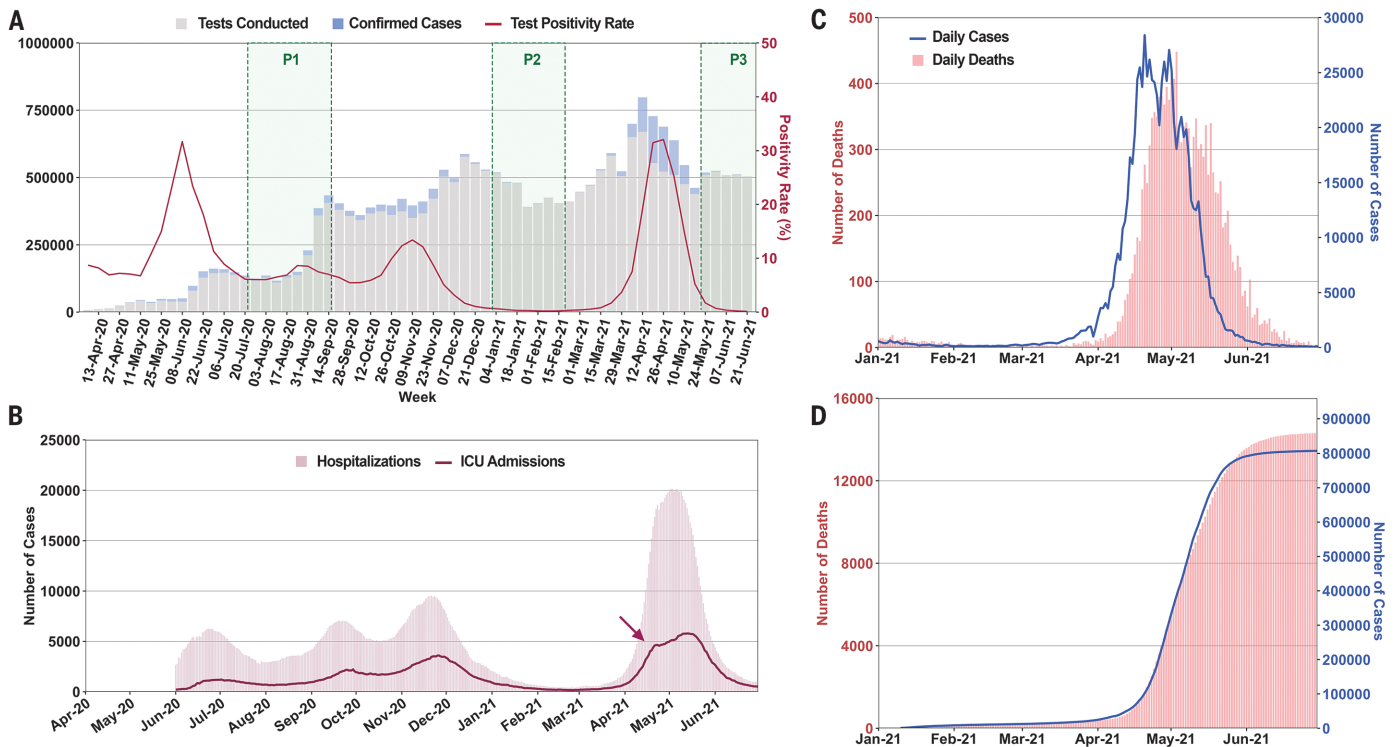


Fig. 1. Multiple surges of SARS-CoV-2 infections in Delhi with an overwhelming outbreak in April and May 2021. (A) Weekly tests, confirmed cases, and test positivity rate in Delhi from April 2020 to June 2021. Sample collection period for CSIR serosurveys is marked as P1 to P3. (B) Number of hospitalizations and ICU admissions plotted on a daily basis from June 2020 to June 2021. The arrow marks the possible saturation of ICU

capacity (3). (C) Daily cases and daily deaths from January 2021 to June 2021. (D) Time-advanced and scaled cumulative cases, fitted to cumulative deaths. Time advancement of cumulative reported cases by 8 days was done for maximal coincidence with scaled cumulative deaths. Case fatality ratio = averaged scaling factor (cumulative deaths / time-advanced cumulative cases). Mean \pm SD, 0.019 \pm 0.003.

