



HOKKAIDO UNIVERSITY

Title	Environmental Variability and Adaptative Evolution : Insights from Neural-Network-based Artificial Agents and Virus Dynamics
Author(s)	砂川, 純也
Degree Grantor	北海道大学
Degree Name	博士(生命科学)
Dissertation Number	甲第16781号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16781
Doc URL	https://hdl.handle.net/2115/99626
Type	doctoral thesis
File Information	Junya_Sunagawa.pdf, 全文



博士学位論文

Environmental Variability and Adaptative Evolution: Insights from Neural-
Network-based Artificial Agents and Virus Dynamics

(環境変動と適応進化：ニューラルネットワークを用いた人工エージェントとウイルス動態に
基づく知見)

砂川 純也

北海道大学大学院生命科学院

2026 年 3 月

Contents

Preface	4
Acknowledgement.....	8
Chapter1 : Evolving neural networks through bio-inspired parent selection in dynamic environments	9
Introduction	10
Related works.....	13
Materials and methods.....	19
Results	21
Conclusion.....	31
References	33
Figures.....	38
Chapter2 : Contact-number-driven virus evolution: A multi-level modeling framework for the evolution of acute or persistent RNA virus infection	61
Introduction	62
Results	65
Discussion	69
Materials and methods.....	77
References	84
Figures.....	91
Supplementary Tables	98
Supplementary Figures.....	101
Chapter3 : Isolation may select for earlier and higher peak viral load but shorter duration in SARS-CoV-2 evolution.....	113
Introduction	114
Results	117
Discussion	128

Methods	134
References	142
Figures	150
Supplementary Tables	158

Preface

An organism's ability to reproduce across generations determines its evolutionary success. This is called "fitness", a quantitative measure of reproductive success. Genotype (or phenotype) and the environment the organism lives in shape its traits and functions that affect its fitness. To illustrate how genetic and phenotypic differences affect fitness changes, evolutionary biology often employs the concept of a fitness landscape. In this metaphor, every possible genotype (or phenotype) corresponds to a point in a multidimensional space, and each point is assigned a fitness value. Plotting these values creates a landscape with peaks and valleys like a mountain, where higher peaks represent genotype (or phenotype) with greater fitness. In such a landscape, a population composed of multiple organisms can evolve (seen as 'walks') and adapt (seen as 'climbs') toward higher positions on the fitness surface.

Numerous studies across diverse fields have characterized fitness landscapes in specific environments and identified their optima, exploring how the peak can be reached through evolutionary dynamics. For example, in evolutionary computation, an artificial agent learns to move forward and backward in a planar environment by optimizing its body architecture. Identifying a high-performing architecture involves searching the parameter space that might govern joint angles or the number of limbs, which process can be interpreted as exploring the fitness landscape. Likewise in the virus, the viruses evolve and adapt their phenotypes ranging from increased transmissibility against uninfected cells to evasion of host immune responses. The fitness landscape defined by these parameters can highlight accessible peaks for a viral population to persist within a host and spread within a host population.

However, less effort has been devoted to how different environment

reshapes the optimal genotypic (or phenotypic) state of a focal organism. The situation corresponds to an agent that encounters obstacles and needs to re-construct the architecture to navigate around them, or to a virus that prevails in a different host population. In addition, when one or more environmental shifts occur during adaptive evolution, the location of the optimum changes. As a result, the population experiences a decrease in fitness and is forced to alter its adaptive trajectory, leading to a non-linear evolutionary path.

This thesis aims to investigate how different environment shapes distinct evolutionary consequences for populations, from theoretical aspects focusing on two research areas (i.e., evolutionary computation and virus evolution). Another critical objective is to examine how population endure fitness declines due to environmental shift and how difficult it is for the population to keep track of the moving optimum in a changing world.

The contents are summarized as follows.

Chapter1: Evolving neural networks through bio-inspired parent selection in dynamic environments

Environmental variability often degrades the performance of algorithms designed to capture the global convergence of a given search space. Several approaches have been developed to challenge environmental uncertainty by incorporating biologically inspired notions, focusing on crossover, mutation, and selection. This study proposes a bio-inspired approach called NEAT-HD, which focuses on parent selection based on genetic similarity. The originality of the proposed approach rests on its use of a sigmoid function to accelerate species formation and contribute to population diversity. Experiments on two classic control tasks were performed to demonstrate the performance of the proposed method. The results show

that NEAT-HD can dynamically adapt to its environment by forming hybrid individuals originating from genetically distinct parents. Additionally, an increase in diversity within the population was observed due to the formation of hybrids and novel individuals, which has never been observed before. Comparing two tasks, the characteristics of NEAT-HD were improved by appropriately setting the algorithm to include the distribution of genetic distance within the population. My key finding is the inherent potential of newly formed individuals for robustness against dynamic environments.

Chapter2: Contact-number-driven virus evolution: A multi-level modeling framework for the evolution of acute or persistent RNA virus infection

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, I considered virus evolution based on the host environment, specifically, the effect of the contact history of infected hosts. I 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.

Chapter3: Isolation may select for earlier and higher peak viral load but shorter duration in SARS-CoV-2 evolution

During the COVID-19 pandemic, human behavior change as a result of nonpharmaceutical interventions such as isolation may have induced directional selection for viral evolution. By combining previously published empirical clinical data analysis and multi-level mathematical modeling, I found that the SARS-CoV-2 variants selected for as the virus evolved from the pre-Alpha to the Delta variant had earlier and higher peak in viral load dynamics but a shorter duration of infection. Selection for increased transmissibility shapes the viral load dynamics, and the isolation measure is likely to be a driver of these evolutionary transitions. In addition, I show that a decreased incubation period and an increased proportion of asymptomatic infection are also positively selected for as SARS-CoV-2 mutated to adapt to human behavior (i.e., Omicron variants). The quantitative information and predictions I present here can guide future responses in the potential arms race between pandemic interventions and viral evolution.

Acknowledgement

I would like to express my sincere gratitude to all those who have supported me throughout the course of these studies.

First and foremost, I am deeply grateful to Professor Shinji Nakaoka, my supervisor, for his invaluable guidance, insightful advice, and continuous encouragement. His expertise and dedication have greatly shaped these works. I would also like to thank Assistant Professor Ryo Yamaguchi for his stimulating discussions, constructive feedbacks, and friendly supports throughout my studies. I would like to extend my sincere appreciation to Professor Toshiyuki Nakagaki of the Laboratory of Physical Ethology for his time and effort of my doctoral dissertation. I also would like to thank Prof. Shingo Iwami, Assistant Prof. Hyeongki Park, Assistant Prof. Kwang Su Kim, Assistant Prof. Ryo Komorizono for their great help.

My heartfelt thanks go to the members of our laboratory. I would like to extend special thanks to Keiko Saito (Assistant) and Kozue Yoshimura (Assistant) for their kind and comprehensive supports, both academically and personally, throughout my research and student life. I would like to say big thank you to all the members of Laboratory of Mathematical Biology in Hokkaido University.

I would like to acknowledge the financial support provided by Japan Society for the Promotion of Science (JSPS), which greatly supports my works.

Finally, I show my gratitude for my family. My father Tomimasa Sunagawa, my mother Hisayo Sunagawa, and my brother Tomoyasu Sunagawa.

Chapter1

Evolving neural networks through bio-inspired parent selection in dynamic environments

This study in this chapter was published in J. Sunagawa, R. Yamaguchi, S. Nakaoka, “Evolving neural networks through bio-inspired parent selection in dynamic environments.” *Biosystems.*, 218:104686 (2022).

Introduction

Neuroevolution (NE) is an area of machine learning that develops artificial neural networks (ANNs) using an evolutionary algorithm (EA) or a similar bio-inspired stochastic search/optimization algorithm [1,2]. ANNs originate from biological neural networks in the brain. They are generally composed of interconnected neurons and weights to determine their behavior. EAs evolve ANNs to find suitable network parameters and topology structures inspired by the principles of Darwinian evolution. The concept of EA is to maintain a population of solution candidates through the search space and process a stochastic search that has the advantage of sampling many points simultaneously by selection, crossover, and mutation. Thus, the population can avoid falling to the local optima, leading to global optima. The potential of NE has been widely recognized and applied in fields ranging from quantum physics [3] to video games [4] and chemical reaction networks [5]; which are considered competitive with deep learning [6].

One approach to gaining considerable attention in EA is the neuroevolution of augmenting topologies (NEAT) algorithm [7]. NEAT adopts a framework in which an ANN adapts to an environment by evolution, owing to unique properties represented by “speciation”. Speciation is an idea from biology that each ANN is regarded as an individual in terms of genetic relationships. Genetically close individuals (i.e., the same species) can cross over, while genetically distinct individuals (i.e., different species) cannot. Speciation prevents the population from losing innovative ANNs having near-global convergence in subsequent generations.

In methods such as Genetic Algorithm (GA), speciation often matches individuals to improve the performance of the algorithm. Various

selection rules have already been proposed for parent selection, such as roulette wheel selection, tournament selection, and rank selection. These well-known methods find parents with better fitness [8]. On the other hand, in the field of natural ecology and evolution, various studies have argued that the hybridization process between distinct individuals can significantly affect the evolution of the population [9]. Empirical and theoretical studies in biology have shown that hybridization between ecologically divergent taxa results in a rise in genetic variation, allowing the hybrids to adapt to a novel environment, thus forming a new species [10–12]. The introduction of genetic relationships as a novel parent selection process would improve the performance of GA.

In this work, we propose a new method called NEAT with biological Hybridization for a Dynamic environment (NEAT-HD), which extends NEAT by replacing the parent selection process of NEAT with a bio-inspired operation based on the genetic distance. NEAT-HD allows the population to crossover within a given species group and between genetically distinct species. This property results in the emergence of a hybrid individual, defined as offspring produced from parents of two different species. Despite the low probability of emergence, we demonstrate that a population including both interspecies and intraspecies crossover effectively searches a search space in dynamic environments where the solution changes over time. More specifically, we show that NEAT-HD can achieve the suppression of considerable fitness reduction and diversity preservation, referred to as fitness endurance contributed by genetic diversity when the environment changes dynamically.

The remainder of this paper is organized as follows. Section 2 discusses previous studies on parent selection as a critical attribute of the search for exploration and exploitation in GA and the NE methodology, which includes parent selection and variability for dynamic environments.

In Section 3, we present and discuss the features of NEAT-HD. Section 4 explains our experimental methodology and our testing results and discussion for a simple task. In addition, we examined the performance of NEAT-HD against another more difficult task. Concluding remarks and suggestions for future research are summarized in Section 5.

Related works

Exploration and Exploitation in parent selection

To develop a successful algorithm, EAs, there is a need to address the exploration and exploitation of a search space [13]. The former aims to approach entirely new areas of search space to find better solutions. The latter seeks to visit a search space of previously visited points to sophisticate the solutions already found. As a rule of thumb, maintaining a good ratio between exploration and exploitation is required to establish a practical algorithm, that is, to control operators such as selection of individuals that enable offspring, mutation, and crossover. Parent selection is also among the significant components that influence the balance between exploration and exploitation.

Parent selection in genetic algorithms

Fitness proportional-based methods such as roulette wheel selection, tournament selection, and rank selection [8] are widely used for parent selection. All these operators are based on the common assumption that individuals with higher fitness are more likely to generate offspring than those with lower fitness. Parents are chosen by non-uniform sampling, which depends only on fitness.

Parent selection and diversity preservation

Some algorithms are designed to trigger innovation in search space exploration by utilizing parent selection, allowing crossover between different types of relatively dissimilar solutions. For instance, [14] and [15] proposed algorithms in which one parent was chosen randomly from the population. The second parent was selected from randomly chosen K

individuals and chooses the most dissimilar one. This second parent is often neither the most similar nor the most dissimilar from the first parent. It has been reported to preserve the diversity of the whole population.

NeuroEvolution of augmenting topologies (NEAT)

NEAT is a highly sophisticated algorithm that finds the optimal neural network structures that solve a wide range of tasks [3–6]. NEAT has gained attention in the area of NE because it does not rely on a fixed neural network structure. However, it allows the evolution of its topology structures that perform significantly well. Compared to previous works, NEAT has some efficiency, such as assigning an ID called historical markings that avoids the loss of information embedded in the topology and making a sub-group to prevent an innovative network from being excluded from the population, called speciation.

