

## RESEARCH ARTICLE

# Contact-number-driven virus evolution: A multi-level modeling framework for the evolution of acute or persistent RNA virus infection

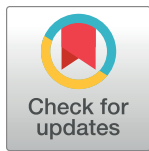
Junya Sunagawa<sup>1</sup>, Ryo Komorizono<sup>2</sup>, Hyeongki Park<sup>3</sup>, William S. Hart<sup>4</sup>, Robin N. Thompson<sup>5,6</sup>, Akiko Makino<sup>2,7</sup>, Keizo Tomonaga<sup>2,7,8†</sup>, Shingo Iwami<sup>3,9,10,11,12,13‡</sup>, Ryo Yamaguchi<sup>1,14\*</sup>

**1** Department of Advanced Transdisciplinary Science, Hokkaido University, Sapporo, Hokkaido, Japan, **2** Laboratory of RNA Viruses, Department of Virus Research, Institute for Life and Medical Sciences (LiMe), Kyoto University, Kyoto, Japan, **3** interdisciplinary Biology Laboratory (iBLab), Division of Natural Science, Graduate School of Science, Nagoya University, Nagoya, Japan, **4** Mathematical Institute, University of Oxford, Oxford, United Kingdom, **5** Mathematics Institute, University of Warwick, Coventry, United Kingdom, **6** Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry, United Kingdom, **7** Laboratory of RNA Viruses, Graduate School of Biostudies, Kyoto University, Kyoto, Japan, **8** Department of Molecular Virology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **9** Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan, **10** Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto, Japan, **11** Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Saitama, Japan, **12** NEXT-Ganken Program, Japanese Foundation for Cancer Research (JFCR), Tokyo, Japan, **13** Science Groove Inc., Fukuoka, Japan, **14** Department of Zoology & Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

☞ These authors contributed equally to this work.

‡ KT and SI also contributed equally to this work.

\* iwami.iblab@bio.nagoya-u.ac.jp (SI); ryamaguchi@sci.hokudai.ac.jp (RY)



## OPEN ACCESS

**Citation:** Sunagawa J, Komorizono R, Park H, Hart WS, Thompson RN, Makino A, et al. (2023) Contact-number-driven virus evolution: A multi-level modeling framework for the evolution of acute or persistent RNA virus infection. PLoS Comput Biol 19(5): e1011173. <https://doi.org/10.1371/journal.pcbi.1011173>

**Editor:** Joel O. Wertheim, University of California San Diego, UNITED STATES

**Received:** September 29, 2022

**Accepted:** May 10, 2023

**Published:** May 30, 2023

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1011173>

**Copyright:** © 2023 Sunagawa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All codes necessary to repeat the simulation described in this study have been made available. Python source codes

## Abstract

Viruses evolve in infected host populations, and host population dynamics affect viral evolution. RNA viruses with a short duration of infection and a high peak viral load, such as SARS-CoV-2, are maintained in human populations. By contrast, RNA viruses characterized by a long infection duration and a low peak viral load (e.g., borna disease virus) can be maintained in nonhuman populations, and the process of the evolution of persistent viruses has rarely been explored. Here, using a multi-level modeling approach including both individual-level virus infection dynamics and population-scale transmission, we consider virus evolution based on the host environment, specifically, the effect of the contact history of infected hosts. We found that, with a highly dense contact history, viruses with a high virus production rate but low accuracy are likely to be optimal, resulting in a short infectious period with a high peak viral load. In contrast, with a low-density contact history, viral evolution is toward low virus production but high accuracy, resulting in long infection durations with low peak viral load. Our study sheds light on the origin of persistent viruses and why acute viral infections but not persistent virus infection tends to prevail in human society.

are hosted on Github ([https://github.com/j-s9vcp/ContactNumberDriven\\_VirusEvolution](https://github.com/j-s9vcp/ContactNumberDriven_VirusEvolution)) [69]. There are no data to be archived.

**Funding:** This study was supported in part by Grants-in-Aid for JSPS Scientific Research (KAKENHI) Scientific Research B 18KT0018 (to S.I.), 18H01139 (to S.I.), 16H04845 (to S.I.), Scientific Research in Innovative Areas 20H05042 (to S.I.), 19H04839 (to S.I.), 18H05103 (to S.I.), 21K15160 (to R.Y.), JP20H05682 (to K.T.), JP21K19909 (to K.T.); JSPS Overseas Research Fellowships (to R.Y.); ACT-X JPMJAX22AK (to R.Y.); AMED CREST 19gm1310002 (to S.I.); AMED Research Program on HIV/AIDS 19fk0410023s0101 (to S.I.); AMED Japan Program for Infectious Diseases Research and Infrastructure, 20wm0325004s0201, 20wm0325012s0301, 20wm0325015s0301 (to S.I.); AMED Research Program on Emerging and Re-emerging Infectious Diseases 19fk0108156h0001, 20fk0108140s0801 and 20fk0108413s0301 (to S.I.); AMED Program for Basic and Clinical Research on Hepatitis 19fk0210036h0502 (to S.I.); AMED Program on the Innovative Development and the Application of New Drugs for Hepatitis B 19fk0310114h0103 (to S.I.); Moonshot R&D Grant Number JPMJMS2021 (to S.I.) and JPMJMS2025 (to S.I.); JST MIRAI (to S.I.); Mitsui Life Social Welfare Foundation (to S.I.); Shin-Nihon of Advanced Medical Research (to S.I.); Suzuken Memorial Foundation (to S.I.); Life Science Foundation of Japan (to S.I.); SECOM Science and Technology Foundation (to S.I.); The Japan Prize Foundation (to S.I.), and Daiwa Securities Health Foundation (to S.I.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

As exemplified by the SARS-CoV-2 variants of concern and influenza A virus variants, we need to predict the future evolution of viral properties to counter an oncoming pandemic. While some RNA viruses such as influenza A viruses adopt an acute infection strategy, others, such as Borna disease virus, adopt a persistent infection strategy. In what kind of environment did these viruses evolve? This study presents a modeling framework to investigate the evolution of proliferative ability and accuracy throughout the viral life cycle. Specifically, we calculated eco-evolutionary dynamics from virus infection in the host to transmission between hosts and capture the transmission potential of the virus. The acute infection phenotype evolves under a contact history involving frequent contacts between hosts. By contrast, the persistent infection phenotype evolves under an environment with a small mean and a large variance in the host contact history. These findings have direct implications for the fight against the continuous evolution of various viruses, which have easily prevailed in the setting of dense human contacts. We believe it is essential to readdress the evolution of viruses from an evolutionary and ecological perspective.

## Introduction

Because they lack proofreading activity in their RNA-dependent RNA polymerase, RNA viruses exhibit extremely high mutation rates, generating  $10^{-6}$  to  $10^{-4}$  substitutions per nucleotide, orders of magnitude greater than those of most DNA-based life forms [1–3]. The accuracy (or inaccuracy) of the polymerase is the source of intra-host viral diversity and determines viral propagation and virulence *in vivo* [4,5].

In general, acute infectious RNA viruses, such as SARS-CoV-2, are highly proliferative and release large numbers of progeny viral particles [6]. From an epidemiologic point of view, acute infectious RNA viruses are highly pathogenic [7,8]. Infections with most non-retroviral RNA viruses result in the characteristic symptoms and signs of acute illness. On the other hand, persistent viral infection is a different strategy. The properties of RNA viruses vary widely, among which non-retroviral RNA viruses can often be broadly classified according to acute infection and persistent (chronic) infection. Some non-retroviral RNA viruses, such as hepatitis C virus, coxsackievirus, and borna disease virus, establish persistent infection and replicate in the host for long periods [9]. Whereas acute infectious viruses are eliminated by the immune response within a few weeks, persistent infectious viruses are maintained in the host for at least several months or even years. In most cases, patients are mildly symptomatic or asymptomatic during persistent infection. Given the multiple viral and host factors involved in viral pathogenesis, it is not easy to discern the causes of mortality in the course of a viral infection. Many acutely infectious RNA viruses cause fatal disease soon after infection (e.g., rabies virus, Nipah virus), whereas some persistent infectious RNA viruses cause nonfatal disease after infection (e.g., Borna disease virus, Pegi virus) [10]. A persistent infectious virus can coexist with the host by reducing the amount of progeny virus released from the infected cells, slowing down the replication rate, decreasing the evolutionary rate, or evading the host's immune system to use the host as its own reservoir until the next opportunity to infect [11–15]. These different viral phenotypes of acute and persistent infection express different strategies in the host population, ranging from intracellular replication rates and innate immune response to inter-individual transmission dynamics [9].