Delta becoming the dominant lineage in Delhi during April (Fig. 3B). The proportion of the Delta variant was strongly correlated to the rise in cases and health care stress (fig. S2). Overall, the genomic and epidemiological data were most consistent with the hypothesis that a new variant with higher infectivity, Delta, was driving the unexpected overwhelming surge in Delhi. Recent *in vitro* data supports the possibility of a higher replication rate for Delta, thereby explaining potentially higher viral loads in RT-PCR data and greater transmissibility (7).

We further investigated the sequence of seeding and spread of the Alpha, Kappa, and Delta variants of SARS-CoV-2. Phylogenetic analysis showed common origins between the Alpha variants in Delhi and Punjab, and between the Kappa or Delta variants in Delhi and Maharashtra, where Kappa and Delta were first sequenced (Fig. 3C and fig. S3). There was a recurring pattern of initial smaller outbreaks with Alpha, followed by larger outbreaks coinciding with Alpha-to-Delta transition across all neighboring states (fig. S4). The substantial relative growth advantage of Delta was explored in terms of transmissibility and/or immune escape. The rise of Delta, but not other

lineages, was temporally coincident with a rise in the test positivity rate (TPR) and new cases during the surge (fig. S5). While overall vaccination levels were only about 5% in Delhi, most health care workers had received one or two doses of ChAdOx1-nCov19 (AstraZeneca/Serum Institute, India) or BBV152 (Bharat Biotech, India) (8, 9). We sequenced 24 breakthrough infections starting at least 1 week after the first dose, collected between 22 March and 28 April 2021 at the National Centre for Disease Control (NCDC). The ratio of Delta to non-Delta lineages was 850:1211 from 20 March to 30 April. In contrast, the ratio of Delta to non-Delta lineages was 13:3 in 16 breakthroughs after one dose and 7:1 in eight breakthroughs after the second dose of vaccine. Although the small sample size and lack of a formal control group preclude definitive analysis, estimated higher odds for Delta in vaccination breakthroughs (odds ratio, 7.1; 95% CI, 2.4 to 20.9) corroborate other reports of reduced vaccine effectiveness against Delta (10).

To better characterize how the properties of Delta might differ from other SARS-CoV-2 lineages previously circulating in the city, we used a Bayesian model of SARS-CoV-2 transmission and mortality that simultaneously

models the dynamics of two categories of virus (B.1.617.2 and non-B.1.617.2) (11), while also explicitly incorporating natural waning of immunity derived from prior infection, with the duration of immunity consistent with the results of recent longitudinal cohort studies (12, 13). Details of the model and input data are given in the supplementary methods. Briefly, the model is fitted to COVID-19 mortality data, genomic sequence data presented here and from GISAID [with Phylogenetic Assignment of Named Global Outbreak Lineages (PANGOLIN) classification] (14–16), and serological data presented alongside an additional longitudinal serosurvey carried out in the city from July to December 2020 (17). Substantial uncertainties remain as to the date of introduction of B.1.617.2 into Delhi and the degree of COVID-19 death underascertainment. We therefore explored a range of different scenarios in which we varied underascertainment (10, 33, 50, and 66%) and introduction dates (15 January 2021, 31 January 2021, 14 February 2021, and 28 February 2021).

Using this framework for an introduction date of 14 February 2021 and mortality underascertainment of 50%, our results, shown in

Fig. 2. Serological estimates of prior infections, preexisting immunity, and new infections for the April and May 2021 outbreak.