Encoding

NEAT represents the phenotype directly from the genotype (**Fig. 1**). The genotype contained two lists of genes. A node gene in the genotype corresponds to a node in the network. A connection gene corresponds to a connection between two nodes in the network. The historical markings assigned to each gene make them distinguishable. Mutation and crossover occur by adding a change to the list of genes.

Crossover

When comparing two individuals, the use of historical markings helps to distinguish an identical gene (**Fig. 2**). Two lists of genes are lined up, and identical genes are regarded as matching genes. Otherwise, the

non-identical genes are regarded as disjoint or excessive, producing a difference between the two individuals. The matching genes are always inherited from the parents. Disjoint and excess genes are inherited randomly. In this way, an offspring is generated that avoids the competing convention problem. The competing convention problem is that the two networks have an identical topology structure, although the genotype representation is different, making them incompatible for crossover.

Mutation

Two types of mutations occur in the NEAT (**Fig. 3**). In the add node mutation, a new node splits an existing old connection and re-assigns new historical markings to new split connections (**Fig. 3, top**). The old connection was disabled. A new connection connects the two existing old nodes in the add connection mutation. It assigns a historical marking to the new connection (**Fig. 3, bottom**). These mutations make the individual innovative, which has never been seen previously. Although innovative individuals are likely to die early in subsequent generations, speciation saves innovation by dividing the whole population.

Speciation

Some new connections made by a mutation are added to an individual that might become innovative. This individual is often likely to die in subsequent generations because the weight of the newly added structure is not maturely optimized for the network. A niching scheme, called speciation, was introduced to protect these networks and expects survival. Speciation divides the population into non-overlapping sets of similar individuals based on their shared evolutionary history.

To divide the population into species such that similar structures are in the same species, speciation uses a compatibility function to determine whether two individuals should be in the same species or not. Owing to the historical marking, each gene is comparable to the difference between two individuals. When comparing the genotypes of two selected individuals, identical or non-identical genes were detected. Some genes are in excess or disjoint. These genes are counted as their difference, converted to the compatibility distance to express the relative similarity of the two selected networks. The distance δ between two individuals is a simple linear combination of the number of excess E and disjoint D genes and, as well as the average weight differences of matching genes $\bar{W}$, including disabled genes.

$$\delta = \frac{c_1 E}{N} + \frac{c_2 D}{N} + c_3 \bar{W}, \quad (1)$$

where c_1 , c_2 , and c_3 are the adjusted coefficients, and N is the number of genes in the larger genome. The distance measure δ allows us to use a compatibility threshold to divide the group into species. In each generation, the individuals were sequentially placed into the species. Suppose the distance between the focal individual and the representative individual of the species is under the compatibility threshold. In that case, the focal individual is assigned as a member of the species. Otherwise, if the distance between the focal one and the representative individual of every species is greater than the compatibility threshold, a new species originates with the focal individual as its representative.

Having formed a new species, stagnation promotes species extinction in which fitness is not updated within the species for some generations. Apart from stagnation, species can also be extinct when none of the individuals are assigned to that species because of the compatibility threshold.

Speciation and diversity in EC

In EC, speciation restricts access to other species, thus maintaining diversity. As in the case of NEAT, the notion of preserving diversity is helpful because it prevents pre-optimized individuals from being excluded, which might be a novel solution later. Speciation strategies have been developed in algorithms derived from NEAT, such as grouping the population-based hidden neurons [16] or clades sharing a common ancestor [17]. To develop diversity, Novelty Search (NS) performed well [18]. Its primary strategy is not to set a fitness function related to the domain field but rather to set it related to the individual's dissimilarity from the population. NS has enabled the development of new quality diversity algorithms (QDs) to generate large collections of diverse high-performance solutions [19,20]. Additionally, these developments appear to be related to open-endedness, which attempts to understand processes that continue to generate exciting artifacts indefinitely in artificial life (Alife) [21,22]. One recent study incorporates the idea of coevolution, in which both individuals and environments evolve interactively and permanently [23,24]. While NS or QD can actively produce novel individuals, our work aims to incorporate an idea from a biological context. Although some controversy exists that novel algorithms based on metaphors of natural processes lead the area away from scientific rigor [25], still, special attention has been paid to natural computations that copy the inherent advantages of different behaviors observed in biological systems. Bio-inspired computation has been explored to solve complex EC tasks [26].

Challenge to environmental variability

Many real-world optimization problems are subject to changing

environments. They have become one of the most active areas in the EC field for decades [27,28]. That is, an algorithm not only finds the optimum in the vast search space but also tracks it over time. Various techniques have been developed to adapt to a dynamic environment, such as memory-based [29–31], prediction-based [32–34], and diversity-based [35,36]. Memory-based methods retrieve archived historical information aimed at convergence. Prediction-based methods collect archives from the past that the population passes along and integrate them into various predictions of a changing optimum. Diversity-based methods enhance population variations for some individuals to address environmental changes. From the perspective of diversity-based methods, one example is that an extension algorithm of NEAT [37] which is capable of maintaining species size based not only on the previous generation but also the arbitral past generations, has shown significant performance over NEAT. The strategy used by the authors involved maintaining the number of species and allowing long-life species to form when the fitness of individuals dramatically changes between generations. As a diversity-based method, we challenge environmental variability by maintaining diversity by choosing parents according to their similarities.

Materials and methods

Approach: distance-based parent selection

We propose NEAT-HD to employ a parent selection operator, allowing genetically distinct individuals to be chosen as parents (**Fig. 4A**). From a biological perspective, genetically close ones (e.g., creatures belonging to the same species) are likely to crossover and leave offspring. Conversely, the more distinct the biological relationship between two individuals, the less likely they are to be parents. Our approach generates a small proportion of hybrid offspring. We demonstrate that NEAT-HD preserves innovation through the speciation scheme and also makes the population robust to dynamic environmental changes by hybrid offspring. Algorithm 1 represents the pseudocode of the novel distance-based parent selection (**Fig. 4B**).

The algorithm selects each species, excluding low-ranked individuals from the population (**Step 2 in Algorithm 1**). At each iteration, a set of first parent is generated within the species. For each parent, the candidate of the second parent is selected through the population (**Step 4**). In **Step 5**, genetic distances D between individual i and j are calculated [38] as follows:

$$D(i, j) = c_1 \frac{G}{N'+1} + c_2 \bar{W}, \quad (2)$$

where G is the number of non-matching genes of the two individuals, and $\bar{W}$ is the average weight difference of matching genes between individual i and j . The normalization constant N' is the number of added genes to the minimal network of two individuals. c_1 and c_2 are the gene and weight coefficients, respectively. In **Step 6**, GeneticIncompatibility I is calculated using the genetic distance using (2). Genetic incompatibility refers to the difficulties faced by the offspring by mating two genetically

distinct individuals. When genetic incompatibility increases, the probability of creating offspring decreases. Herein, we used a sigmoidal function to associate the genetic distance with genetic incompatibility, which reports a high likelihood of speciation even while allowing hybridization [39].

$$I(i, j) = \frac{1}{1 + \exp(-a(-b + D))}, \quad (3)$$

where a and b are parameters that determine the steepness and coordination of incompatibilities of 0.5, respectively. Finally, in **Step 7**, the probability of being chosen as a parent is given by MatingProbability P :

$$P(i, j) = 1 - I, \quad (4)$$

Using these metrics from (2) to (4), the mating pair is chosen from the population based on the above mating probability.

We expect that utilizing this function will make hybrid offspring more likely to maintain diversity, stability, and reactivity within the population in changing environments.

The NEAT algorithm adopts speciation such that the genetic distance between the representative of a species and each member of the species is under the compatibility threshold defined by an optimizer. This indicates that the genetic distance between two individuals within a species might be greater than the compatibility threshold. Unlike this previous approach, such as implicitly assuming the possibility of mating between genetically distinct individuals, we test an explicit approach that calculates the genetic distance between individuals. Therefore, it might have a chance of crossover between different species. NEAT-HD is tested in two classic control problems: single-cart pole swing-up and lunar lander experiments in Sections 4.

Results

We conducted two experiments, the cart-pole swing-up, and the lunar lander tasks, using the same parameters (**Table 1**). NEAT was executed for 50 runs of 300 generations, where 128 agents were the total population using the code from [38]. NEAT-HD used the parameters $a = 0.2$, $b = 1.3$ in (3).

Cart-pole swing-up

Choosing a domain with simple input and a relatively small network architecture is desirable to see how the population includes individuals from the same species and different species. We tested the cart-pole swing-up task as our first examination to validate the effectiveness of NEAT-HD, one of the most classic and straightforward control problems.

The goal of the cart-pole swing-up task is to balance the pole upright by moving the cart left and right (**Fig. 5A**). The trained ANN has to add nodes by mutation because the cart encounters nonlinear phases such as swinging up the pole or balancing the pole. The cart has six inputs, including the information of the cart and the pole; two of them are the position and velocity of the cart, three are the sine and cosine of the angle and angular velocity of the pole, and the other is the bias term. This setup is identical to that employed by [38].

Fitness is evaluated by how long the pole is straight up from the cart and how close the cart is to the optimal position in a fixed time. The fitness is expressed as follows:

$$F = F_{pole} \times F_{cart}, \quad (5)$$

$$F_{pole} = \frac{\cos \theta + 1}{2}, \quad (6)$$

$$F_{cart} = \frac{x-(optimal\ position)}{2.4} \times \frac{\pi}{2}, \quad (7)$$

where θ is the angle from the perpendicular line of the ground to the pole, x is the coordinate of the cart. Over 300 generations, the dynamic change is set at the pre-specified 100th and 200th generations. The *optimal position* at $x = 0$ remains until the first change to $x = -1.5$, and finally, at the second change transition to $x = 1.5$ (**Fig. 5B**). We intend to change the magnitude of the transition to observe the differences. The second change was more severe than the first change because the magnitude was more significant. Thus, the second change seems to greatly influence the population, which is close to convergence significantly. Note that the magnitudes of the first and second changes were comparable. The first change has little effect on the second change due to the long-generation pass after the first change, in which the population can re-organize for convergence.

Parents are regarded as the same or different species by genetic distance, proceeding with the algorithm. Here, utilizing a threshold can determine the species identity, called the compatibility threshold, which was initially introduced by [7]. When the genetic distance between a parent is under the compatibility threshold, the two individuals are regarded as the same species. As the compatibility threshold directly affects the parent sameness, we changed the value of the compatibility threshold to some extent and explored the difference within the range.

The following three performance metrics were employed to check how the algorithm could endure a dynamic environment. We utilized and modified the performance measures developed by [40] to this specific study for that purpose.

-Accuracy indicates how the population can suppress dynamic changes that occur suddenly. This metric can be measured by recording the best-fitness

observed from the population at generation t when the optimal position of the cart changes. The accuracy at generation t is defined as:

$$Accuracy_t = \frac{Best_i - Min_i}{Max_i - Min_i}, \quad (8)$$

where $Best_i$ is the best-fitness at generation t , Max_i is the possible max-fitness value at generation t in the task, Min_i is the possible worst-fitness value at generation t in the task. The accuracy ranges from 0 to 1, where 1 is the best possible value.

-Stability is the ratio of the number of iterations that successfully recovered after the dynamic change at generation t . It is counted as “recovered” when the mean value of the accuracy after the dynamic change is greater than or equal to the mean value of the accuracy before the dynamic change. The stability at generation t is defined as:

$$Stability_t = \frac{\sum_k^{(total\ number\ of\ iterations)} (stability_{iteration=k})}{(total\ number\ of\ iterations)}, \quad (9)$$

$$Stability_t = \begin{cases} 1, & \text{if } \frac{Best_{t \leq i \leq t+99}(accuracy_i)}{Mean_{t-100 \leq i \leq t-1}(accuracy_i)} \geq 1, \\ 0, & \text{otherwise} \end{cases}, \quad (10)$$

where $Mean_{t-100 \leq i \leq t-1}(accuracy_i)$ is the mean value of the accuracy between generations $t-100$ and $t-1$. This value corresponds to the accuracy criterion before the dynamic change.