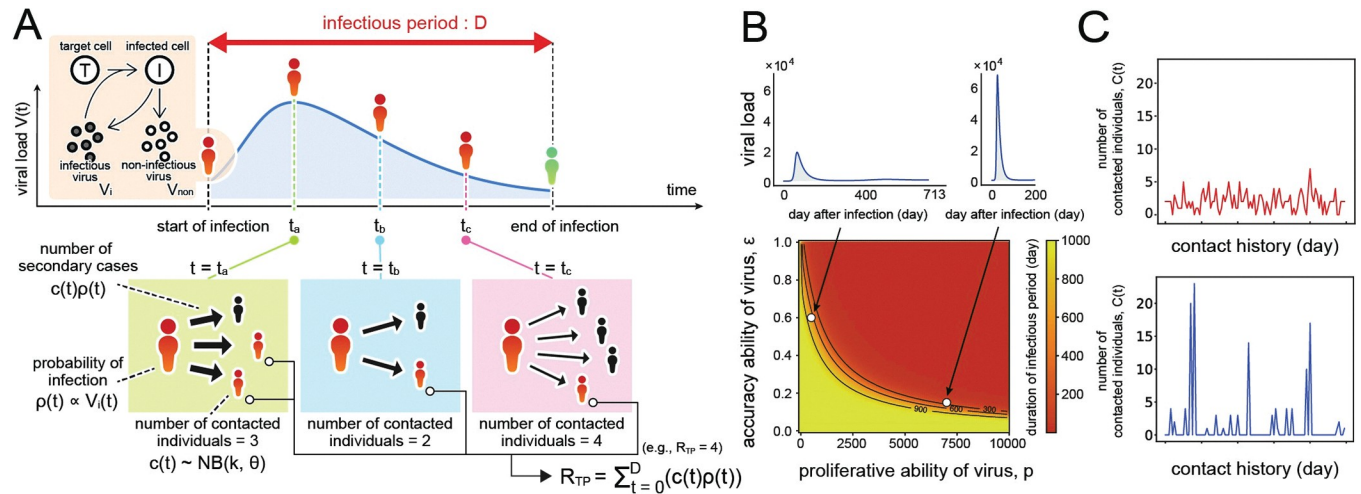
For gaining a better understanding of how viral properties are optimized, the lifestyles of the host organisms are of critical importance. Different host behaviors provide different environments, with some environments favoring particular viral characteristics. Human activities in the Anthropocene have greatly affected the global environment, increasing the frequency of human-animal contact and various human-induced evolutionary changes in animals [16]. Traditionally, it is considered that acute infection evolved because of population-level transmission potential, that is, the potential for many susceptible hosts to come into contact with an infected individual [17]. This suggests the question, “Why did persistent infectious RNA viruses evolve?” Commonly, the population-level transmission potential is characterized by the (basic) reproduction number [8]. However, the environmental conditions in the aspect of contact history that favor the evolution of persistent infectious RNA viruses have not yet been explored. In addition, in terms of the within-host virus life cycle, it is unclear how viral properties such as the proliferation rate and the accuracy of viral replication affect virus evolution under different environmental conditions. Another unresolved issue is why acute infections are generally more prevalent than persistent infections in the current profile of viral diversity in human populations [11].

To tackle these challenges, we employed a multi-level modeling approach including both individual-level virus infection dynamics and population-scale transmission. Because we aimed to seek an optimal strategy across a broad range of viruses rather than explain the properties of a specific virus, we simply assume that an optimization occurs at an epidemiologic equilibrium (i.e., we focus on equilibrium solutions). Although elaborate discussions are made possible by considering the (basic) reproduction number and balance with other high-level population criteria like lifespan of the host in models specific to each virus (e.g., [18,19]; also see [20] as an example for limitations to the idea that natural selection will always maximize the number), we examine the long-term equilibrium evolutionary outcome in the context of stable endemic diseases which solely maximizes “transmission potential” as the reproduction number (i.e., the total number of secondary cases generated throughout the infectious period). Our simplification is valid unless considering the situation where pathogens often showing non-equilibrium dynamics such as antigenic escape from host immune defenses through evolution [21]. Under different environments characterized by different contact numbers in populations, we evaluated how measures of the proliferation rate and accuracy in the viral life cycle affect population-level transmission potential and therefore RNA virus evolution.

## Results

### Virus evolution in different contact scenarios

To explore how virus evolution is influenced by the environment, we first developed a probabilistic multi-level model to characterize population-level transmission potential,  $R_{TP}$ . Using a genetic algorithm (see [Materials and Methods](#) for further details), we investigated the evolution of the proliferation rate and the accuracy of replication in the viral life cycle,  $p$  and  $\epsilon$ , respectively. Our model [Eqs (7–8)] calculates  $R_{TP}$  as the sum of secondary cases generated throughout the infectious period by a primary case (see [Fig 1A](#), [Materials and Methods](#) in detail). Calculating  $R_{TP}$  means that we investigate the virus evolution as a long-term equilibrium (i.e., population-level transmission potential) here rather to explore non-equilibrium co-evolutionary dynamics with host individuals. The duration of the infectious periods  $D(V_{total})$  is assumed to depend on the total viral load. Specifically, the duration of the infectious period decreases with the cumulative viral load. In each parameter combination (i.e.,  $p$  and  $\epsilon$ ),  $D(V_{total})$  is calculated by Eqs (5–6) based on the virus infection model [Eqs (1–4)] with other fixed



**Fig 1. Schematic illustration of multi-level disease transmission.** A schematic of our model is depicted in (A). The probability of infection  $p(t)$  of susceptible individuals from an infected hosts depends on their viral load  $V_i(t)$ . At each time step (i.e., day), the focal infected individual has contacts with multiple susceptible individuals. Here the contact numbers  $C(t)$  are assumed to follow a negative binomial distribution, which does not depend on the viral load. The sum of newly infected individuals (i.e., [number of contacted individuals per day]  $\times$  [probability of infection per contacted individual]) during the infectious period is calculated as  $R_{TP}$ . The duration of the infectious periods for each parameter combination ( $p$  and  $\epsilon$ ) characterizing the proliferative ability and accuracy of viruses are calculated in (B) based on the virus infection model. The left panel corresponds to an infectious period of 713 days with a relatively low viral load, and the right panel represents an infectious period of 200 days with a high viral load. Examples of generated daily contact numbers characterizing environmental conditions are shown in (C). The contact numbers under a negative binomial distribution with  $k=100$  and  $\theta=0.03$  (the mean and variance are 3.0 and 3.1) and  $k=0.12$  and  $\theta=10$  (the mean and variance are 1.2 and 13.3) are shown in the top and bottom panels, respectively.

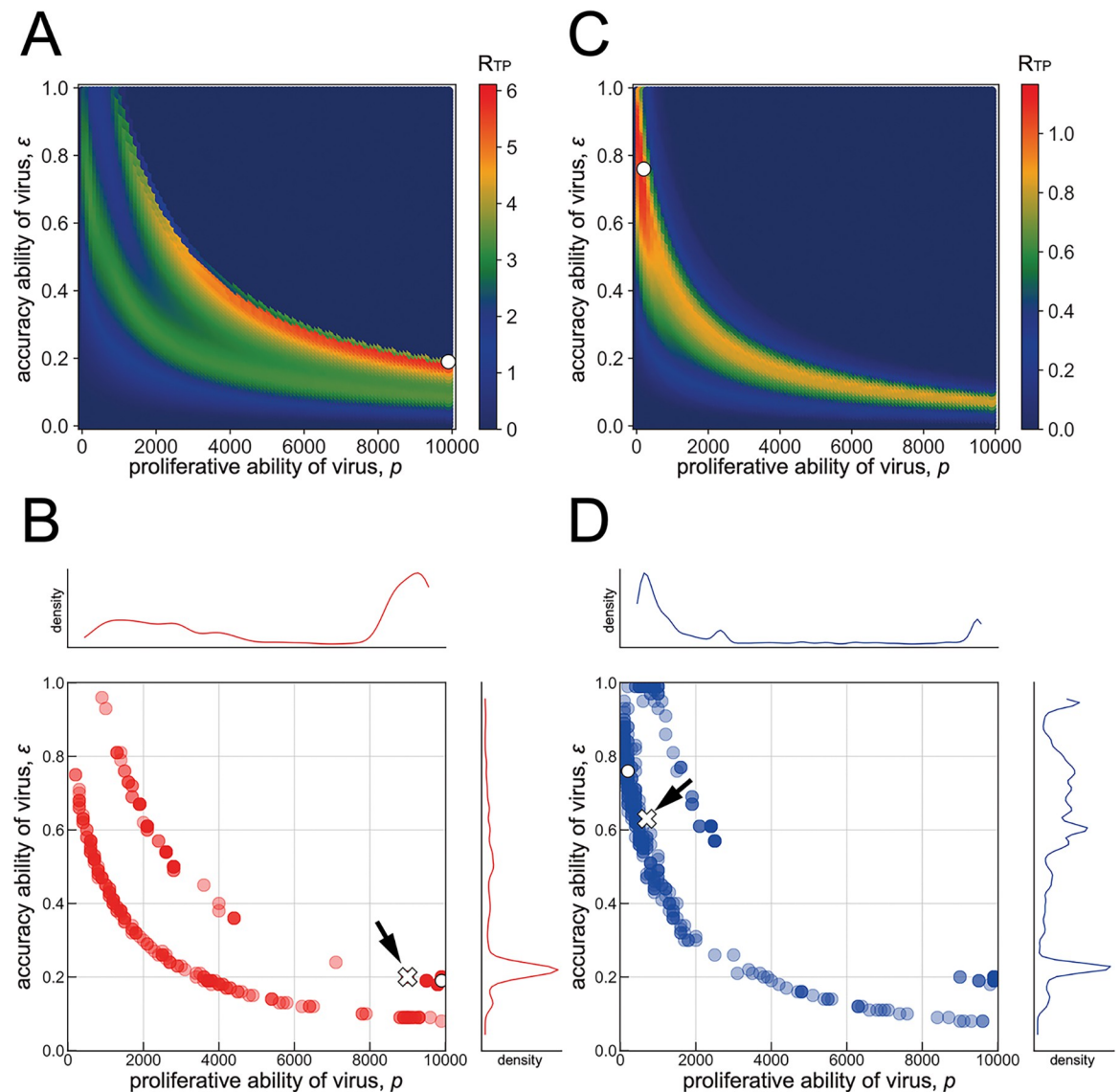
<https://doi.org/10.1371/journal.pcbi.1011173.g001>

parameters (see **Table A in S1 Text**). The calculated infectious period for each parameter pair is described in **Fig 1B**.