(A) Seropositivity in the CSIR cohort, subdivided by nature of employment and use of public transport, plotted for different time periods (phase I to phase III, proportion \pm 95% CI). Details are provided in table S1. (B) Variability and temporal decline in neutralization capacity estimated by sVNT assay between phases I and II ($n = 52$ subjects). (C) Serial antibody concentration measurements in initially seropositive subjects ($n = 91$). Pattern suggestive of reinfections is shown (decline followed by rise, $n = 25$). (D) Remaining data ($n = 66$ subjects), with four indeterminate reinfection cases indicated with arrows. Antibody concentration is reported in multiples of the assay cutoff index value (Col).

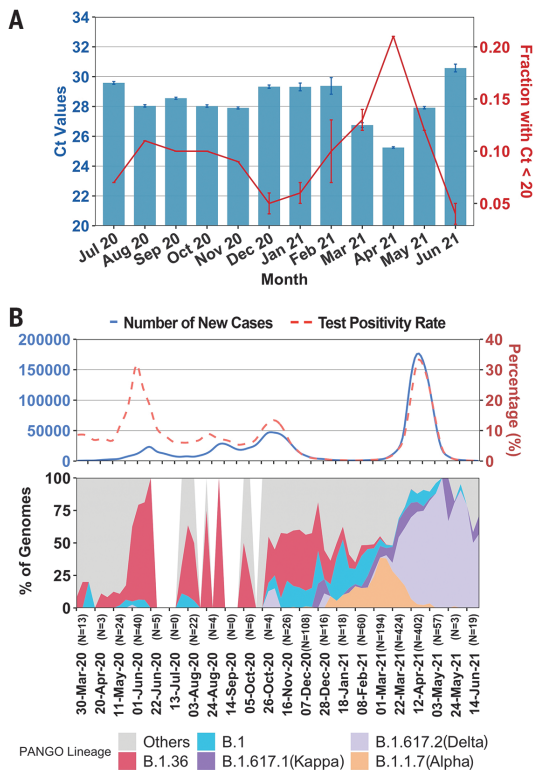
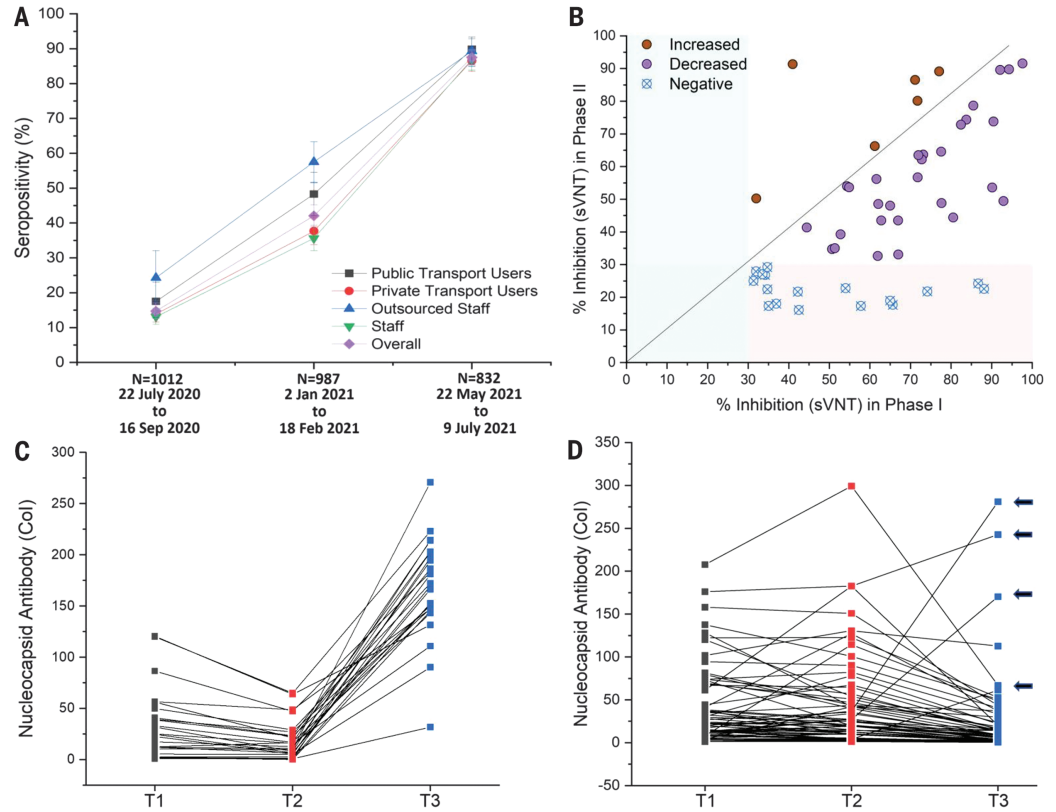
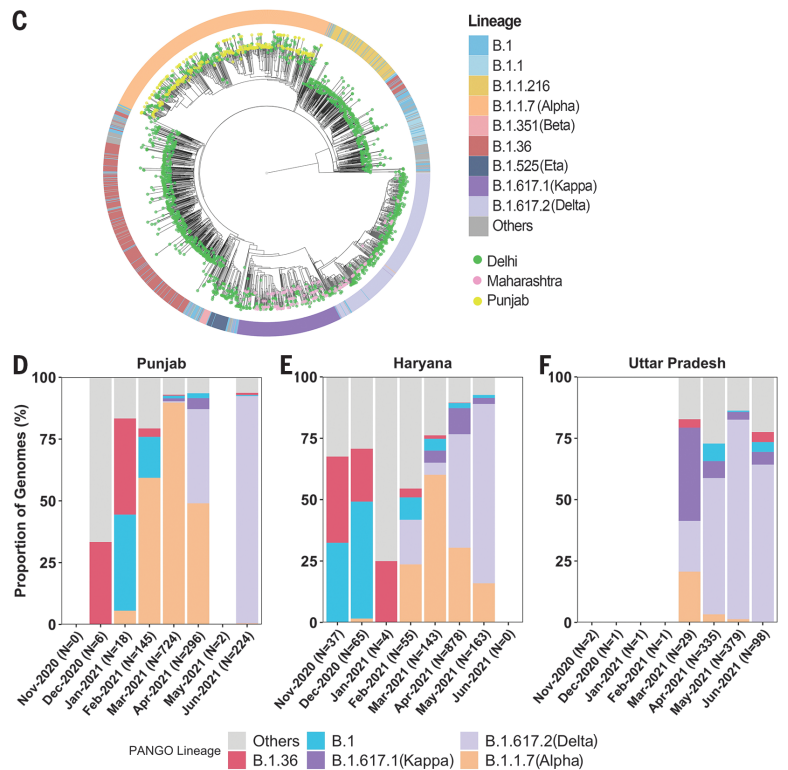