$Best_{t \leq i \leq t+99}(accuracy_i)$ is the best accuracy value between generation t and $t+99$. This value corresponds to the best value that an algorithm can increase from dynamic change. For both

$Mean_{t-100 \leq i \leq t-1}(accuracy_i)$ and $Best_{t \leq i \leq t+99}(accuracy_i)$, 100 generations before and after the change are considered because of the duration of 100 generations between two changes at the 100th and 200th

generations.

-Reactivity represents the speed at which the population adjusts their strategies and recovers their fitness from the environmental change at generation t . An assumption in this metric is that the iteration is regarded as recovered in the stability metric. Measurements are applied, such as recording how many generations take to catch up with the fitness before the environmental change, the mean value of the accuracy. The reactivity at generation t is defined as

$$Reactivity_t = Min \left\{ (t' - t) \left| \frac{accuracy_i}{Mean_{t-100 \leq i \leq t-1}(accuracy_i)} \geq 1 \right. \right\} \cup \{maxgen - t\}, \quad (11)$$

where $t, t' \in \mathbb{N}$ and $t < t' \leq maxgen$, with $maxgen$ referring to the total number of generations. Lower values indicate a better and faster reaction to changes.

Application results of NEAT-HD on cart-pole swing-up task

First, it is necessary to verify that hybrid individuals are born, as shown in **Fig. 6**. We can confirm that hybrids were formed more often in NEAT-HD than in NEAT in the low region of the compatibility threshold. This allows mating between genetically distinct parents. **Fig. 7** shows an overview of the best-fit values along each compatibility threshold. We confirmed that the fitness dropped in both algorithms (NEAT and NEAT-HD) at the 100th and 200th generations, where the optimal position of the cart-pole changes. In addition, it is observed that the drop at the 200th generation is more significant than that at the 100th generation because of the significant dynamic change at the 200th generation. When the compatibility threshold is 0.5, and 0.9, the best fitness of NEAT-HD is slightly below that of NEAT throughout the generation. As the

compatibility threshold increased, both algorithms exhibited fewer differences. In the high region of the compatibility threshold, the essential difference between the two algorithms decreases, thus resulting in similar trajectories.

We also investigated the three metrics to test the robustness of the proposed algorithm to dynamic environmental changes (**Fig. 8 and Table 2**). The accuracies at the 100th and 200th generations are illustrated in **Fig. 8A and B**, respectively. When a small change occurred at the 100th generation, no significant difference was observed between both algorithms. At the 200th generation, on the other hand, some significance was found at compatibility thresholds of 0.5, 0.9, and 1.3 ($p < 0.05$; Welch's t-test). The difference between the two algorithms appears in the lower region of the compatibility threshold, demonstrating that NEAT-HD endures the dynamic change. These results indicate a trade-off between the best-fitness trajectory and endurance against the change. In the case of stability, it is determined whether the best-fitness recovers or remains after the dynamic change (**Table 2**). Both algorithms have similar stability because the best-fitness recovers after the dynamic change. More specifically, at the lower region of the compatibility threshold, NEAT-HD showed higher stability. However, NEAT was higher at the high region of compatibility threshold as well. In the case of reactivity, it detected the number of generations required to recover; the shorter the generations that need to recover, the higher their reactivity (**Fig. 8C and D**). At both the 100th and 200th generations, in the low region of the compatibility threshold, significant differences are found so that NEAT-HD only requires a short duration to recover from dynamic change.

To determine the diversity of the population through simulations, the number of species was calculated, as shown in **Fig. 9**. At a compatibility threshold of 0.5, the species number is significantly affected

when dynamic changes occur, as the species number falls suddenly. When the threshold was increased, the effect was relaxed. At the compatibility thresholds of 1.3 and 1.7, the decrease in species number stopped after some generations in response to the change, such as the 160th and 260th generations. This is identical to the stagnation duration of approximately 60 generations (see **Table 1** for the parameters used) that species hold their individuals when fitness has not been updated. In some cases (e.g., compatibility threshold of 1.3, 1.7), NEAT-HD contributed to maintaining the diversity compared to NEAT, as NEAT-HD holds more species in terms of the mean value. In the other region, it is difficult to say that the contribution of NEAT-HD is greater than that of NEAT.

The numbers of newly formed and extinct species are shown in **Fig. 10**. In the lower region of the compatibility threshold, the number of newly formed and extinct species was higher in NEAT-HD than in NEAT. As the threshold increases, the difference decreases toward disappearance. **Fig. 11** shows another consequence for observing a significant difference in the species' duration between NEAT-HD and NEAT at the compatibility threshold of 0.5, 0.9, 1.3, and 1.7 ($p < 0.001$; Welch's t-test). This result implies that the cycles of the current model from formation to extinction are shorter due to the high rate of species formation and extinction. **Fig. 8C and D**, 10, and 11 show the sigmoid effect proposed in [39]. More specifically, as a new species is generated, the species becomes extinct shortly after that, and another new species is formed. This short cycle constantly generates new species during the speciation process of NEAT-HD. As species formation is rapid in NEAT-HD, species extinction is fast, resulting in a decrease in the species number in the compatibility threshold of 0.5 (**Fig. 9 A**).

Fig. 12 depicts the number of novel networks observed through the generations, which were not observed in previous generations. We defined

“novel” as to whether the topology of the focal network was ever seen before in the whole prior generations. The weights on the edges between nodes are neglected to determine the pure difference in the topological architecture. Over generations, the number of novel networks increased in both models. In the last 300th generation, NEAT-HD was significantly higher than NEAT at the compatibility threshold of 0.5, 0.9, 1.3, and 1.7 ($p < 0.05$; Welch's t-test), indicating the effect of forming hybrid individuals from different species.

From the overall results, as the compatibility threshold decreases, the properties of NEAT-HD have been increasingly highlighted, such as the formation of hybrid offspring, the acceleration the emergence of new species, and the extinction with short species duration, all of which influence the endurance in terms of accuracy. We note one disadvantage of NEAT-HD, which affects the best-fitness to decrease to some extent, as shown in **Fig. 7**.

Lunar-Lander

Difficult domains include more inputs that are given more information about the environment. Another concern is that the domains that can be dynamic are desired in this study, where the optimal position varies across generations. We adopted a Lunar-Lander task as a challenge, a control task often used in machine learning, and a toolkit from Open AI [41]. In this task, a lander aiming to land on the moon's surface in a realistic physics simulation was controlled by ANNs (**Fig. 13A**).

The task is as follows. The lander is set at a certain height pre-specified and starts to land smoothly using three engines. The engine can adjust its position above the surface and stability of the rotation, with one throttle pointing downward and the other two pointing in the left and right

directions. The lander has nine inputs, including the information of the lander and whether the lander is attached to the surface; four of them are the x- and y-axes of the position and velocity, two of which are the angle and angular velocity of the lander, two are for attachment to the surface, and the other is the bias term. The possible output actions are to fire the left orientation engine, fire the main orientation engine, fire the right orientation engine, and do nothing. The task ends if the lander crushes, receiving a penalty reward or rest for a while receiving an additional reward.

Fitness is given as the sum of the rewards using an addition and subtraction method. As mentioned above, if the lander is crushed, the minus reward is assigned, and if the lander rests on the pad for a while, the plus reward is assigned. Additionally, some elements give the lander better rewards through the landing, such as maintaining a stable posture and using less engine fuel. Around the landing pad, the presence of an attachment on the ground provides a reward. We set this task to be dynamic as well as the cart-pole task by changing the optimal position for some generations. By default, the optimal position, the center of the landing pad, is at coordinates (0, 0). We changed the position to $(-5, 0)$ at the 100th generation and $(5, 0)$ at the 200th generation, respectively, over 300 generations (**Fig. 13B**). Similar to the cart-pole task, the second change was set to be severe. We tested this task, which was more difficult because of the increased input data; thus, the algorithm must find the optimal solution in the expanded search space compared to the cart-pole task.

Application results of NEAT-HD on the Lunar-Lander task

The number of hybrid individuals is shown in **Fig. 14**. Similar to the cart-pole task, we confirmed that hybrids were formed more often in

NEAT-HD than in NEAT in the low region of the compatibility threshold.

Over 50 iterations, the trajectory of the best-fitness is plotted in **Fig. 15**. The mean of the best-fitness seemed to show a similar result, as shown in **Fig. 7**. In the lower region of the compatibility threshold, NEAT performs negligibly better than NEAT-HD. As the threshold increased, the difference decreased.

The accuracy shown in **Fig. 16A and B** indicates that some significant differences were observed at the 100th generation (compatibility threshold of 0.9, and 1.7, $p < 0.05$; Welch's t-test). In addition, at the 200th generation, significant differences were observed (compatibility threshold of 2.1, $p < 0.05$; Welch's t-test). In all cases where differences were significant, the mean value of best-fitness in NEAT was higher than in NEAT-HD. A trade-off observed in the cart-pole task was not apparent here. The stabilities were all 1.0 for each compatibility threshold and the 100th and 200th generations (Table 3). **Fig. 16C and D** illustrates the reactivity at the 100th and 200th generations. At the 200th generation, NEAT-HD showed high reactivity, which was significant compared to NEAT at the compatibility threshold of 0.5, 0.9, and 1.7 ($p < 0.05$; Welch's t-test). These results indicate that NEAT-HD exhibited low endurance in terms of accuracy when dynamic changes occurred, although it recovered quickly in reactivity.

The species number in **Fig. 17** as a diversity metric shows similar results for the two tasks compared with **Fig. 9**. At the compatibility threshold of 0.5, as shown in **Fig. 17A**, the number decreases soon after a sudden rise, similar to **Fig. 9A**. **Fig. 18** illustrates the total number of species and the number of extinctions. **Fig. 18** implies that NEAT-HD generates more species and extinctions, especially in the low region of the compatibility threshold. As the threshold increased, both algorithms overlapped, showing minor differences. **Fig. 19** shows the duration of each

species. At the low region of compatibility threshold (e.g., 0.5, and 0.9), significant differences were found in the sense that species generated by NEAT-HD have a shorter duration until extinction than NEAT ($p < 0.001$; Welch's t-test). Similar to the cart-pole task, a short cycle from the formation of new species to extinction is observed. In addition, at a compatibility threshold of 0.5 (**Fig. 17A**), the decrease in the species number can be attributed to the rapid formation of species and extinction.

The novel networks that emerged over generations are summarized in **Fig. 20**. Comparing the last 300 generations, hybrid offspring are generated significantly more often at the low region of compatibility threshold of 0.5, 0.9, and 1.7 ($p < 0.05$; Welch's t-test).

In this task, from the perspective of species-cycle, the short duration from species formation to extinction affects the diversity, thus expecting to help the endurance against dynamic change. However, the accuracy metrics show the opposite results. We need to determine the distribution of the population diversity that crucially affects the fitness evaluation.

Conclusion

This article proposes NEAT-HD, an algorithm that adapts a dynamic environment with a population including offspring from different species. This method attempts to preserve genetic/species diversity by forming hybrid offspring between genetically distinct parents to adapt to environmental changes. An essential advantage of NEAT-HD is that it employs the sigmoid function to determine the probability of mating, as supported by a study in which speciation was accelerated [39]. We argue that some offspring from distinct parents can help endure environmental changes.

We compared NEAT-HD with NEAT for two control tasks: (1) a single cart-pole swing-up task and (2) a lunar lander task. We tested how NEAT-HD behaves in these tasks under various compatibility thresholds. Our results showed two different results. One showed endurance in terms of accuracy when a dynamic change occurred. In addition, the contribution to the diversity was observed in the sense that the production of new hybrid individuals drastically accelerated the formation of species. However, one of the major pitfalls of this study is that the best-fitness is relatively low in NEAT-HD compared with NEAT at the complementary expense for endurance in terms of accuracy. Moreover, in some cases, diversity did not influence endurance in terms of accuracy. Our continued efforts will be made to implement and evaluate new approaches derived from this study.