To evaluate how the proliferative ability and accuracy of replication of viruses are optimized to increase the transmission potential, we assumed that environments are characterized by the number of contacts per day. We made two categories of contact history with daily numbers of contacts sampled from the negative binomial distribution (see **Materials and Methods** in detail). The first category is one with large mean and small variance, in which the contact number is always around the mean throughout the contact history and does not decrease too low or increase too high (**Fig 1C**, top). The other has small mean and large variance, in which the contact number is usually 0 but is sometimes very high (**Fig 1C**, bottom). For contact numbers with large mean and small variance,  $k=100$  and  $\theta=0.03$  (the mean number is 3) are fixed to generate the contact history. In the history with small mean and large variance,  $k=0.12$  and  $\theta=10$  (the mean number is 1.2). Here  $k$  and  $\theta$  represent the shape parameter and scale parameter of the negative binomial distribution, respectively.

Given a contact history generated from the negative binomial distribution, as described above, we searched for the optimal point in the fitness landscape (i.e., the largest  $R_{TP}$ ), that is, the optimal combination of  $p$  and  $\epsilon$ , by using a genetic algorithm for each contact number (see **Materials and Methods** in detail). For the contact history with large mean and small variance (i.e., high-density contact history:  $(k, \theta) = (100, 0.03)$ , **Fig 1C**, top), optimal fitness involves high virus production but low accuracy, which results in a short duration of infectious period with high peak viral load (we call this the *acute infection* phenotype) (**Figs 2A and 2D in S1 Text**). In contrast, for the contact history with small mean and large variance (i.e., low-density contact history:  $(k, \theta) = (0.12, 10)$ , **Fig 1C**, bottom), the virus is more likely to be optimized toward low virus production but high accuracy, resulting in a long infection duration with low peak viral load (we call this the *persistent infection* phenotype) (**Figs 2C and D in S1 Text**).





**Fig 2. Virus evolution in a given environment.** The optimal parameter combinations ( $p$  and  $\epsilon$ ) that maximize  $R_{TP}$  are calculated for a parameter set of  $k = 100$  and  $\theta = 0.03$  in (A). The white dot represents the optimal point of  $R_{TP}$  in a single simulation. Using the same parameter set, 500 optima based on independent contact histories are calculated in (B). The kernel plots illustrated alongside the top and right parts are the marginal distributions of  $p$  and  $\epsilon$ , respectively. The white dot exactly corresponds to the optimal point shown in a single simulation in (A). The white cross indicated by the black arrow corresponds to the median of optimized parameter set of  $p$  and  $\epsilon$  based on the 500 simulation runs. In a similar manner, the parameter set of  $k = 0.12$  and  $\theta = 10$  is used to calculate the optimal parameter combinations ( $p$  and  $\epsilon$ ) in (C). Using the same parameter sets, 500 evolutionary consequences for independently drawn contact histories are shown in (D). The white dot and white cross correspond to the optimal point of the parameter set and the median point, respectively.

<https://doi.org/10.1371/journal.pcbi.1011173.g002>

The white dots in [Fig 2A and 2C](#) correspond to the coordinates of the combination of  $p$  and  $\epsilon$  that maximize  $R_0$  with a given contact pattern.

With 500 independent iterations of this process with the same parameter sets, the evolutionary endpoint of each iteration is distributed as shown in [Fig 2B and 2D](#). The white dots in [Fig 2B and 2D](#) correspond to the evolutionary endpoint in each single simulation run, [Fig 2A](#) and [Fig 2C](#), respectively. The white crosses indicated by the black arrows in [Fig 2B and 2D](#)

correspond to the median parameter set of  $p$  and  $\varepsilon$  after evolution under the given environment sequence of the contact history.

In addition to these contact histories, we considered the evolution of viral phenotype under other contact histories in scale-free networks and small-world networks. In the analysis on networks, we interpret the evolutionary consequences of the phenotypes similarly. Under networks where hubs exist or where individuals are in contact with each other evenly and with high frequency, the average value of the contact history is stable at a high value, and the acute infection type is likely to be optimal, and vice versa (**Figs E and F in S1 Text**). As described in the next section, we investigated how the trend of virus evolution changes under the different environments.

### Virus evolution depending on the difference in mean contact number

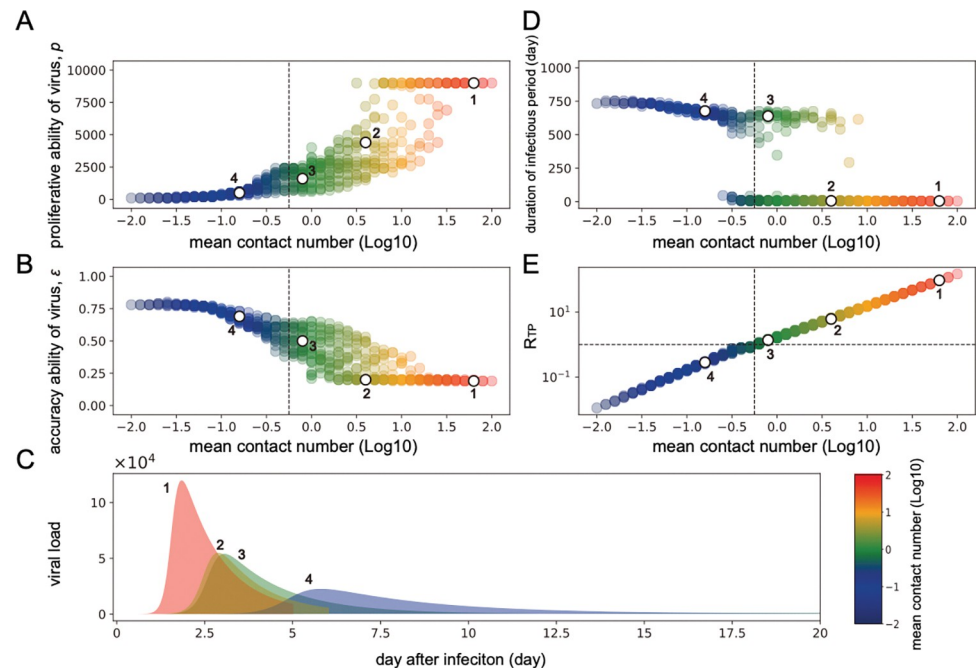
We explored the optimal properties of the virus (proliferation rate and accuracy of replication,  $p$  and  $\varepsilon$ , respectively) for transmission by gradually changing the mean number of daily contacts in the contact history. To represent various contact histories, we explored a total of 441 patterns of contact histories ranging from 0.01 to 100. That is, we used the following parameter sets: shape parameter  $k$  ranges from  $10^{-1.0}$ ,  $10^{-0.9}$ ,  $10^{-0.8}$ , ...,  $10^{0.9}$ ,  $10^{1.0}$  (sum up to 21 points), as is also the case with scale parameter  $\theta$ , and each contact history is made by the combination of each  $k$  and  $\theta$ . We conducted a total of 500 independent iterations for each parameter set of  $(k, \theta)$  and determined the evolutionary consequences of the median parameter set of  $p$  and  $\varepsilon$ .

We found that viruses are optimized toward low production but high accuracy (the persistent infection phenotype) in regions of low-density contact history (the mean contact number from 0.01 to 1) regardless of the variance in daily contacts (white dots 3 and 4 in **Fig 3A and 3B**). As discussed above, these optimized virus infection dynamics showed a longer duration of infectious period with low peak viral load (white dots 3 and 4 in **Fig 3C and 3D**). In contrast, in the region of a highly dense contact history ranging from 30 to 100, viruses were optimized with high production but low accuracy (white dots 1 and 2 in **Fig 3A and 3B**), representing the acute infection phenotype with the shorter duration of infectious period and high peak viral load (white dots 1 and 2 in **Fig 3C and 3D**). In addition, it is noteworthy that large variance in daily contacts can select for greater accuracy/lower proliferative capacity (**Fig 4A and 4B**). Viruses that adapted to the highly variable environment were characterized by a persistent infectious phenotype and a low transmission potential, especially under a small mean contact number (**Fig 4C and 4D**). With respect to the transmission potential,  $R_{TP}$ , a logarithmic increase is observed as the mean contact number increases (**Fig 3E**). Note that some of the persistent infection phenotypes evolving under the low-density contact history (white dot 4 in **Fig 3A, 3B, 3D and 3E**) might have a low transmission potential characterized by  $R_{TP} < 1$  (see **Discussion**).

Importantly, we demonstrated that, if the contact history is of intermediate density (with mean contact numbers ranging from 1 to 30), viruses can be optimized toward either acute or persistent infection, showing a critical transition on the virus infection phenotypes (**Fig 3D**). This bistable-like pattern may appear by chance owing to the stochasticity of contact number history, even when drawn from the same distribution. Further discussion is provided below.