Fig. 3. Genomic-epidemiologic correlations. (A) Time trends of Ct values (mean \pm SE) and high viral load samples (proportion \pm SE) for *Orf1* gene (E gene data, fig. S1). (B) Smoothed graph of main lineages in Delhi from March 2020 to May 2021 in biweekly increments. New cases and TPR are aligned and plotted on the same timeline. (C) Phylogenetic analysis for variant of concern (VOC) strains



between Delhi and states (Punjab and Maharashtra) with known VOC outbreaks before April 2021. Further analysis suggesting a superspreading event for Alpha is shown in fig. S3. (D to F) Month-wise proportions of different lineages ($n > 3$ sequenced samples) in states surrounding Delhi. Additional data are shown in figs. S4 and S5.

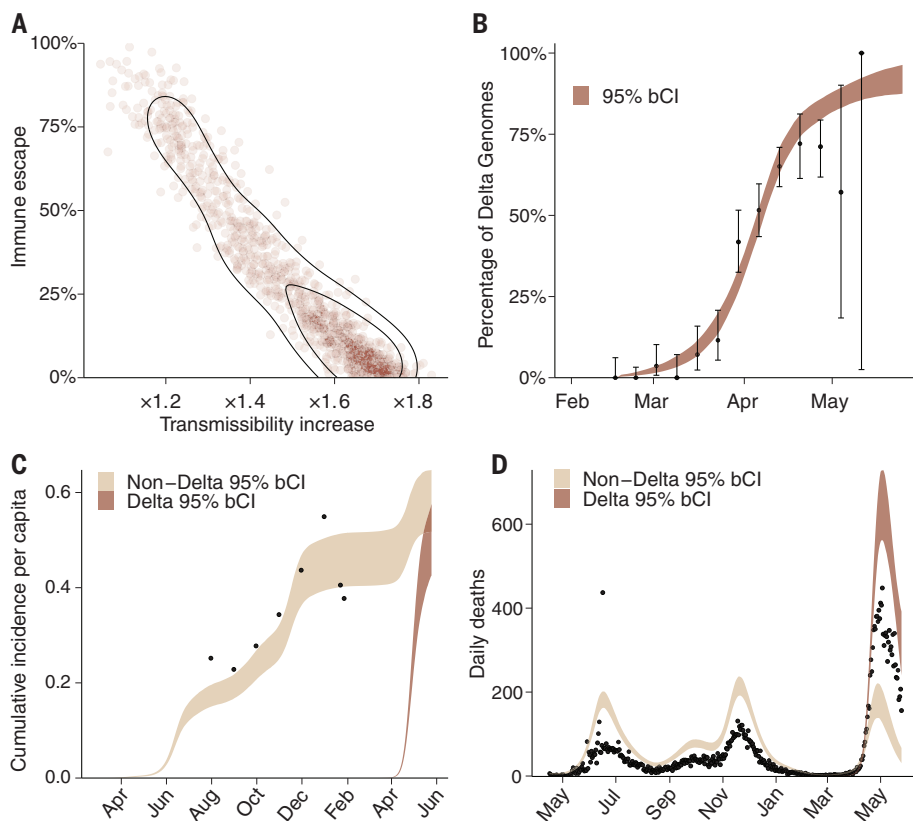


Fig. 4. Estimates of the epidemiological characteristics of the Delta variant. Values were inferred from a two-category Bayesian transmission model fitted to mortality, serosurvey, and genomic data from Delhi, India. **(A)** Joint posterior distribution, with isoclines corresponding to the 90% and 50% enclosures of posterior density of the Delta variant immune escape and transmissibility increase relative to non-Delta categories. Immune escape has a median of 20% with 50% Bayesian credible interval (bCI) of 10 to 50%, and transmissibility has a median increase of 1.5 with 50% bCI of 1.3 to 1.7. **(B)** Delta fraction over time, inferred by the model. Black dots represent genome sampling data points, with exact binomial confidence intervals. **(C)** Serosurvey data (black dots) and inferred cumulative incidence for Delta and non-Delta variant categories. **(D)** Mortality data (black dots) and inferred deaths assuming 50% underreporting. Other underascertainment scenarios are presented in the supplementary materials.