In future work, it will be necessary to combine NEAT-HD with other diversity-preserving methods, especially in dynamic environments. More specifically, other methods, such as NS or QD, are subject to testing within the same task. For example, the method proposed by [23] and [24] uses a mutation-only NS that spreads behavioral diversity. Our distance-based scheme can also be applied. Thus, diversity measurements must be

unified in the study to compare them appropriately. Furthermore, it is interesting to employ the idea of genetic incompatibility with HyperNEAT [42], an extended version of NEAT. HyperNEAT can treat very large neural network topologies. Thus the impacts on distance-based crossover are expected to be significant. Another concern is the use of our proposed algorithm in the area of multimodal optimization problems. In multimodal optimization problems, approaches that utilize species such as species conservation [43] are thought of an effective way to capture many global optimums. Our approach also has an expectation of improving the population diversity so that the population can handle multiple solutions within a single run. For future improvement of our proposed method, we must consider the criterion of the algorithm. In this study, genetic distance represents the difference between the two ANNs. Other metrics, such as behavioral representation, might be an alternative option that contributes to the population diversity used in the field of evolutionary robotics [44].

Code availability

The source code is available in Python using NEAT-HD (<https://github.com/j-s9vcp/NEAT-HD>).

References

1. Floreano D, Dürri P, Mattiussi C. Neuroevolution: from architectures to learning. *Evol Intell.* 2008;1: 47–62. doi:10.1007/s12065-007-0002-4
2. Papavasileiou E, Cornelis J, Jansen B. A Systematic Literature Review of the Successors of “NeuroEvolution of Augmenting Topologies.” *Evol Comput.* 2021;29: 1–73. doi:10.1162/evco_a_00282
3. Aaltonen T, Adelman J, Akimoto T, Álvarez González B, Amerio S, Amidei D, et al. Observation of Electroweak Single Top-Quark Production. *Phys Rev Lett.* 2009;103: 092002. doi:10.1103/PhysRevLett.103.092002
4. Stanley KO, Bryant BD, Miikkulainen R. Real-Time Neuroevolution in the NERO Video Game. *IEEE Trans Evol Comput.* 2005;9: 653–668. doi:10.1109/TEVC.2005.856210
5. Dinh HQ, Aubert N, Noman N, Fujii T, Rondelez Y, Iba H. An Effective Method for Evolving Reaction Networks in Synthetic Biochemical Systems. *IEEE Trans Evol Comput.* 2015;19: 374–386. doi:10.1109/TEVC.2014.2326863
6. Stanley KO, Clune J, Lehman J, Miikkulainen R. Designing neural networks through neuroevolution. *Nat Mach Intell.* 2019;1: 24–35. doi:10.1038/s42256-018-0006-z
7. Stanley KO, Miikkulainen R. Evolving Neural Networks through Augmenting Topologies. *Evol Comput.* 2002;10: 99–127. doi:10.1162/106365602320169811
8. Thierens D, Goldberg D. Convergence models of genetic algorithm selection schemes. In: Davidor Y, Schwefel H-P, Männer R, editors. *Parallel Problem Solving from Nature — PPSN III.* Berlin, Heidelberg: Springer Berlin Heidelberg; 1994. pp. 119–129. doi:10.1007/3-540-58484-6_256
9. Mallet J. Hybrid speciation. *Nature.* 2007;446: 279–283. doi:10.1038/nature05706

10. Grant PR, Grant BR. How and Why Species Multiply: The Radiation of Darwin's Finches. Princeton University Press; 2020.
doi:10.2307/j.ctvx5wc23
11. Grant PR, Grant BR. Hybridization increases population variation during adaptive radiation. *Proc Natl Acad Sci*. 2019;116: 23216–23224.
doi:10.1073/pnas.1913534116
12. Yamaguchi R, Otto SP. Insights from Fisher's geometric model on the likelihood of speciation under different histories of environmental change. *Evolution*. 2020;74: 1603–1619.
13. Črepinšek M, Liu S-H, Mernik M. Exploration and exploitation in evolutionary algorithms: A survey. *ACM Comput Surv*. 2013;45: 1–33.
doi:10.1145/2480741.2480752
14. Drezner Z, Marcoulides GA. A Distance-Based Selection of Parents in Genetic Algorithms. *Metaheuristics: Computer Decision-Making*. Boston, MA: Springer US; 2003. pp. 257–278. doi:10.1007/978-1-4757-4137-7_12
15. Farias RGG, De Magalhaes CS. Parent Selection Strategies in Niching Genetic Algorithms. 2018 IEEE Congress on Evolutionary Computation (CEC). Rio de Janeiro: IEEE; 2018. pp. 1–8.
doi:10.1109/CEC.2018.8477847
16. Hadjiivanov A, Blair A. Complexity-based speciation and genotype representation for neuroevolution. 2016 IEEE Congress on Evolutionary Computation (CEC). Vancouver, BC, Canada: IEEE; 2016. pp. 3092–3101.
doi:10.1109/CEC.2016.7744180
17. Knapp JS, Peterson GL. Natural Evolution Speciation for NEAT. 2019 IEEE Congress on Evolutionary Computation (CEC). Wellington, New Zealand: IEEE; 2019. pp. 1487–1493.
doi:10.1109/CEC.2019.8790153
18. Lehman J, Stanley KO. Abandoning Objectives: Evolution Through the Search for Novelty Alone. *Evol Comput*. 2011;19: 189–223.

doi:10.1162/EVCO_a_00025

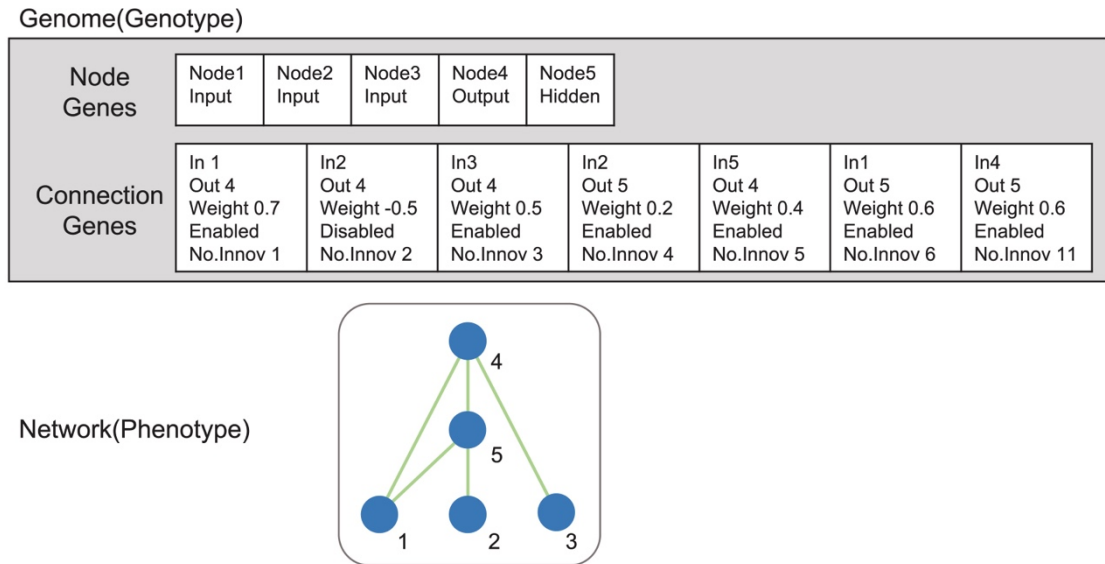
19. Cully A, Clune J, Tarapore D, Mouret J-B. Robots that can adapt like animals. *Nature*. 2015;521: 503–507. doi:10.1038/nature14422
20. Pugh JK, Soros LB, Stanley KO. Quality Diversity: A New Frontier for Evolutionary Computation. *Front Robot AI*. 2016;3. doi:10.3389/frobt.2016.00040
21. Packard N, Bedau MA, Channon A, Ikegami T, Rasmussen S, Stanley KO, et al. An Overview of Open-Ended Evolution: Editorial Introduction to the Open-Ended Evolution II Special Issue. *Artif Life*. 2019;25: 93–103. doi:10.1162/artl_a_00291
22. Taylor T, Bedau M, Channon A, Ackley D, Banzhaf W, Beslon G, et al. Open-Ended Evolution: Perspectives from the OEE Workshop in York. *Artif Life*. 2016;22: 408–423. doi:10.1162/ARTL_a_00210
23. Brant JC, Stanley KO. Benchmarking open-endedness in minimal criterion coevolution. *Proceedings of the Genetic and Evolutionary Computation Conference*. Prague Czech Republic: ACM; 2019. pp. 72–80. doi:10.1145/3321707.3321756
24. Brant JC, Stanley KO. Diversity preservation in minimal criterion coevolution through resource limitation. *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. Cancún Mexico: ACM; 2020. pp. 58–66. doi:10.1145/3377930.3389809
25. Sörensen K. Metaheuristics—the metaphor exposed. *Int Trans Oper Res*. 2015;22: 3–18. doi:10.1111/itor.12001
26. Del Ser J, Osaba E, Molina D, Yang X-S, Salcedo-Sanz S, Camacho D, et al. Bio-inspired computation: Where we stand and what’s next. *Swarm Evol Comput*. 2019;48: 220–250. doi:10.1016/j.swevo.2019.04.008
27. Jin Y, Branke J. Evolutionary Optimization in Uncertain Environments—A Survey. *IEEE Trans Evol Comput*. 2005;9: 303–317. doi:10.1109/TEVC.2005.846356

28. Nguyen TT, Yang S, Branke J. Evolutionary dynamic optimization: A survey of the state of the art. *Swarm Evol Comput.* 2012;6: 1–24. doi:10.1016/j.swevo.2012.05.001
29. Chi-Keong Goh, Kay Chen Tan. A Competitive-Cooperative Coevolutionary Paradigm for Dynamic Multiobjective Optimization. *IEEE Trans Evol Comput.* 2009;13: 103–127. doi:10.1109/TEVC.2008.920671
30. Jiang S, Yang S. A Steady-State and Generational Evolutionary Algorithm for Dynamic Multiobjective Optimization. *IEEE Trans Evol Comput.* 2017;21: 65–82. doi:10.1109/TEVC.2016.2574621
31. Wang Y, Li B. Multi-strategy ensemble evolutionary algorithm for dynamic multi-objective optimization. *Memetic Comput.* 2010;2: 3–24. doi:10.1007/s12293-009-0012-0
32. Jiang M, Huang Z, Qiu L, Huang W, Yen GG. Transfer Learning-Based Dynamic Multiobjective Optimization Algorithms. *IEEE Trans Evol Comput.* 2018;22: 501–514. doi:10.1109/TEVC.2017.2771451
33. Muruganantham A, Tan KC, Vadakkepat P. Solving the IEEE CEC 2015 Dynamic Benchmark Problems Using Kalman Filter Based Dynamic Multiobjective Evolutionary Algorithm. In: Lavangnananda K, Phon-Amnuaisuk S, Engchuan W, Chan JH, editors. *Intelligent and Evolutionary Systems*. Cham: Springer International Publishing; 2016. pp. 239–252. doi:10.1007/978-3-319-27000-5_20
34. Aimin Zhou, Yaochu Jin, Qingfu Zhang. A Population Prediction Strategy for Evolutionary Dynamic Multiobjective Optimization. *IEEE Trans Cybern.* 2014;44: 40–53. doi:10.1109/TCYB.2013.2245892
35. Deb K, Rao N. UB, Karthik S. Dynamic Multi-objective Optimization and Decision-Making Using Modified NSGA-II: A Case Study on Hydro-thermal Power Scheduling. In: Obayashi S, Deb K, Poloni C, Hiroyasu T, Murata T, editors. *Evolutionary Multi-Criterion Optimization*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. pp. 803–817. doi:10.1007/978-3-540-70928-2_60