### Discussion

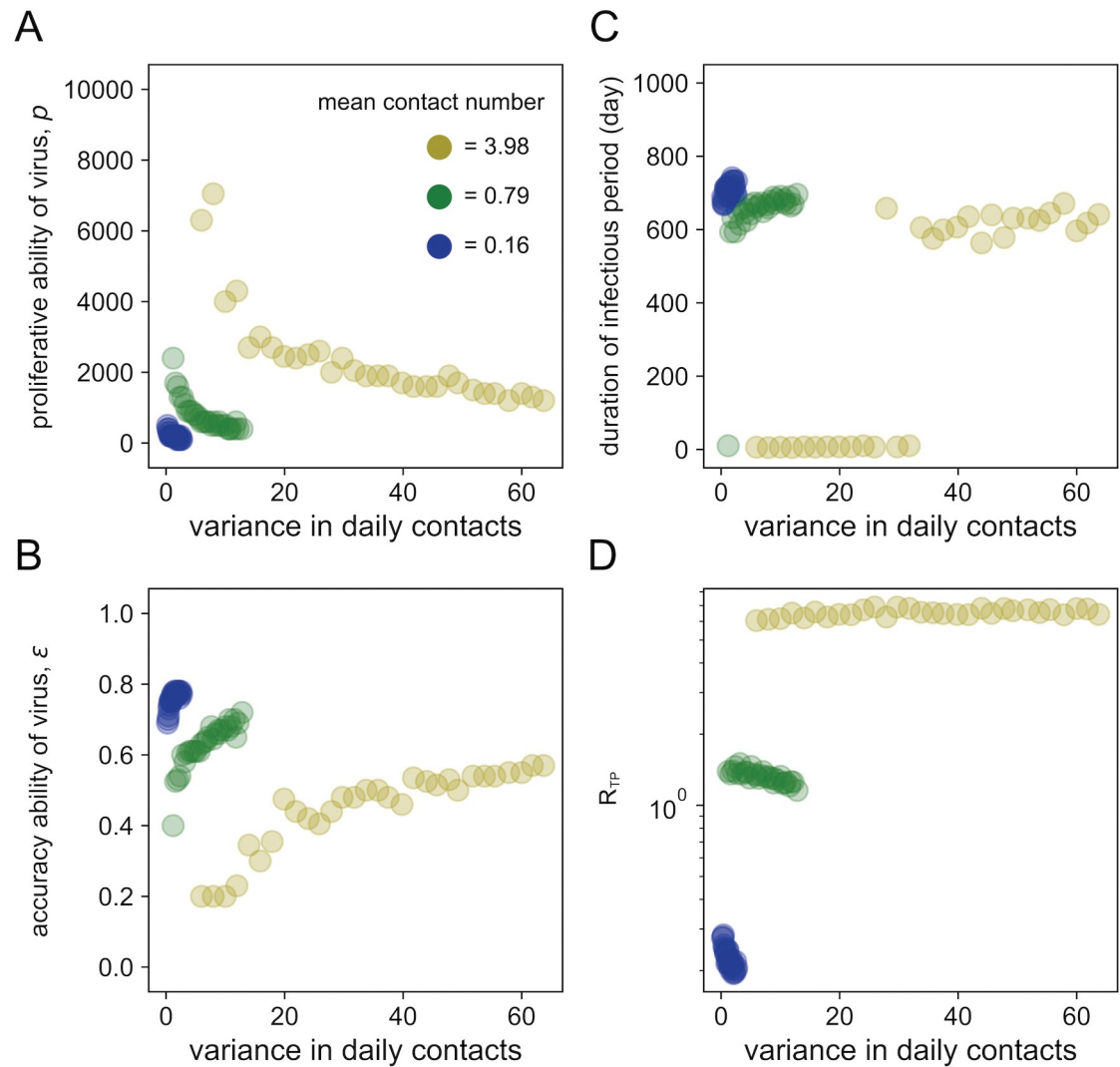
In this study, we investigated the impact of the proliferation rate and the accuracy of replication on viral evolution by considering optimal combinations for transmission under different contact histories of a host organism. Specifically, assuming an epidemiologic equilibrium, we developed a probabilistic multi-level model that characterizes population-level transmission



**Fig 3. Virus evolution in various scenarios of contact history.** A total of 441 contact history patterns are explored to determine the viral evolution to an optimal  $R_{TP}$  in a given environmental scenario. 500 iterations of each parameter combination of  $k$  and  $\theta$  are conducted to determine the optimized parameter set of  $p$  and  $\varepsilon$  characterizing final virus infection dynamics (i.e., different viral infection phenotypes). The median of the optimized parameter set of  $p$  and  $\varepsilon$  based on the 500 simulation runs is shown against the mean contact number of histories (with logarithmic increase) determined by each parameter set of  $k$  and  $\theta$ . The combination of  $p$  and  $\varepsilon$  is plotted in (A) and (B), separately. Colors represent the degree of the mean contact number: blue corresponds to a low mean contact number, whereas red corresponds to a high mean contact number. How virus infection dynamics vary with the mean contact number is shown in (C). Several examples of the virus infection dynamics are shown to represent the difference in the infectious period and the peak viral load. Numbers 1, 2, 3, and 4 in (C) are the corresponding dynamics illustrated as white dots in (A), (B), (D), and (E), respectively. The duration of the infectious period and  $R_{TP}$  varied with the mean contact number are plotted in (D) and (E), respectively. In (E), the vertical dashed line separates the area of mean contact number by an index,  $R_{TP} = 10^0 (= 1)$ . The corresponding dashed line is also illustrated in (A), (B), and (D).

<https://doi.org/10.1371/journal.pcbi.1011173.g003>

potential based on individual-level virus infection dynamics. The result presented above show that, depending on host density, acutely infectious viruses acquire a phenotype of a high mutation rate and a high replication rate, whereas persistently infectious viruses acquire a phenotype of a low mutation rate and a slow replication rate. A similar phenomenon has been observed in relation to the characteristics of actual RNA viruses. Borna disease virus (BDV), a persistent infectious RNA virus, has been reported to have a suppressed intracellular replication speed and a very slow evolutionary rate [22,23]. The number of virus particles released from BDV-infected cells is also extremely low [24]. Moreover, when coronaviruses shift their phenotype to persistent infectivity through the accumulation of mutations, translation efficiency and replication speed are markedly suppressed [25,26]. Similarly, bovine viral diarrhea virus (BVDV) strains that establish persistent infection evolve at a slower rate than acutely infectious strains. The acutely infectious strain of BVDV is a highly virulent strain that replicates more rapidly, has a higher viral load, and is more virulent than persistently infectious strains [27–31]. Therefore, the phenotypic shift in RNA viruses defined in this study has been confirmed in epidemiologic studies. As with many evolutionary biological models, our general model can be applied across many taxa to interpret effects on the diversity of shape of viral load. It should be noted that some existing studies have already addressed within- and



**Fig 4. Virus evolution in various scenarios of variance in daily contacts under three different mean contact numbers.** Each color represents the mean contact numbers of 3.98, 0.79, 0.16, respectively, corresponding to the colored dots in Fig 3 in the main text. We used the following parameter sets: the scale parameter  $\theta$  ranges from 0.5, 1.0, 1.5, ..., 14.5, 15.0 (30 points in total), and the shape parameter  $k$  ranges to keep the specific mean contact number. We conducted a total of 500 independent iterations for each parameter set of  $k$  and  $\theta$ , and showed the mean values. For different variance in daily contact numbers shown in the horizontal axis, the optimal parameter combination of  $p$  and  $\varepsilon$  that maximizes  $R_{TP}$  is plotted in (A) and (B), separately. The duration of the infectious period and  $R_{TP}$  based on the parameters determined above are plotted in (C) and (D), respectively.

<https://doi.org/10.1371/journal.pcbi.1011173.g004>

between-host(s) dynamics. There are, however, several differences between our model and previous ones.

First, although we explicitly defined the duration of infectious period as a function of viral load (Eqs (5–6)), some previous pieces of research considered the relationship between viral load and viral duration of infection under the invasion-persistence trade-off (e.g., [19,32]; also see [33] for immuno-epidemiologic multi-scale models that mechanistically capture the relationship). Those previous studies discussed a continuum of evolutionarily stable pathogen strategies. Pathogens evolve to greater acuteness to maximize between-host transmission in large well-mixed populations at one end, and selection favors less acute pathogens, which achieve enhanced between-population transmission when the host population is broken into



many small populations at the other end [20,34]. This study, without considering population structure of hosts, employs modelling of continuous evolution and phenotypic shifts of pathogens depending on mean and variance of contact rates, demonstrating the flexibility of adaptive evolution of pathogens that is observed in epidemiology in practice.

Second, research considering contact histories as an arena of virus evolution did not assume stochasticity [35] or extreme rare events [32]. There is a growing list of literature on the evolution of acuteness, especially shedding light on the importance of critical community size, which is approximately the population size above which pathogen extinction is not observed, implying an unbroken within-population chain of transmission (e.g., [36]). However, the previous research rarely considered variation in population size as a time series pattern. We adopted a negative binomial distribution to generate various contact histories, which corresponds to a broad range of host species' behavior. Using the current multi-level model, we especially focused on the condition where persistent viruses evolve. This study provides a detailed trace of how RNA viruses adapt to variable environments through the rare but significant events of beneficial mutations, which alter the phenotype and are inherited to the next generation of progeny viruses in a stochastic manner. As these mutations subsequently spread into the host population, they become fixed in the viral population, thereby shedding valuable insight into the complex stochastic process of viral evolution.

The acute and persistent infection phenotypes are classified in terms of the duration of the infectious period and the peak viral load, which is determined by the combination of the accuracy of genome RNA replication,  $\varepsilon$ , and its production rate,  $p$ . The acute infection phenotype (with a short infectious period and high peak viral load) is achieved in the parameter range where the production rate of virus particles is high and the replication accuracy is low. Although our model does not explicitly include the mutation process by which new viral phenotypes are generated (cf. [32]), this corresponds to a scenario with a low accuracy of replication in the current model, which is the case for influenza viruses with around 90% of noninfectious activities [37,38]. On the contrary, for the persistent infection phenotype, a long infectious period with a low peak viral load is obtained when there is high fidelity of viral replication and a low rate of genome replication. The properties characterized in this parameter region correspond to pathogens such as bornaviruses [39,40].

Our main finding is that the acute infection phenotype is optimal under conditions of a contact history involving frequent contacts between hosts (Figs 2AB and 3). In this environment, the virus does not need to stay in the infected host for a long time, because the focal infected individual can quickly obtain another opportunity to infect (called a “leaving-home” strategy in [41]). Therefore, an effective strategy is to sustain a high peak viral load and adopt an increased probability of infection even for a short period of time (Fig D in S1 Text). On the other hand, in an environment with a small mean and a large variance in contact history, since there is a possibility that no contact will occur at all in a short period of infection (Figs 2CD and 3), the virus must stay in the infected host for a long time (called the “stay-at-home” strategy in [41]). Hence,  $R_{TP}$  increases by maintaining an infectious period with a long duration while suppressing the peak viral load (Fig D in S1 Text). This corresponds to the evolution of the persistent infection phenotype. Thus, the contact history of infected hosts determines the direction of viral evolution balanced between the duration of infectious period and peak viral load. That is, persistent-acute infection phenotype switching is induced depending on the environment.