Fig. 4, indicate that the Delta variant is 1.3- to 1.7-fold [50% Bayesian credible interval (bCI)] more transmissible than earlier and co-circulating SARS-CoV-2 lineages in Delhi, including the Alpha variant. Notably, the model also indicates that the Delta variant can partially evade immunity elicited by prior infection, with prior infection providing only 50 to 90% (50% bCI) of the protection against infection with the Delta variant that it provides against previous lineages. There is an inherent trade-off between transmissibility and immune escape, and the worst-case scenario of both very high transmissibility and immune escape is rejected a posteriori by the data. Figure 4A also highlights the nature of uncertainty in the exact level of immune escape and transmissibility increase, because these inferred characteristics of the Delta variant are collinear given the modeling framework used and data currently available. The

main limitations of the model are due to biases in the data and our choice of priors. For example, there is an unknown degree of underreporting, and serological estimates are likely to be systematically biased. Similarly, there are uncertainties in estimates of the temporal waning of immunity, the date of the first introduction of Delta, and the true infection fatality rate of SARS-CoV-2 variants in Delhi. We explicitly consider these limitations and partially mitigate such biases, as described in the supplementary materials. Overall, the main findings are robust to variation in prior assumptions, including both changing underascertainment and the date of introduction (supplementary methods and tables S2 and S3). The results are valid for a population where most of the immunity arose from prior infection (rather than vaccination), which is true for Delhi. On the basis of median estimates of the model (Fig. 4A)

and high transmissibility of the background Alpha lineage (18), Delta should potentially be at least twice as transmissible as the wild-type lineage.

We note that after the massive Delta outbreak, new cases in Delhi and other North Indian states have stayed extremely low, with TPR in Delhi below 0.1% as of September 2021. This fits the serological picture presented in Fig. 2, with a very high fraction of the population being recently infected and with good immunity to Delta. However, Delta outbreaks have continued in parts of India and elsewhere in the world, despite moderately high seropositivity or vaccination levels that were previously considered to be adequate (7, 19, 20). We conclude that the Delta variant is capable of initiating fast-rising outbreaks in populations with immune responses to prior variants, resulting in reinfections and vaccination breakthroughs. Public health strategies may need to be revised to account for variants with heightened transmissibility and immune escape.

REFERENCES AND NOTES

- M. V. Murhekar *et al.*, *Int. J. Infect. Dis.* **108**, 145–155 (2021).
- S. Jain (@SatyendarJain), Twitter post, 2 February 2021, 5:19 a.m.; <https://twitter.com/satyendarjain/status/135654781742726624?lang=en>.
- Government of National Capital Territory of Delhi, Orders F.23/Misc/COVID-19/DGHS/NHC/2020/Pt XVIII/90-97 and No.01/H&FW/COVID-Cell/2020/07/ss4hfw/726, 12 April 2021; http://health.delhigovt.nic.in/wps/wcm/connect/doit_health/Health/Home/Covid19/Covid+19+Related+order+April+2021.
- S. Naushin *et al.*, *eLife* **10**, e66537 (2021).
- J. M. Dan *et al.*, *Science* **371**, eabf4063 (2021).
- J. A. Hay *et al.*, *Science* **373**, eabh0635 (2021).
- P. Mlicochova *et al.*, *Nature* **10.1038/s41586-021-03944-y** (2021).
- P. M. Folegatti *et al.*, *Lancet* **396**, 467–478 (2020).
- R. Ella *et al.*, medRxiv 2021.06.30.21259439 [Preprint] (2021). <https://doi.org/10.1101/2021.06.30.21259439>.
- J. Lopez Bernal *et al.*, *N. Engl. J. Med.* **385**, 585–594 (2021).
- S. Flaxman *et al.*, *Nature* **584**, 257–261 (2020).
- C. H. Hansen, D. Michlmayr, S. M. Gubbels, K. Mølbak, S. Ethelberg, *Lancet* **397**, 1204–1212 (2021).
- V. J. Hall *et al.*, *Lancet* **397**, 1459–1469 (2021).
- Y. Shu, J. McCauley, *Euro Surveill.* **22**, 30494 (2017).
- A. Rambaut *et al.*, *Nat. Microbiol.* **5**, 1403–1407 (2020).
- Á. O'Toole *et al.*, *Virus Evol.* **7**, veab064 (2021).
- A. Velumani *et al.*, medRxiv 2021.03.19.21253429 [Preprint] (2021). <https://doi.org/10.1101/2021.03.19.21253429>.
- N. G. Davies *et al.*, *Science* **372**, eabg3055 (2021).
- D. Planas *et al.*, *Nature* **596**, 276–280 (2021).
- S. Mishra *et al.*, *EclinicalMedicine* **39**, 101064 (2021).
- banijolly/ncov-Delhi-Epidemiology: nCoV Delhi-Epidemiology, version 1.0, Zenodo (2021); <https://doi.org/10.5281/zenodo.5521073>.
- T. Mellan, M. Swapnil, C. Whittaker, S. Bhatt, S. Flaxman, ImperialCollegeLondon/Delta_Variant_Delhi: Genomic characterization and Epidemiology of an emerging SARS-CoV-2 variant in Delhi, India, version 1.0, Zenodo (2021); <https://doi.org/10.5281/zenodo.5524442>.