36. Li C, Yang S. A General Framework of Multipopulation Methods With Clustering in Undetectable Dynamic Environments. *IEEE Trans Evol Comput.* 2012;16: 556–577. doi:10.1109/TEVC.2011.2169966
37. Krčah P. Effects of Speciation on Evolution of Neural Networks in Highly Dynamic Environments. In: Hamadi Y, Schoenauer M, editors. *Learning and Intelligent Optimization*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. pp. 425–430. doi:10.1007/978-3-642-34413-8_39
38. Gaier A, Ha D. Weight Agnostic Neural Networks. *arXiv*; 2019. doi:10.48550/ARXIV.1906.04358
39. Yamaguchi R, Iwasa Y. A tipping point in parapatric speciation. *J Theor Biol.* 2017;421: 81–92. doi:10.1016/j.jtbi.2017.03.018
40. Weicker K. Performance Measures for Dynamic Environments. In: Guervós JJM, Adamidis P, Beyer H-G, Schwefel H-P, Fernández-Villacañas J-L, editors. *Parallel Problem Solving from Nature — PPSN VII*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002. pp. 64–73. doi:10.1007/3-540-45712-7_7
41. Brockman G, Cheung V, Pettersson L, Schneider J, Schulman J, Tang J, et al. OpenAI Gym. *arXiv*; 2016. doi:10.48550/ARXIV.1606.01540
42. Stanley KO, D’Ambrosio DB, Gauci J. A Hypercube-Based Encoding for Evolving Large-Scale Neural Networks. *Artif Life.* 2009;15: 185–212. doi:10.1162/artl.2009.15.2.15202
43. Li J-P, Balazs ME, Parks GT, Clarkson PJ. A Species Conserving Genetic Algorithm for Multimodal Function Optimization. *Evol Comput.* 2002;10: 207–234. doi:10.1162/106365602760234081
44. Mouret J-B, Doncieux S. Encouraging Behavioral Diversity in Evolutionary Robotics: An Empirical Study. *Evol Comput.* 2012;20: 91–133. doi:10.1162/EVCO_a_00048

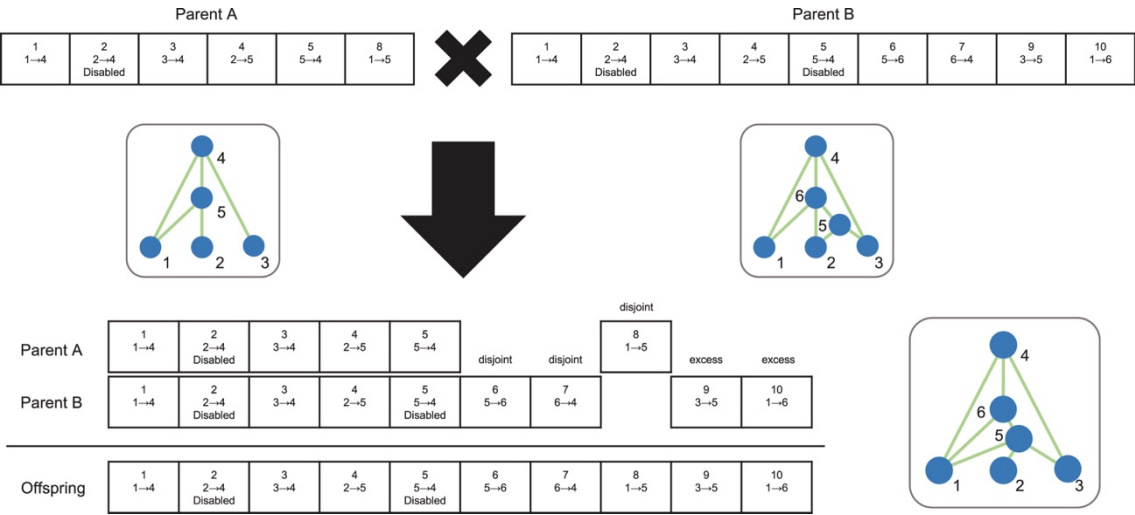
Figures

Figure 1: Example of mapping from genotype to phenotype.



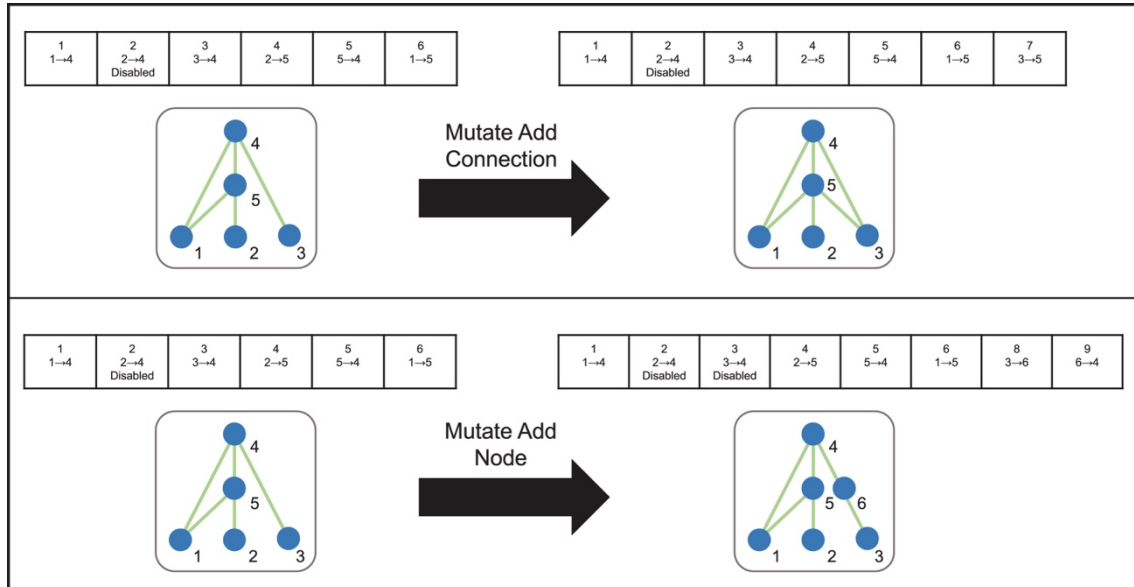
For one genotype, two genes were listed: node genes and connection genes. The numbers (1, 2, ...) and the type (Input, Output, Hidden) are assigned in the node genes. In the connection nodes, the direction of the connection (e.g., In 1–4), and the weight are assigned. Additionally, the genes allocated as “Disabled” were not expressed in the phenotype. The historical markings are assigned to each gene to identify each gene (e.g., No.innov 1).

Figure 2: An example of the crossover between two individuals (parent A and parent B).



The top number of the genes is the historical markings. Although the structures of the individuals are different, each gene is aligned. Thanks to historical markings, the identical genes, called matching genes, are aligned together. Matching genes are directly inherited to offspring. For the non-identical genes such as disjoint and excess are randomly inherited to offspring. The offspring illustrated is an example when all genes are inherited.

Figure 3: Two types of mutations, add connection mutation (top) and add node mutation (bottom).



For the add connection mutation, a new connection is added between selected two nodes, and a new historical marking is assigned to the new connection. For the add node mutation, a new node splits the already-existing connection into two new connections. The new two connections are assigned new historical markings, and the old connection is re-defined as “Disabled”.

Figure 4: The concept and process of biological hybridization.

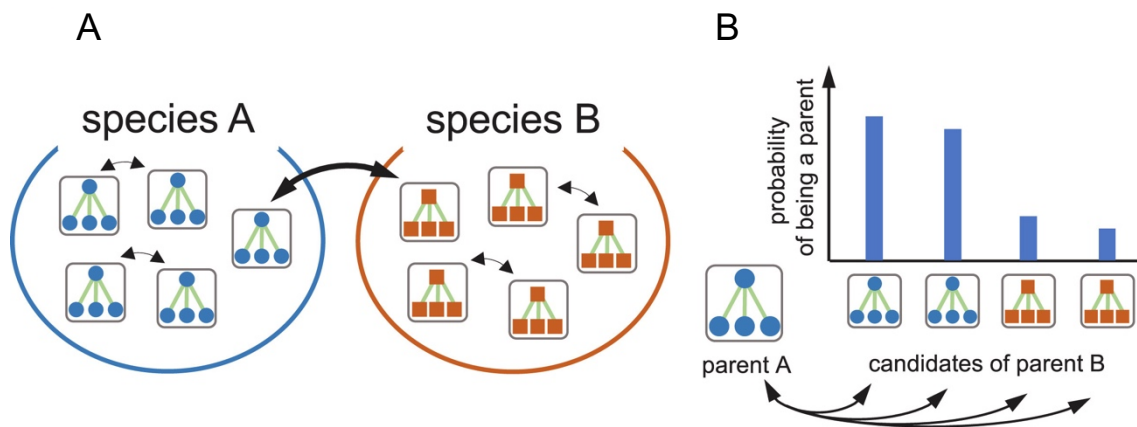
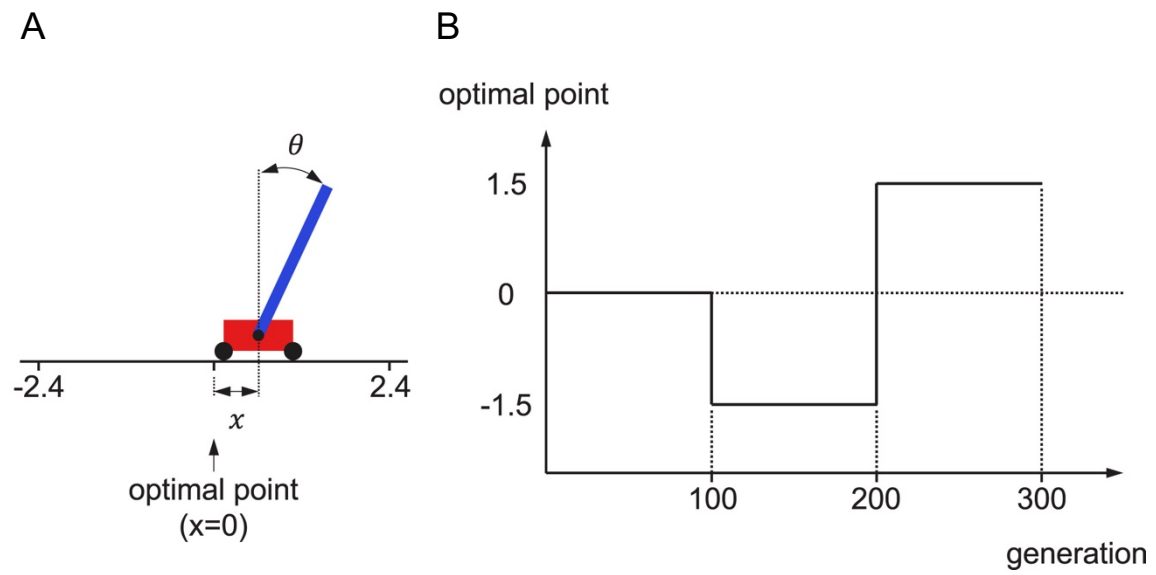


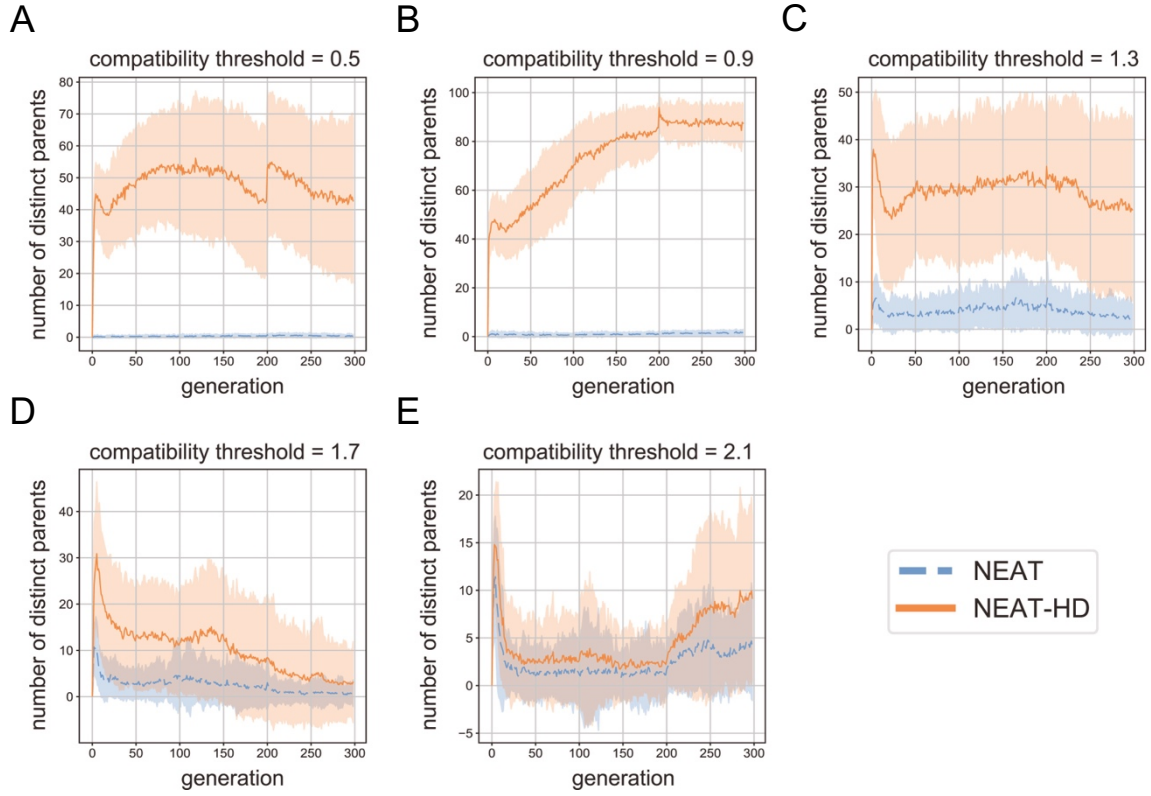
Fig. 4. (A) The concept of biological hybridization. In most major cases, crossover happens within the same species (e.g., only species A or species B, represented by small black arrows). However, in a rare case, crossover happens between different species (e.g., species A and B, represented by a large black arrow) that are genetically distinct and generate hybrid offspring. (B) The process of selecting a pair to generate offspring. Select one parent (parent A), and calculate the mating probability through all possible pairs (represented by black arrows) from the whole population. Then choose the other parent (parent B) by non-equal sampling. Though the most possibilities are that individuals within the same species get parents, some rare cases happen so that parents from different species generate offspring.