In this study, one of the parameters that influenced the infection phenotype was the replication accuracy  $\varepsilon$ . The RNA-dependent RNA polymerase encoded by RNA viruses lacks proof-reading activity, unlike DNA viruses. Owing to this low fidelity of viral polymerase, the genome sequence exists as diversified “quasispecies” in infected individuals [42–44]. These

intra-host quasispecies are involved in viral fitness *in vivo* and affect the virulence and transmission ability of a virus [4,42,45]. Interestingly, low-fidelity mutants of coxsackievirus B3, which primarily infects the heart muscle, have lost the ability to establish persistent infection *in vivo* [14]. This result on the relationship between the fidelity of viral polymerase and persistent infection *in vivo* are consistent with our study and support the model that replication accuracy  $\varepsilon$  determines the establishment of the persistent infection phenotype.

Note that we can also easily extend our interpretation of  $\varepsilon$  by including multiple factors such as the accuracy of virion assembly, the rate of virus particle aggregation, and others, in addition to the main factor of the accuracy of replication (see Eqs (1–4)). From the assumption that the virus is optimized to increase its transmission potential, the genetic algorithm heuristically searches for the optimal  $R_{TP}$  as a population-based method with a group of solutions made up of combinations of  $p$  and  $\varepsilon$ . In other words,  $p\varepsilon I(t)$  in our model represents “infectious virus production”, meaning the parameter  $\varepsilon$  can include the multiple above-mentioned factors.

From an evolutionary perspective, organisms, including viruses, adapt to the environments they experienced in the past, and their traits are optimized accordingly [46]. Our exhaustive simulations with the multi-level model suggest that the acute infection phenotype is more likely to be optimal when individuals live in high densities or when there is a highly connected network of hosts (as is often the case for humans and other social animals). On the other hand, the persistent infection phenotype may be optimal in organisms that do not form groups and where the environmental variance for the virus is largely dependent on their life history; for example, if the contact rate increases only during the breeding season. The unpredictability of the presence of any individual, which is referred to “biotic drift”, can influence evolutionary consequences [47]. The current model can be extended to analyze the seasonal (periodic) forcing in contact rates without loss of generality by assuming that the average value in the contact history is a periodic function.

Our simulations also show that the evolutionary endpoints of  $R_{TP}$  are larger in the acute infection phenotype than in the persistent infection phenotype (Fig 3E). For example, with low-density contact histories ranging from 0.01 to 0.57,  $R_{TP}$  decreases below 1, meaning that viruses with the persistent infection phenotype are unlikely to be sustained from the perspective of population-level transmission, even though the virus phenotype adapts to the low-density contact histories. This implies that acute virus infection, such as with SARS-CoV-2, will easily prevail in human society. BDV, an RNA virus that can establish persistent infection without cytopathic effects for a long time, has been observed as an endemic disease in Germany but has not reached a large-scale epidemic [22,48,49]. On the other hand, even if viruses with the persistent infection phenotype are not transmitted among individuals in the context of low-density contact history, an increase in contact numbers may lead to an optimized virus with the acute infection phenotype, increasing transmission potential within a host population. In fact, many acute infectious viruses that infect humans have been proposed to have been optimized through continuous mass infection after the formation of highly dense human communities for cultural reasons, such as for agriculture [50–53].

Large-scale genome sequencing and phylogenetic analysis of SARS-CoV-2 during the epidemic further the conclusions of this study. Epidemiologic studies have shown that SARS-CoV-2 evolved from pre-alpha to alpha, beta, and delta variants via host-to-host transmission [54]. These variants of concern (VOCs) transmit more quickly in the host population compared to background, suggesting that they may have shorter intrinsic generation times [55]. Moreover, the molecular clock modeling using the GISAID database and Bayesian analysis suggested that the substitution rate of SARS-CoV-2 is increasing with the emergence of VOCs relative to the background [56]. In the above, it can be inferred that the continuous chain of

infection within the host population facilitates viral evolution and contributes to the establishment of further acutely infectious viral phenotypes.

In terms of evolutionary biology, maximizing “fitness” like  $R_{TP}$  is used as a metric for evaluating the direction of evolution. For the sake of simplicity, in the current study, we considered only the evolution of the virus at an epidemiologic equilibrium. We ignored the evolution of the host for simplicity, although we expect this to be reasonable because the generation time of a virus is much shorter than that of the host organism. However, to capture a long-term non-equilibrium evolution, it may be necessary to consider coevolutionary dynamics between hosts and viruses. A recent study [21] showed that pathogens often showing antigenic escape from host immune defenses evolve through non-equilibrium dynamics, indicating that  $R_{TP}$  is no longer the measure to understand a viral evolutionarily stable strategy. Evolutionary game theory might be able to predict an evolutionarily stable strategy under the stochastic environmental scenario presented in this paper [57,58].

In addition, our approach has several limitations. For example, we do not explicitly include a host immune response or antiviral treatment, both of which may affect the duration of the infectious period and peak viral load. However, we do not expect this to change our qualitative conclusions, as these effects may be mimicked by changing the values of the parameters of our model. For example, the effect of cellular and humoral immune responses are implicitly included in the death rate of the infected cell (i.e.,  $\delta$  in Eq (3)) and the virus infection rate (i.e.,  $\beta$  in Eqs (1–2)), respectively. In our mathematical model, the viral propagation is limited by target cell availability rather than these immune responses. In fact, it seems that the target cell limitation is particularly important for infectious diseases in which the host’s immune response can reduce symptoms or negative health outcomes but does not effectively suppress viral replication.

Our mathematical model for population-level transmission, Eqs (7–8), does not account for “a transmission chain”, that is, we focused only on the sum of secondary cases generated throughout the infectious period by a primary case,  $R_{TP}$ . If entire chains of transmission are considered, it may be necessary to consider stochastic effects that may lead to particular viral phenotypes failing to become established in host populations. This would happen particularly in the region of a low-density contact history (the mean contact number from 0.01 to 1) for which  $R_{TP} \approx 1$  [59]. Future studies may explore this further by nesting a viral dynamics model within a stochastic population-scale transmission model (using, for example, a stochastic adaptation of the multi-scale epidemiologic modeling framework proposed by [60]). Furthermore, explicitly considering transmission mode, host heterogeneity and detailed life history may also impact the outcomes of viral evolution qualitatively.

Overall, we anticipate our study to be a key advance in modeling theory that associates the type of virus optimized to the host environment. We face emerging respiratory acute virus infections such as SARS-CoV-2, and host environments continue to change as transportation networks are developed and urban areas expanded. In particular, the frequency of contacts between people has increased yearly. As we showed here, such an environment favors acute viral infections. To prevent such emerging infectious diseases, it is crucial that infectious disease surveillance systems be strengthened, particularly in locations with high contact densities. In such a way, extracting the contact history of a host allows us to extract the predict and understand the evolution of future viruses and introduce interventions quickly to prevent the transmission of novel viruses and variants. This is of clear benefit for public health.

## Materials and methods

To understand the dependence of viral evolution on the “environment”, we employed a multi-level population dynamics model (i.e., coupling a population-level virus transmission model

and an individual-level virus infection model) and evaluated how the proliferative ability and accuracy of replication of viruses optimized to increase the transmission potential,  $R_{TP}$  (see Fig 1A). For the sake of simplicity, we considered an epidemiologic equilibrium. We assumed that environments are characterized by the number of contacts per day (i.e., contact numbers among populations), and that an infected individual has a probability of infection (per contact with a susceptible individual) that depends on viral load. Note that the “noninfectious” period of infected individuals has not been directly accounted for in the infection probability (see below). We assumed that a newly infected individual has the potential to spread the infection to others immediately after infection, although the probability of infection is very small during this period. Below, we describe the details of the model as separate sections.

### Virus infection dynamics

We employed a simple mathematical model for virus infection dynamics considering the accuracy of viral replications [61]:

$$\frac{dT(t)}{dt} = \lambda - \beta T(t) V_i(t) - \mu T(t), \quad (1)$$

$$\frac{dI(t)}{dt} = \beta T(t) V_i(t) - \delta I(t), \quad (2)$$

$$\frac{dV_i(t)}{dt} = p \varepsilon I(t) - c V_i(t), \quad (3)$$

$$\frac{dV_{\text{non}}(t)}{dt} = p(1 - \varepsilon) I(t) - c V_{\text{non}}(t), \quad (4)$$

where the variables  $T(t)$ ,  $I(t)$ ,  $V_i(t)$  and  $V_{\text{non}}(t)$  are the number of uninfected target cells, infected target cells, and the amount of infectious virus (i.e., infectious viral load) and noninfectious virus at time  $t$  since infection, respectively. The parameters  $\beta$ ,  $\delta$ ,  $p$ , and  $c$  represent the rate constant for virus infection, the death rate of infected cells, the per-cell viral production rate (i.e., proliferative ability of virus), and the per capita clearance rate of the virus, respectively. In addition, we assumed a fraction  $\varepsilon$  of the produced virus is infectious but  $1 - \varepsilon$  is noninfectious (the value of  $\varepsilon$  represents the accuracy of replication). To mimic viral cytopathogenesis depending on the viral replication level, we defined  $\delta = \delta_{\text{max}} p / (p + p_{50})$ , that is, an increasing Hill function of  $p$ . The parameters  $\delta_{\text{max}}$  and  $p_{50}$  are the maximum value of  $\delta$  and the viral production rate satisfying  $\delta = \delta_{\text{max}}/2$ , respectively. We fixed the parameter values (except for  $p$  and  $\varepsilon$ ) and initial values to correspond to biologically reasonable values or ranges (see below and Table A in S1 Text). The sensitivity analysis on  $\delta$  is provided in Fig G in S1 Text.