ACKNOWLEDGMENTS

Suggestions by N. K. Arora, G. Padmanabhan, and members of the Scientific and Clinical Advisory Group of INSACOG are

acknowledged. Coordinators of CSIR labs and centers are acknowledged for CSIR-Cohort. We gratefully acknowledge the authors from the originating laboratories and the submitting laboratories, who generated and shared via GISAID genetic sequence data on which this research is based. A full list acknowledging GISAID contributors for the data used in this study can be found in data S3. **Funding:** Support from the Ministry of Health and Family Welfare, the Council of Scientific and Industrial Research (CSIR-Phenome India Grant, MLP-2007; InGen-CoV2, MLP-2005), and the Department of Biotechnology (INSACOG), India, is acknowledged. **Author contributions:** Conceived of the study: A.A., P.R., S.K.S., R.K.G., S.Ka., S.S., M.F., D.D., R.P., S.B., S.Fl., R.C.B., and B.J. Designed the study and experiments: A.A., P.R., S.K.S., R.K.G., S.Ka., S.S., M.F., R.P., U.S., P.Si., M.D., R.C.B., K.P., R.V.S., R.Mar., and M.S.D. Performed experiments: M.S.D., R.Mar., R.V.S., R.C.B., S.N., M.D., P.Si., U.S., R.U., M.K.D., M.K.S., M.I., V.Se., and P.Sh. Collected and analyzed patient data: M.S.D., R.Mar., R.V.S., K.P., M.D., P.Si., U.S., M.K.S., U.C., N.S., H.L., H.G., P.Ma., S.Ku., H.C., and P.R. Performed bioinformatics analyses: M.S.D., R.Mar., R.V.S., K.P., B.J., S.Fa., N.B., R.Mau., N.J., P.Me., V.A., A.V.R., and L.T.

Performed statistical analyses: A.A., V.Sa., T.A.M., S.M., C.W., S.Fl., S.B., and D.D. Interpreted data: A.A., P.R., M.S.D., R.Mar., R.V.S., K.P., B.J., V.Sa., S.N., M.R., T.A.M., S.M., C.W., R.U., N.B., L.T., S.Fl., S.B., R.P., D.D., M.F., S.S., S.Ka., R.K.G., and S.K.S.

Competing interests: The authors declare that they have no competing interests. **Data and materials availability:** All data used in the analysis are accessible in the supplementary materials as data files or in Zenodo (21), including accession numbers for all genomic data that have been deposited at GISAID. Code and data for replication of the Bayesian inference are available in Zenodo (22). Biological samples from infectious patients and patient details were collected for the purpose of public health and cannot be shared for reasons of confidentiality and to comply with Indian government regulations. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a

third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abj9932](https://doi.org/10.1126/science.abj9932)
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S3
List of INSACOG Consortium Members
References (23–42)
MDAR Reproducibility Checklist
Data S1 to S9

[View/request a protocol for this paper from Bio-protocol.](#)

15 June 2021; accepted 6 October 2021
Published online 14 October 2021
10.1126/science.abj9932