Figure 5: Overview of the cart-pole swing-up task.



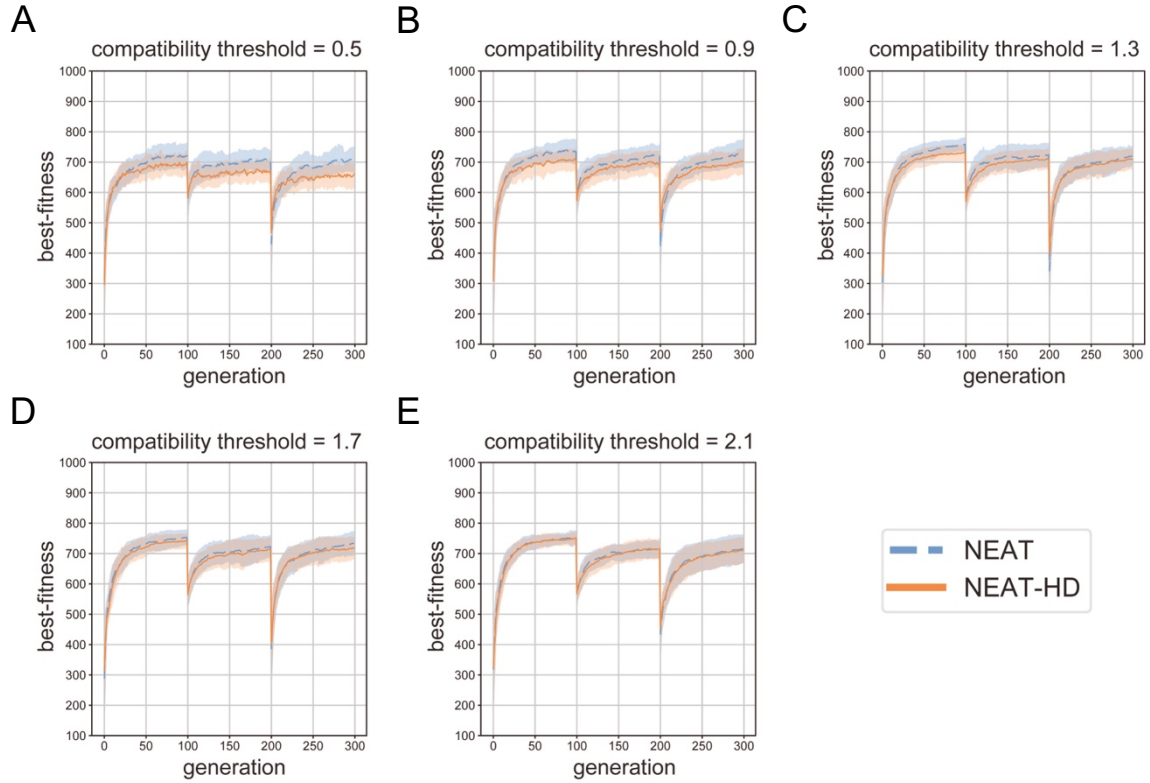
(A) The goal is to balance the pole upright by moving the cart left and right. (B) The trajectory of the optimal point.

Figure 6: The number of hybrid offspring through generation at each compatibility threshold.



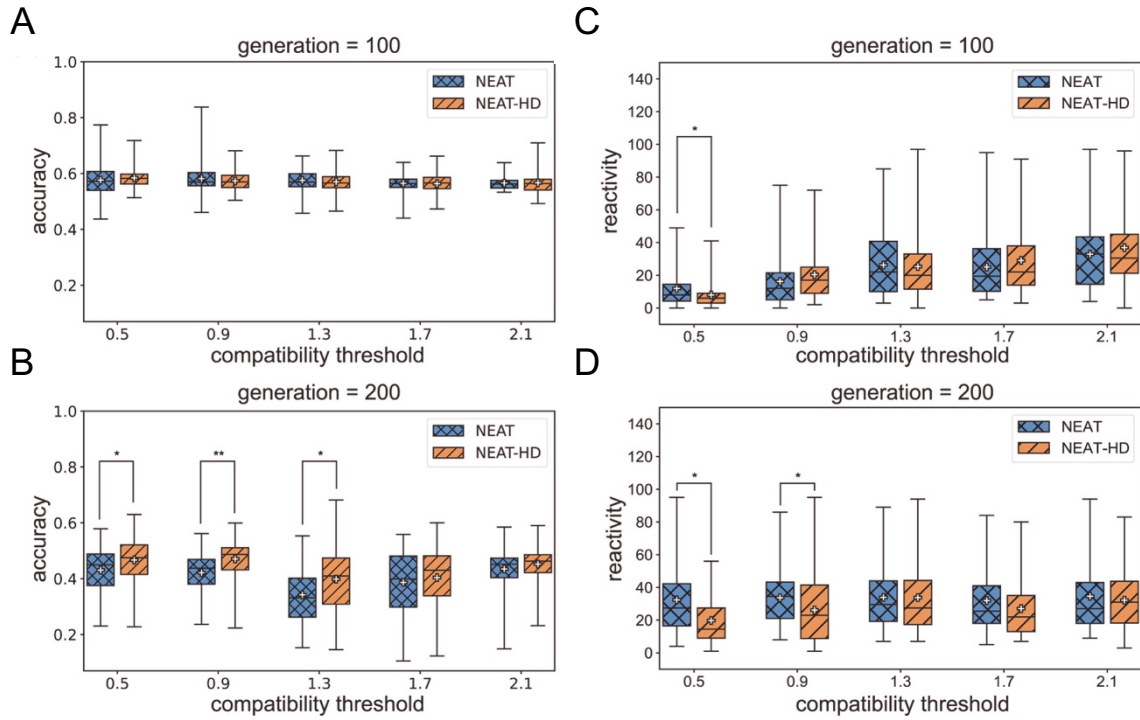
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 7: Best-fitness through generation at each compatibility threshold in the cart-pole task.



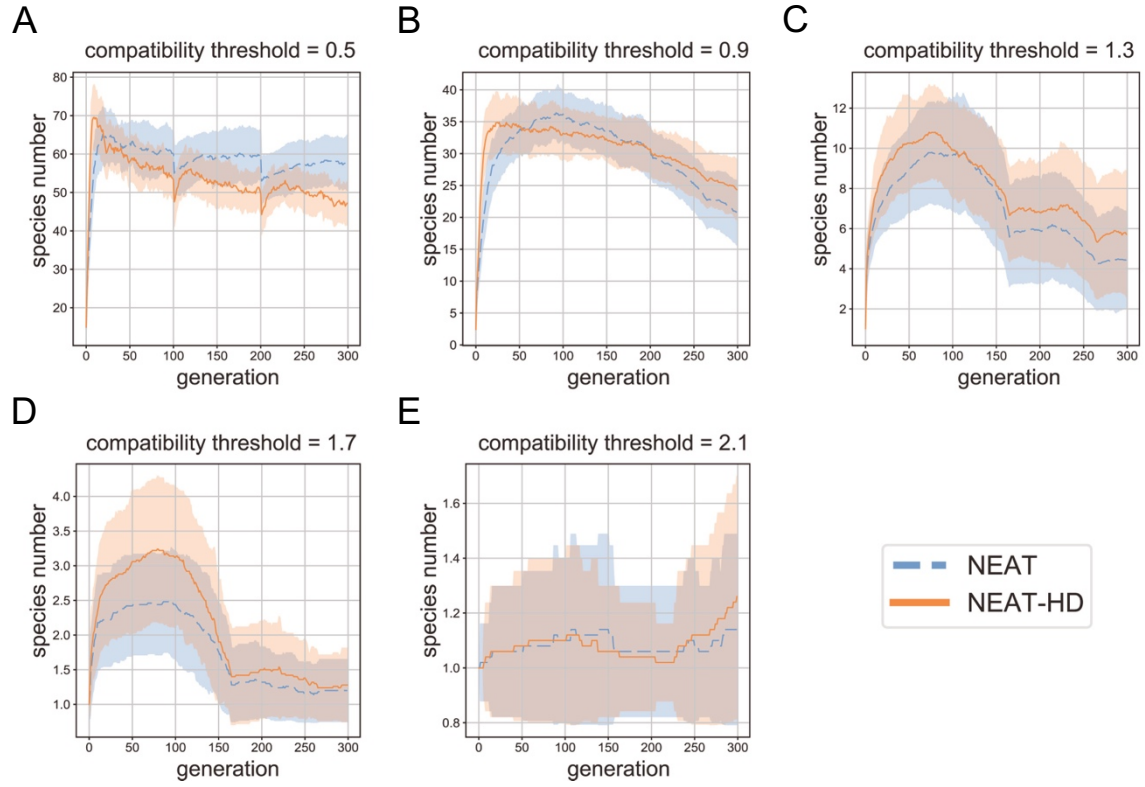
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 8: Accuracy and reactivity.