### Duration of infectious period as a function of viral load

To describe the relation between the duration of infectious period and the viral load across a broad range of viruses (i.e., from acute phenotypes to persistent phenotypes) rather to explain properties of a specific virus, we assumed the duration of infectious period depends on the total viral load (i.e., the cumulative viral load) as observed for several pathogens [62,63]:

$$D(V_{\text{total}}) = \frac{D_{\text{max}} D_{50}^{D_K}}{V_{50}^{D_K} + V_{\text{total}}^{D_K}}, \quad (5)$$

$$V_{\text{total}} = \int_0^{D_{\text{max}}} V_i(t) dt, \quad (6)$$

where  $D_{\text{max}}$  is the maximum duration of the infectious period,  $V_{50}$  is the viral load at which the duration is half of its maximum, and  $D_k$  is the steepness at which duration decreases with increasing viral load [63]. We also conducted a sensitivity analysis on the duration of the infectious period, using different function shapes instead of assuming Eq (5) (**Figs H, I, and J in S1 Text**). Note that other functions may be suitable to evaluate only a specific acute phenotypic virus, such as influenza virus or SARS-CoV-2, because the infected individual no longer sheds viruses after recovery or death.

### Probability of infection

To evaluate the daily transmissibility of an infected individual through the duration of infectious period, we assumed that for any given contact, each 10-fold increase in infectious viral load will lead to an  $r$ -fold increase in infectiousness [64]; thus, the probability of infection  $\rho$  at time  $t$  is described by:

$$\rho(t) = b \cdot r^{\log_{10}(V_i(t))}, \quad (7)$$

where  $b$  is the basilar probability of infection immediately after being infected with the virus.

### Number of encounters

To mimic a daily contact history, we assume the number of contacts on any day is drawn from a negative binomial distribution [65]. Since a negative binomial distribution is identical to a Poisson distribution where the mean parameter  $\lambda$  follows a gamma distribution, the daily contact numbers  $C(t)$  of a focal infectious individual at time  $t$  are generated by sampling from the following distributions (see **Fig 1C** as examples):

$$\lambda \sim \text{gamma}(k, \theta),$$

$$C(t) \sim P_o(\lambda),$$

where  $k$  and  $\theta$  are the shape parameter and the scale parameter, respectively. We note that  $k$  influences the skewness, and  $\theta$  influences the variance of the distribution, respectively. The mean contact number through the history is the product of  $k$  and  $\theta$ , that is,  $k\theta$ .

As a supplementary analysis, we also considered a different approach for generating the contact numbers  $C(t)$ , by considering a contact scenario on different networks: the Barabási–Albert (BA) scale-free network model and the Watts–Strogatz (WS) small-world network model (see the following sections: **Virus evolution based on network-generating contact Table B, Figs E, and F in S1 Text**). The BA model generates a network such that a graph of  $n$  nodes is grown by attaching new nodes, each with  $m$  edges, that are preferentially attached to existing nodes with high degree [66]. The WS model generates a network such that starting from a ring over  $n$  nodes, which is connected with its  $q$  nearest neighbors, shortcuts are created by replacing some edges with probability  $p$  [67]. Contact numbers on a generated graph are calculated by counting the number of nodes that are connected with the one randomly chosen node on the network at each time step.



## Transmission potential and fitness landscape

We calculated the total number of secondary cases generated throughout the infectious period of a primary case ( $R_{TP}$ ). On each day, the number of secondary cases is considered as the multiplication of the number of encounters (i.e., the contact number) and the probability of infection. Thus,  $R_{TP}$  is calculated as below:

$$R_{TP} = \sum_{t=0}^{D(V_{\text{total}})} (C(t)\rho(t)), \quad (8)$$

where  $D(V_{\text{total}})$  is the duration of infectious period of an infected individual defined by Eq (5). Once fixing the daily contact number,  $R_{TP}$  is dependent on the parameters appearing in the mathematical model for the virus infection dynamics, Eqs (1–4). We hereafter focus on the dependence of the rate of proliferation and accuracy in the viral life cycle,  $p$  and  $\varepsilon$ , respectively. We calculate  $R_{TP}$  in various patterns of  $(p, \varepsilon)$ , and thus the fitness landscape of  $R_{TP}$  is constructed as a function of  $p$  and  $\varepsilon$ .

## Evolution of virus to increase the transmission potential

One simple assumption is that the viral population will eventually be dominated by the virus with the largest  $R_{TP}$ . From this perspective, a genetic algorithm [68], a population-based method of evolutionary computation, is implemented to search for the optimal point in the fitness landscape (i.e., maximizing  $R_{TP}$ ) regarding the combination of  $p$  and  $\varepsilon$ . The genetic algorithm heuristically searches for the optimal  $R_{TP}$  with a group of solutions made up of combinations of  $p$  and  $\varepsilon$  (see **Alg.1** and **Table B in S1 Text** for details). Initially, solutions are randomly placed around the fitness landscape (Initialize) and are assigned  $R_{TP}$  as fitness (Evaluate). Over successive iterations, the combinations of  $p$  and  $\varepsilon$  with higher  $R_{TP}$  are selected to build the next solutions (selectSolutions) in a fitness proportionate manner where the solution with higher  $R_{TP}$  has a greater chance to be chosen for the next iteration. Note that the two highest-performing solutions are always chosen for the next iteration (Elitism). To find better combinations of  $p$  and  $\varepsilon$ , solutions are recombined (crossover), and each parameter is slightly altered by adding a random variable drawn from a Uniform distribution with a pre-determined range (Mutate). Finally, the algorithm searches for a better solution with the expectation of finally converging to the combination of  $p$  and  $\varepsilon$  with the highest  $R_{TP}$ . These iterations can be considered as the virus "evolving" its properties of  $p$  and  $\varepsilon$  toward an optimal solution so that the virus can increase its own fitness in the evolutionary history.

## Virus evolution based on network-generating contact history

We investigate how the properties of a virus evolve based on contact history generated from the following well-established network models: Barabási-Albert scale-free network model (BA model) and Watts-Strogatz small world network model (WS model). The contact number of a contact history is the number of nodes connected to a node chosen at random from all nodes in the network (see **Table C in S1 Text** for the parameters used). As shown by 50 independent iterations, a virus under the contact history generated by both the BA model and the WS model evolves high virus production but low accuracy (**Figs EC and FC in S1 Text**), which corresponds to the characteristics of the acute infection phenotype.

Algorithm 1: Genetic Algorithm

Bb

Input:  $P(t = i) = \{s_1, s_2, \dots, s_j\}$ , population of solutions in iteration  $i$

```

t ← 0;
Initialize (P(t = 0)) with random p and ε;
Evaluate (P(t = 0));
while not termination do
  P(t)e ← Elitism (P(t));
  P(t)p ← selectSolutions (P(t));
  P(t)c ← crossover (P(t)e + P(t)p);
  Mutate (P(t)c);
  Evaluate (P(t)c);
  P(t+1) ← P(t)c;
  t ← t + 1;
end while

```

## Supporting information

**S1 Text. Supplementary tables and figures.** Table A. Parameters used for virus infection dynamics in the main text. Table B. Parameters used in GA Algorithm. Table C. Parameters used for the network models. Fig D. Relation between duration of infectious period and peak viral load. Fig E. Evolution of virus on a contact history generated from the Barabási-Albert scale free network model. Fig F. Evolution of virus on a contact history generated from the Watts-Strogatz small world network model. Fig G. Virus evolution in various scenarios of contact history. Fig H. Different function types in the duration of infectious period assumed to depend on the cumulative viral load. Fig I. Virus evolution in various scenarios of contact history where a monotonically decreasing linear relationship is assumed between the duration of the infectious period and the cumulative viral load. Fig J. Virus evolution in various scenarios of contact history where a monotonically decreasing exponential relationship is assumed between the duration of the infectious period and the cumulative viral load (DOCX)

## Author Contributions

**Conceptualization:** Ryo Komorizono, William S. Hart, Robin N. Thompson, Akiko Makino, Keizo Tomonaga, Shingo Iwami, Ryo Yamaguchi.

**Formal analysis:** Junya Sunagawa, Shingo Iwami, Ryo Yamaguchi.

**Funding acquisition:** Keizo Tomonaga, Shingo Iwami, Ryo Yamaguchi.

**Investigation:** Junya Sunagawa.

**Methodology:** Shingo Iwami, Ryo Yamaguchi.

**Software:** Junya Sunagawa.

**Supervision:** Shingo Iwami, Ryo Yamaguchi.

**Validation:** Junya Sunagawa.

**Visualization:** Junya Sunagawa, Shingo Iwami, Ryo Yamaguchi.

**Writing – original draft:** Junya Sunagawa, Ryo Komorizono, Hyeongki Park, William S. Hart, Robin N. Thompson, Akiko Makino, Keizo Tomonaga, Shingo Iwami, Ryo Yamaguchi.

**Writing – review & editing:** Junya Sunagawa, Ryo Komorizono, Hyeongki Park, William S. Hart, Robin N. Thompson, Akiko Makino, Keizo Tomonaga, Shingo Iwami, Ryo Yamaguchi.