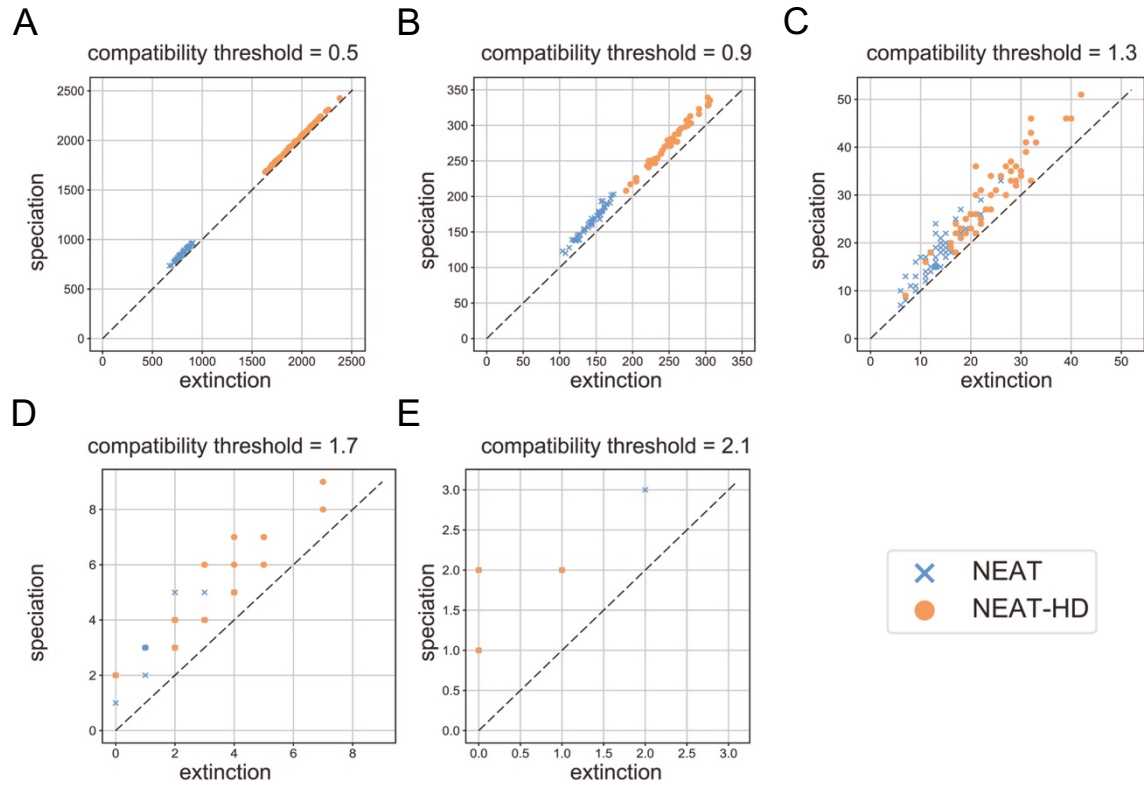
Left side are accuracies at the 100th (A) and 200th (B) generation at each compatibility threshold. Right side are reactivity at the 100th (C) and 200th (D) generation at each compatibility threshold. The white cross in the box represents the mean value.

Figure 9: The species number through generation at each compatibility threshold.



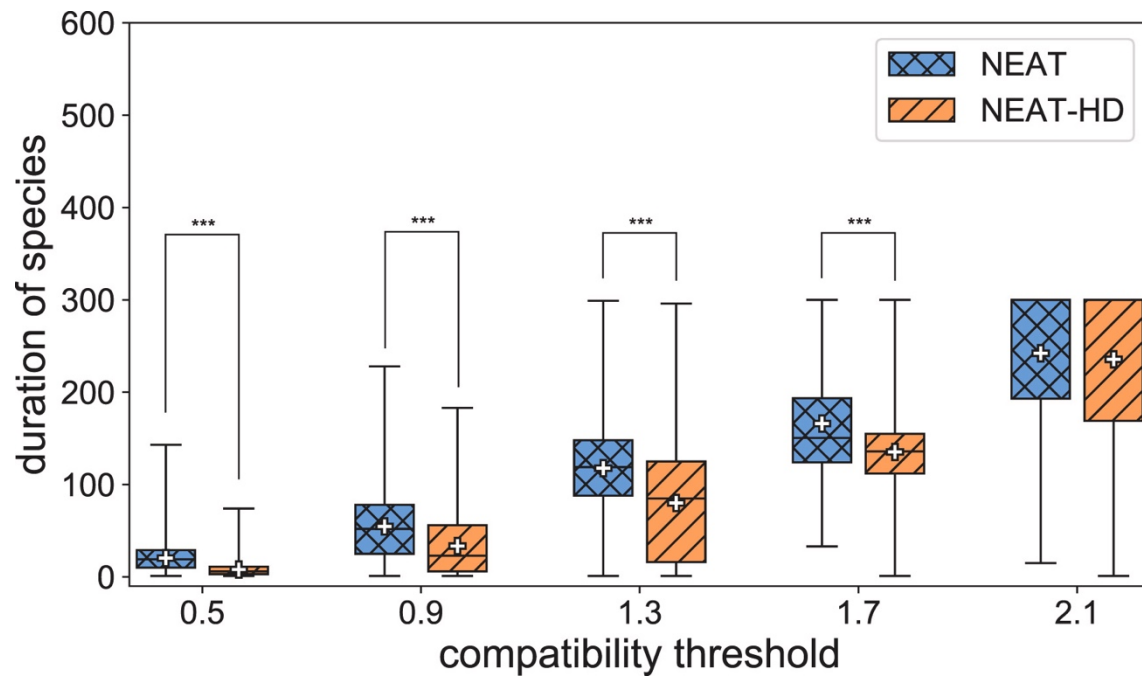
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 10: The relationship between the total number of the species and the number of species extinction counted through the run at each compatibility threshold.



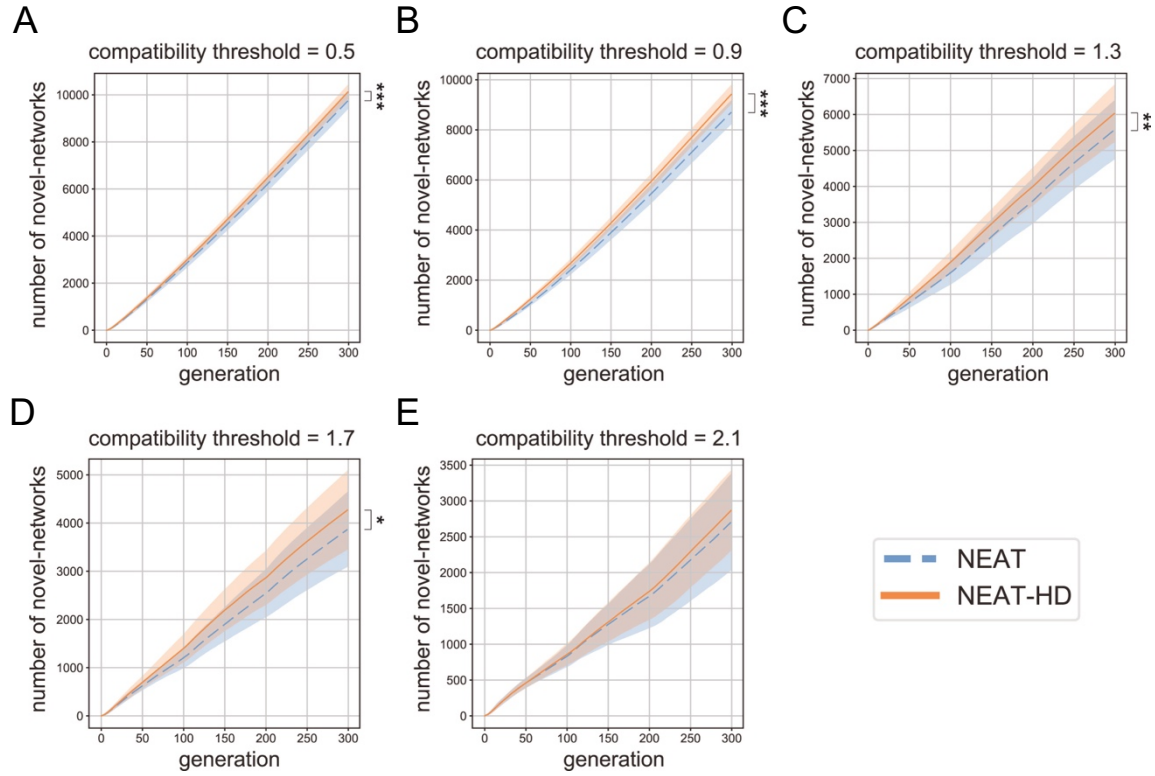
The orange circle represents NEAT-HD, and the blue cross represents NEAT. The baseline drawn by dashed line indicates that all species formed through the generation were entirely extinct until the last generation.

Figure 11: Duration of species at each compatibility threshold.



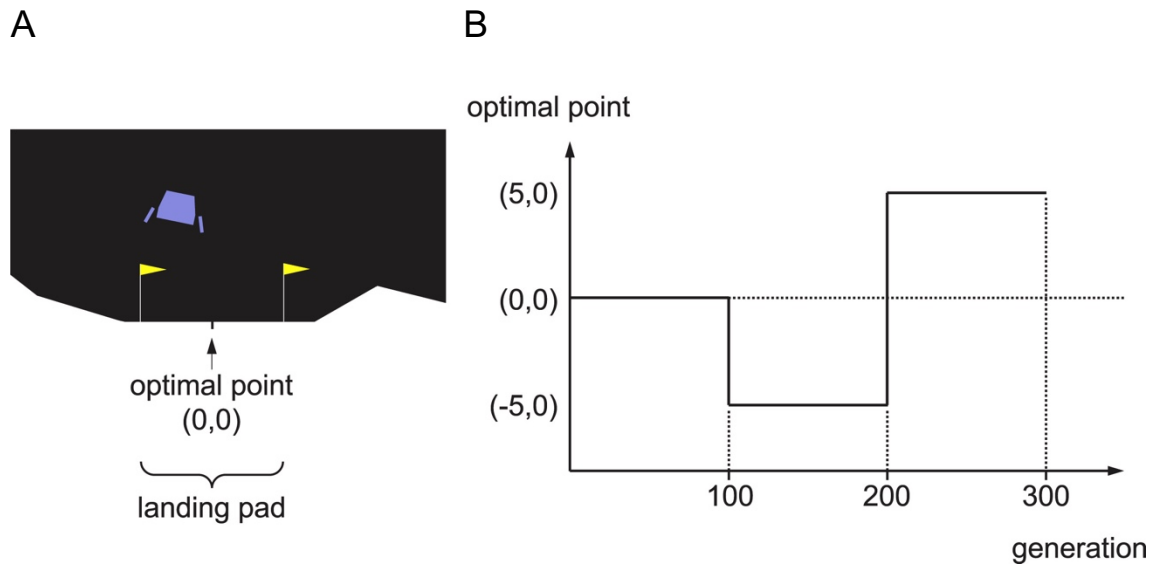
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 12: The number of novel networks through generation at each compatibility threshold.



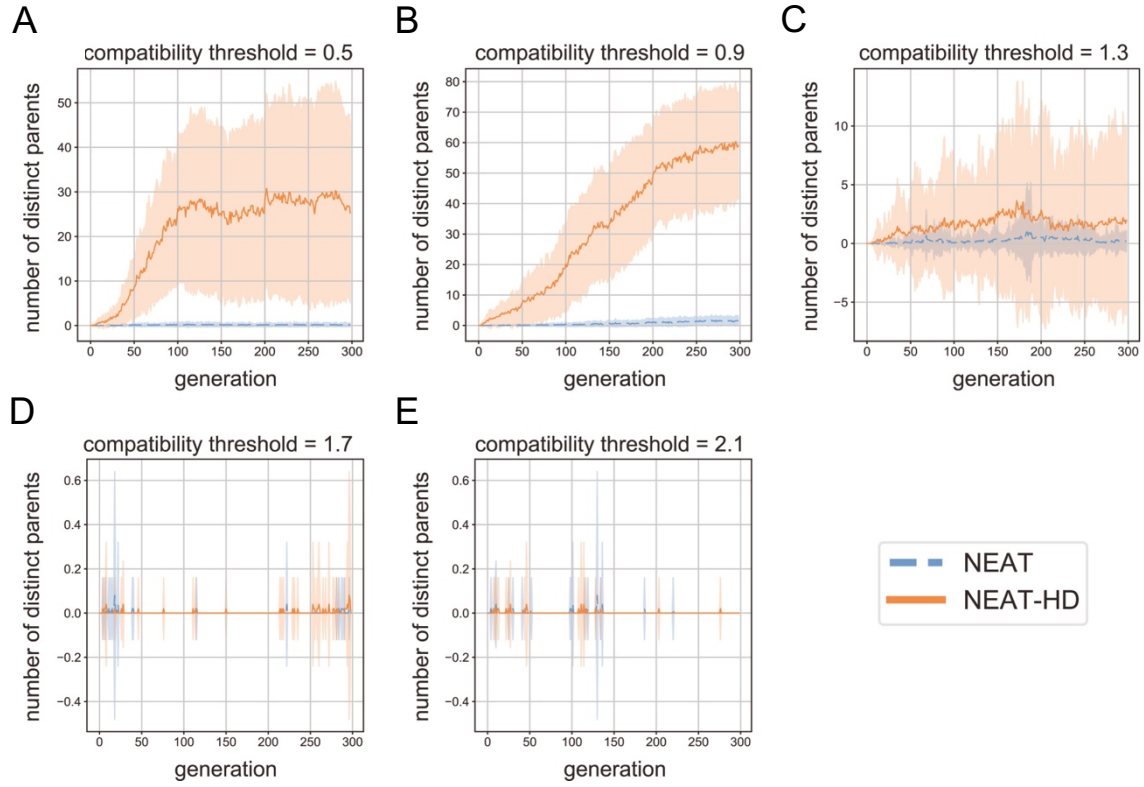
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 13: Overview of the lunar lander task.