## References

- Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S. Rapid evolution of RNA genomes. *Science*. 1982; 215(4540):1577–85. <https://doi.org/10.1126/science.7041255> PMID: 7041255.
- Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. *J Virol*. 2010; 84(19):9733–48. Epub 20100721. <https://doi.org/10.1128/JVI.00694-10> PMID: 20660197; PubMed Central PMCID: PMC2937809.
- Lauring AS, Frydman J, Andino R. The role of mutational robustness in RNA virus evolution. *Nature Reviews Microbiology*. 2013; 11(5):327–36. <https://doi.org/10.1038/nrmicro3003> PMID: 23524517
- Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature*. 2006; 439(7074):344–8. <https://doi.org/10.1038/nature04388> PMID: 16327776
- Xiao C, Gao L, Hou Y, Xu C, Chang N, Wang F, et al. Chromatin-remodelling factor Brg1 regulates myocardial proliferation and regeneration in zebrafish. *Nature Communications*. 2016; 7(1):13787. <https://doi.org/10.1038/ncomms13787> PMID: 27929112
- Sender R, Bar-On YM, Gleizer S, Bernshtein B, Flamholz A, Phillips R, et al. The total number and mass of SARS-CoV-2 virions. *Proc Natl Acad Sci U S A*. 2021; 118(25). <https://doi.org/10.1073/pnas.2024815118> PMID: 34083352; PubMed Central PMCID: PMC8237675.
- Cleaveland S, Laurenson MK, Taylor LH. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos Trans R Soc Lond B Biol Sci*. 2001; 356(1411):991–9. <https://doi.org/10.1098/rstb.2001.0889> PMID: 11516377; PubMed Central PMCID: PMC1088494.
- Woolhouse ME, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis*. 2005; 11(12):1842–7. <https://doi.org/10.3201/eid1112.050997> PMID: 16485468; PubMed Central PMCID: PMC3367654.
- McCarthy MK, Morrison TE. Persistent RNA virus infections: do PAMPS drive chronic disease? *Current Opinion in Virology*. 2017; 23:8–15. <https://doi.org/10.1016/j.coviro.2017.01.003> PMID: 28214732
- Geoghegan JL, Senior AM, Di Giallonardo F, Holmes EC. Virological factors that increase the transmissibility of emerging human viruses. *Proceedings of the National Academy of Sciences of the United States of America*. 2016; 113(15):4170–5. Epub 2016/03/21. <https://doi.org/10.1073/pnas.1521582113> PMID: 27001840.
- Randall RE, Griffin DE. Within host RNA virus persistence: mechanisms and consequences. *Curr Opin Virol*. 2017; 23:35–42. Epub 20170317. <https://doi.org/10.1016/j.coviro.2017.03.001> PMID: 28319790; PubMed Central PMCID: PMC5474179.
- Xu HC, Wang R, Shinde PV, Walotka L, Huang A, Poschmann G, et al. Slow viral propagation during initial phase of infection leads to viral persistence in mice. *Communications Biology*. 2021; 4(1):508. <https://doi.org/10.1038/s42003-021-02028-x> PMID: 33927339
- Arias A, Thorne L, Ghurburrun E, Bailey D, Goodfellow I, Duprex WP. Norovirus Polymerase Fidelity Contributes to Viral Transmission In Vivo. *mSphere*. 1(5):e00279–16. <https://doi.org/10.1128/mSphere.00279-16> PMID: 27777985
- Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, Gong P, et al. Coxsackievirus B3 mutator strains are attenuated in vivo. *Proceedings of the National Academy of Sciences*. 2012; 109(34):E2294. <https://doi.org/10.1073/pnas.1204022109> PMID: 22853955
- Trimpert J, Groenke N, Kunec D, Eschke K, He S, McMahon DP, et al. A proofreading-impaired herpesvirus generates populations with quasispecies-like structure. *Nature Microbiology*. 2019; 4(12):2175–83. <https://doi.org/10.1038/s41564-019-0547-x> PMID: 31477893
- Rutz C, Loretto M-C, Bates AE, Davidson SC, Duarte CM, Jetz W, et al. COVID-19 lockdown allows researchers to quantify the effects of human activity on wildlife. *Nature Ecology & Evolution*. 2020; 4(9):1156–9. <https://doi.org/10.1038/s41559-020-1237-z> PMID: 32572222
- Frickel J, Feulner PGD, Karakoc E, Becks L. Population size changes and selection drive patterns of parallel evolution in a host-virus system. *Nature communications*. 2018; 9(1):1706–. <https://doi.org/10.1038/s41467-018-03990-7> PMID: 29703896.
- Rouzine IM, Weinberger AD, Weinberger LS. An evolutionary role for HIV latency in enhancing viral transmission. *Cell*. 2015; 160(5):1002–12. <https://doi.org/10.1016/j.cell.2015.02.017> PMID: 25723173; PubMed Central PMCID: PMC4488136.
- Goyal A, Reeves DB, Cardozo-Ojeda EF, Schiffer JT, Mayer BT. Viral load and contact heterogeneity predict SARS-CoV-2 transmission and super-spreading events. *Elife*. 2021; 10. Epub 20210223. <https://doi.org/10.7554/eLife.63537> PMID: 33620317; PubMed Central PMCID: PMC7929560.

20. Shrestha S, Bjørnstad ON, King AA. Evolution of acuteness in pathogen metapopulations: conflicts between “classical” and invasion-persistence trade-offs. *Theoretical Ecology*. 2014; 7(3):299–311. <https://doi.org/10.1007/s12080-014-0219-7> PMID: 25214895
21. Sasaki A, Lion S, Boots M. Antigenic escape selects for the evolution of higher pathogen transmission and virulence. *Nature Ecology & Evolution*. 2022; 6(1):51–62. <https://doi.org/10.1038/s41559-021-01603-z> PMID: 34949816
22. Kinnunen PM, Palva A, Vaheri A, Vapalahti O. Epidemiology and host spectrum of Borna disease virus infections. *J Gen Virol*. 2013; 94(Pt 2):247–62. Epub 20121205. <https://doi.org/10.1099/vir.0.046961-0> PMID: 23223618.
23. Staeheli P, Sauder C, Hausmann J, Ehrensperger F, Schwemmler M. Epidemiology of Borna disease virus. *J Gen Virol*. 2000; 81(Pt 9):2123–35. <https://doi.org/10.1099/0022-1317-81-9-2123> PMID: 10950968.
24. Pauli G, Ludwig H. Increase of virus yields and releases of Borna disease virus from persistently infected cells. *Virus Research*. 1985; 2(1):29–33. [https://doi.org/10.1016/0168-1702\(85\)90057-7](https://doi.org/10.1016/0168-1702(85)90057-7) PMID: 3920840
25. Banerjee A, Subudhi S, Rapin N, Lew J, Jain R, Falzarano D, et al. Selection of viral variants during persistent infection of insectivorous bat cells with Middle East respiratory syndrome coronavirus. *Scientific Reports*. 2020; 10(1):7257. <https://doi.org/10.1038/s41598-020-64264-1> PMID: 32350357
26. Ke T-Y, Liao W-Y, Wu H-Y. A leaderless genome identified during persistent bovine coronavirus infection is associated with attenuation of gene expression. *PLoS One*. 2013; 8(12):e82176. <https://doi.org/10.1371/journal.pone.0082176> PMID: 24349214
27. Peterhans E, Bachofen C, Stalder H, Schweizer M. Cytopathic bovine viral diarrhoea viruses (BVDV): emerging pestiviruses doomed to extinction. *Vet Res*. 2010; 41(6):44. Epub 20100304. <https://doi.org/10.1051/vetres/2010016> PMID: 20197026; PubMed Central PMCID: PMC2850149.
28. Hamers C, Lecomte C, Kulcsar G, Lambot M, Pastoret PP. Persistently infected cattle stabilise bovine viral diarrhoea virus leading to herd specific strains. *Vet Microbiol*. 1998; 61(3):177–82. [https://doi.org/10.1016/s0378-1135\(98\)00185-0](https://doi.org/10.1016/s0378-1135(98)00185-0) PMID: 9631529.
29. Hornberg A, Fernández SR, Vogl C, Vilcek S, Matt M, Fink M, et al. Genetic diversity of pestivirus isolates in cattle from Western Austria. *Vet Microbiol*. 2009; 135(3–4):205–13. Epub 20080926. <https://doi.org/10.1016/j.vetmic.2008.09.068> PMID: 19019571.
30. Vilcek S, Alenius S, Paton DJ, Mittelholzer C, Belák S. Genetic Clustering of Bovine Viral Diarrhoea Viruses in Cattle Farms: Genetic Identification and Analysis of Viruses Directly from Cattle Sera. *The Veterinary Journal*. 1999; 158(1):33–8. <https://doi.org/10.1053/tvj.1999.0363> PMID: 10409414
31. Bachofen C, Stalder H, Braun U, Hilbe M, Ehrensperger F, Peterhans E. Co-existence of genetically and antigenically diverse bovine viral diarrhoea viruses in an endemic situation. *Vet Microbiol*. 2008; 131(1–2):93–102. Epub 20080304. <https://doi.org/10.1016/j.vetmic.2008.02.023> PMID: 18424020.
32. Schreiber SJ, Ke R, Loverdo C, Park M, Ahsan P, Lloyd-Smith JO. Cross-scale dynamics and the evolutionary emergence of infectious diseases. *Virus Evol*. 2021; 7(1):veaa105. Epub 20210420. <https://doi.org/10.1093/ve/veaa105> PMID: 35186322; PubMed Central PMCID: PMC8087961.
33. Gulbudak H, Cannataro VL, Tuncer N, Martcheva M. Vector-Borne Pathogen and Host Evolution in a Structured Immuno-Epidemiological System. *Bulletin of Mathematical Biology*. 2017; 79(2):325–55. <https://doi.org/10.1007/s11538-016-0239-0> PMID: 28032207
34. King Aaron A, Shrestha S, Harvill Eric T, Bjørnstad Ottar N. Evolution of Acute Infections and the Invasion-Persistence Trade-Off. *The American Naturalist*. 2009; 173(4):446–55. <https://doi.org/10.1086/597217> PMID: 19231966
35. Regoes RR, Hamblin S, Tanaka MM. Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. *Proc Biol Sci*. 2013; 280(1750):20122047. Epub 20121107. <https://doi.org/10.1098/rspb.2012.2047> PMID: 23135674; PubMed Central PMCID: PMC3574426.
36. Gunning CE, Wearing HJ. Probabilistic measures of persistence and extinction in measles (meta)populations. *Ecology Letters*. 2013; 16(8):985–94. <https://doi.org/10.1111/ele.12124> PMID: 23782847
37. Brooke CB. Biological activities of ‘noninfectious’ influenza A virus particles. *Future Virol*. 2014; 9(1):41–51. <https://doi.org/10.2217/fvl.13.118> PMID: 25067941; PubMed Central PMCID: PMC4109409.
38. Pinilla LT, Holder BP, Abed Y, Boivin G, Beauchemin CA. The H275Y neuraminidase mutation of the pandemic A/H1N1 influenza virus lengthens the eclipse phase and reduces viral output of infected cells, potentially compromising fitness in ferrets. *J Virol*. 2012; 86(19):10651–60. Epub 20120725. <https://doi.org/10.1128/JVI.07244-11> PMID: 22837199; PubMed Central PMCID: PMC3457267.
39. Binz T, Lebelt J, Niemann H, Hagenau K. Sequence analyses of the p24 gene of Borna disease virus in naturally infected horse, donkey and sheep. *Virus Research*. 1994; 34(3):281–9. [https://doi.org/10.1016/0168-1702\(94\)90128-7](https://doi.org/10.1016/0168-1702(94)90128-7) PMID: 7856315

40. Schneider PA, Briele T, Zimmermann W, Ludwig H, Lipkin WI. Sequence conservation in field and experimental isolates of Borna disease virus. *Journal of Virology*. 1994; 68(1):63–8. <https://doi.org/10.1128/JVI.68.1.63-68.1994> PMID: 8254777
41. Iwanami S, Kitagawa K, Ohashi H, Asai Y, Shionoya K, Saso W, et al. Should a viral genome stay in the host cell or leave? A quantitative dynamics study of how hepatitis C virus deals with this dilemma. *PLoS Biol*. 2020; 18(7):e3000562. Epub 20200730. <https://doi.org/10.1371/journal.pbio.3000562> PMID: 32730280; PubMed Central PMCID: PMC7392214.
42. Andino R, Domingo E. Viral quasispecies. *Virology*. 2015; 479–480:46–51. Epub 20150329. <https://doi.org/10.1016/j.virol.2015.03.022> PMID: 25824477; PubMed Central PMCID: PMC4826558.
43. Holland JJ, De La Torre JC, Steinhauer DA. RNA virus populations as quasispecies. *Curr Top Microbiol Immunol*. 1992; 176:1–20. [https://doi.org/10.1007/978-3-642-77011-1\\_1](https://doi.org/10.1007/978-3-642-77011-1_1) PMID: 1600747.
44. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev*. 2012; 76(2):159–216. <https://doi.org/10.1128/MMBR.05023-11> PMID: 22688811; PubMed Central PMCID: PMC3372249.
45. Dolan PT, Whitfield ZJ, Andino R. Mechanisms and Concepts in RNA Virus Population Dynamics and Evolution. *Annu Rev Virol*. 2018; 5(1):69–92. Epub 20180726. <https://doi.org/10.1146/annurev-virology-101416-041718> PMID: 30048219.
46. Yamaguchi R, Otto SP. Insights from Fisher's geometric model on the likelihood of speciation under different histories of environmental change. *Evolution*. 2020; 74(8):1603–19. <https://doi.org/10.1111/evo.14032> PMID: 32542705
47. Lenormand T, Roze D, Rousset F. Stochasticity in evolution. *Trends Ecol Evol*. 2009; 24(3):157–65. Epub 20090127. <https://doi.org/10.1016/j.tree.2008.09.014> PMID: 19178980.
48. Briele T, de la Torre JC, Lewis A, Ludwig H, Lipkin WI. Borna disease virus, a negative-strand RNA virus, transcribes in the nucleus of infected cells. *Proc Natl Acad Sci U S A*. 1992; 89(23):11486–9. <https://doi.org/10.1073/pnas.89.23.11486> PMID: 1454837; PubMed Central PMCID: PMC50576.
49. Cubitt B, Oldstone C, de la Torre JC. Sequence and genome organization of Borna disease virus. *J Virol*. 1994; 68(3):1382–96. <https://doi.org/10.1128/JVI.68.3.1382-1396.1994> PMID: 7906311; PubMed Central PMCID: PMC236592.
50. Villarreal LP, Defilippis VR, Gottlieb KA. Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology*. 2000; 272(1):1–6. <https://doi.org/10.1006/viro.2000.0381> PMID: 10873743.
51. McNeill WH, McNeill W. *Plagues and peoples*: Anchor; 1998.
52. Oldstone MB, Scheifele DW. Viruses, plagues & history. *Canadian Medical Association Journal*. 1999; 160(3):378.
53. Diamond J, Renfrew C. Guns, germs, and steel: The fates of human societies. *Nature*. 1997; 386(6623):339–.
54. Koelle K, Martin MA, Antia R, Lopman B, Dean NE. The changing epidemiology of SARS-CoV-2. *Science*. 2022; 375(6585):1116–21. Epub 20220310. <https://doi.org/10.1126/science.abm4915> PMID: 35271324; PubMed Central PMCID: PMC9009722.
55. Hart WS, Miller E, Andrews NJ, Waight P, Maini PK, Funk S, et al. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *Lancet Infect Dis*. 2022; 22(5):603–10. Epub 20220214. [https://doi.org/10.1016/S1473-3099\(22\)00001-9](https://doi.org/10.1016/S1473-3099(22)00001-9) PMID: 35176230; PubMed Central PMCID: PMC8843191.
56. Tay JH, Porter AF, Wirth W, Duchene S. The Emergence of SARS-CoV-2 Variants of Concern Is Driven by Acceleration of the Substitution Rate. *Mol Biol Evol*. 2022; 39(2). <https://doi.org/10.1093/molbev/msac013> PMID: 35038741; PubMed Central PMCID: PMC8807201.
57. Nowak MA, Sigmund K. Evolutionary dynamics of biological games. *Science*. 2004; 303(5659):793–9. <https://doi.org/10.1126/science.1093411> PMID: 14764867.
58. Renaud F, de Meeüs T. A simple model of host-parasite evolutionary relationships. *Parasitism: compromise or conflict?* *J Theor Biol*. 1991; 152(3):319–27. [https://doi.org/10.1016/s0022-5193\(05\)80197-3](https://doi.org/10.1016/s0022-5193(05)80197-3) PMID: 1749254.
59. Thompson RN, Gilligan CA, Cuniffe NJ. Detecting Presymptomatic Infection Is Necessary to Forecast Major Epidemics in the Earliest Stages of Infectious Disease Outbreaks. *PLoS Comput Biol*. 2016; 12(4):e1004836. Epub 20160405. <https://doi.org/10.1371/journal.pcbi.1004836> PMID: 27046030; PubMed Central PMCID: PMC4821482.
60. Hart WS, Maini PK, Yates CA, Thompson RN. A theoretical framework for transitioning from patient-level to population-scale epidemiological dynamics: influenza A as a case study. *Journal of The Royal Society Interface*. 2020; 17(166):20200230. <https://doi.org/10.1098/rsif.2020.0230> PMID: 32400267



61. Iwami S, Holder BP, Beauchemin CAA, Morita S, Tada T, Sato K, et al. Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an in vitro experiment and a mathematical model. *Retrovirology*. 2012; 9(1):18. <https://doi.org/10.1186/1742-4690-9-18> PMID: 22364292
62. Iwami S, Sato K, Morita S, Inaba H, Kobayashi T, Takeuchi JS, et al. Pandemic HIV-1 Vpu overcomes intrinsic herd immunity mediated by tetherin. *Scientific reports* [Internet]. 2015 2015/07//; 5:[12256 p.]. Available from: <http://europepmc.org/abstract/MED/26184634> Available from: <https://doi.org/10.1038/srep12256> from: <https://europepmc.org/articles/PMC4505337> Available from: <https://europepmc.org/articles/PMC4505337?pdf=render>. PMID: 26184634
63. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc Natl Acad Sci U S A*. 2007; 104(44):17441–6. Epub 20071022. <https://doi.org/10.1073/pnas.0708559104> PMID: 17954909; PubMed Central PMCID: PMC2077275.
64. Yamin D, Gertler S, Ndeffo-Mbah ML, Skrip LA, Fallah M, Nyenswah TG, et al. Effect of Ebola progression on transmission and control in Liberia. *Ann Intern Med*. 2015; 162(1):11–7. <https://doi.org/10.7326/M14-2255> PMID: 25347321; PubMed Central PMCID: PMC4402942.
65. Lindén A, Mäntyniemi S. Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology*. 2011; 92(7):1414–21. <https://doi.org/10.1890/10-1831.1> PMID: 21870615.
66. Barabási A-L, Albert R. Emergence of Scaling in Random Networks. *Science*. 1999; 286(5439):509–12. <https://doi.org/10.1126/science.286.5439.509> PMID: 10521342
67. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature*. 1998; 393(6684):440–2. <https://doi.org/10.1038/30918> PMID: 9623998
68. Holland JH. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*: The MIT Press; 1992. Available from: <https://doi.org/10.7551/mitpress/1090.001.0001>.
69. Sunagawa J, Komorizono R, Park H, Hart WS, Thompson RN, Makino A, Tomonaga K, Iwami S, Yamaguchi R. 2023. Codes from: Contact-number-driven virus evolution: a multi-level modeling framework for the evolution of acute or persistent RNA virus infection. [https://github.com/j-s9vcp/ContactNumberDriven\\_VirusEvolution](https://github.com/j-s9vcp/ContactNumberDriven_VirusEvolution)