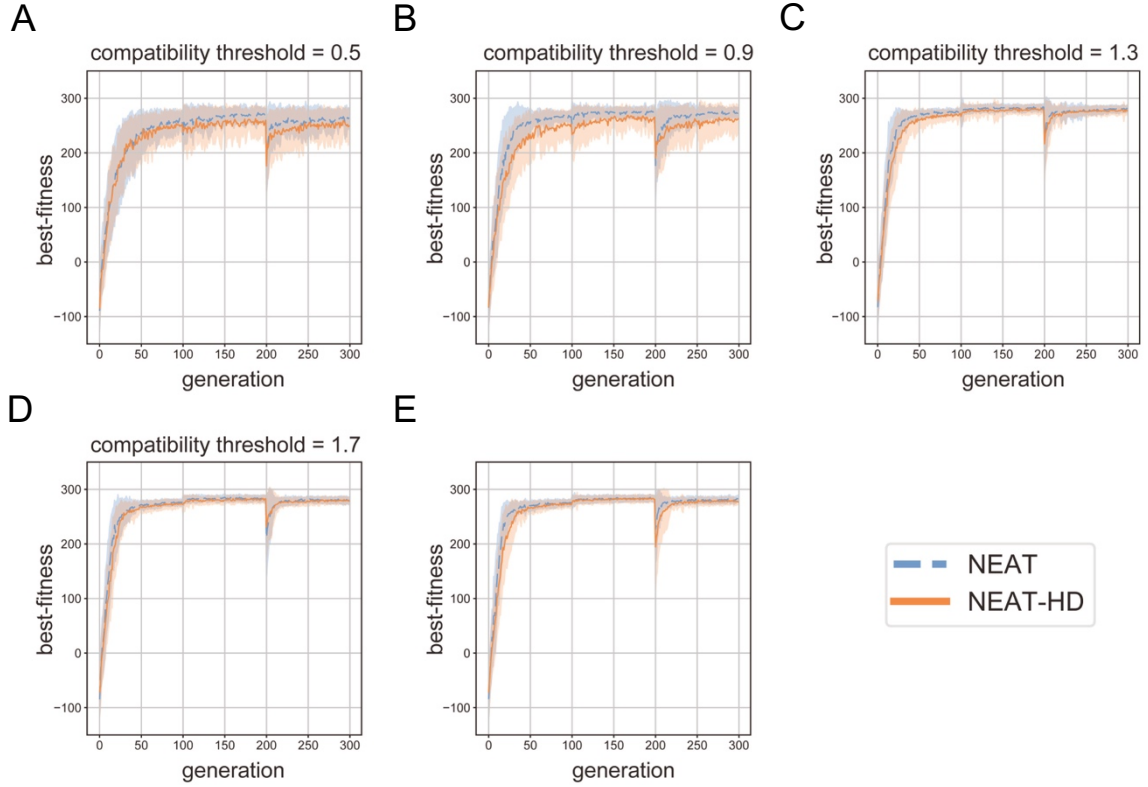
(A) The goal for the lander is to land safely on the ground with less use of engine. The area covered by two flags represents the landing pad. The optimal point is the center of the landing pad. (B) The trajectory of the optimal point.

Figure 14: The number of hybrid offspring through generation at each compatibility threshold.



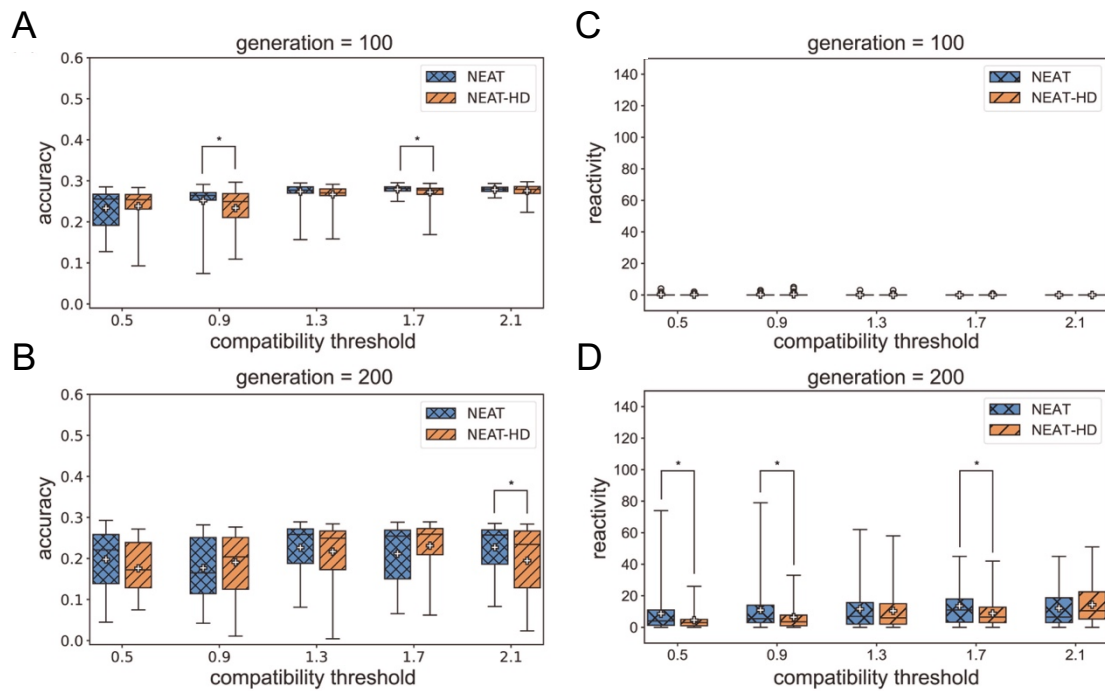
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Figure 15: Best-fitness through generation at each compatibility threshold in the lunar lander task.



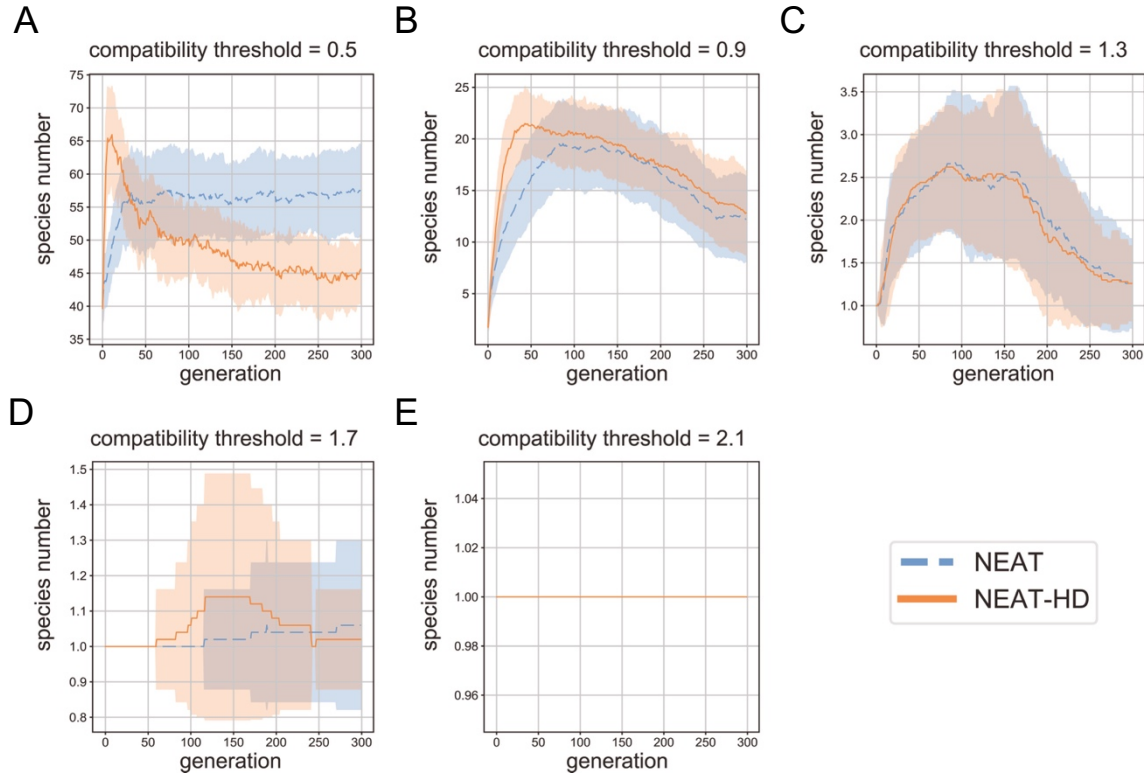
The lines and dashed lines are utilized to separate the algorithm. The shaded regions indicate the standard deviation from the mean.

Figure 16: Accuracy and reactivity.



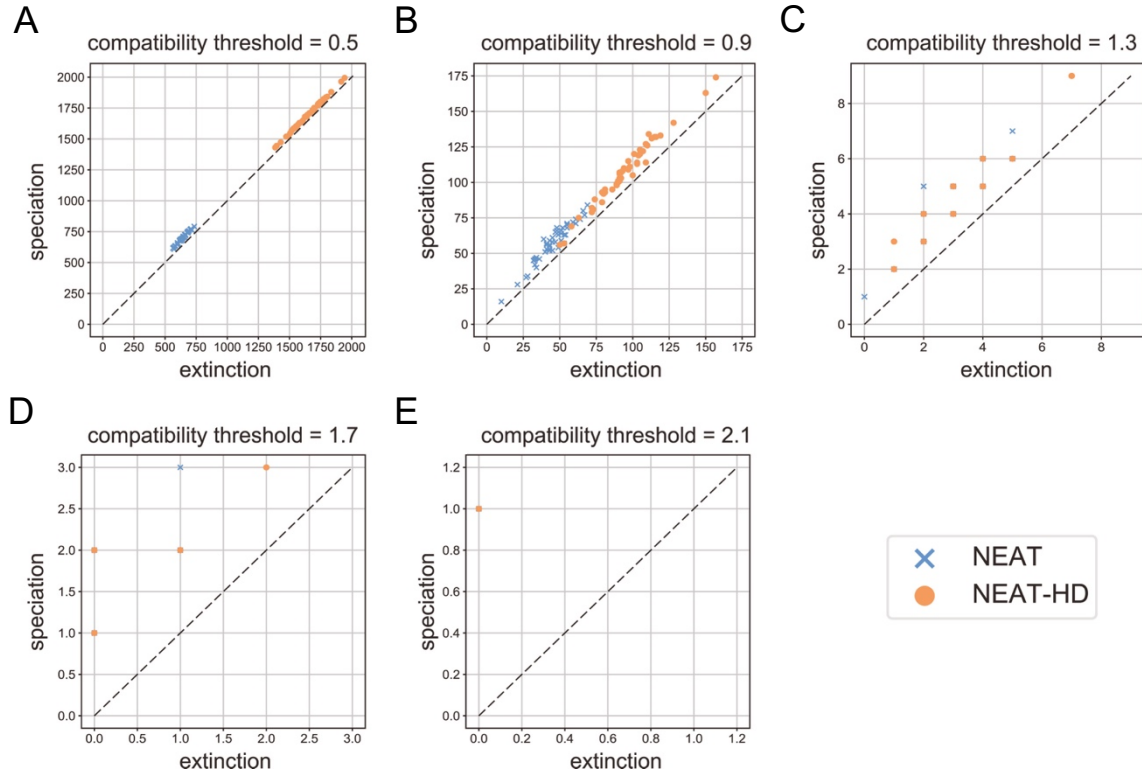
The left side is accuracies at the 100th (A) and 200th (B) generation at each compatibility threshold. The right side shows reactivities at the 100th (C) and 200th (D) generation at each compatibility threshold. The white cross in the box represents the mean value.

Figure 17: The species number through generation at each compatibility threshold.



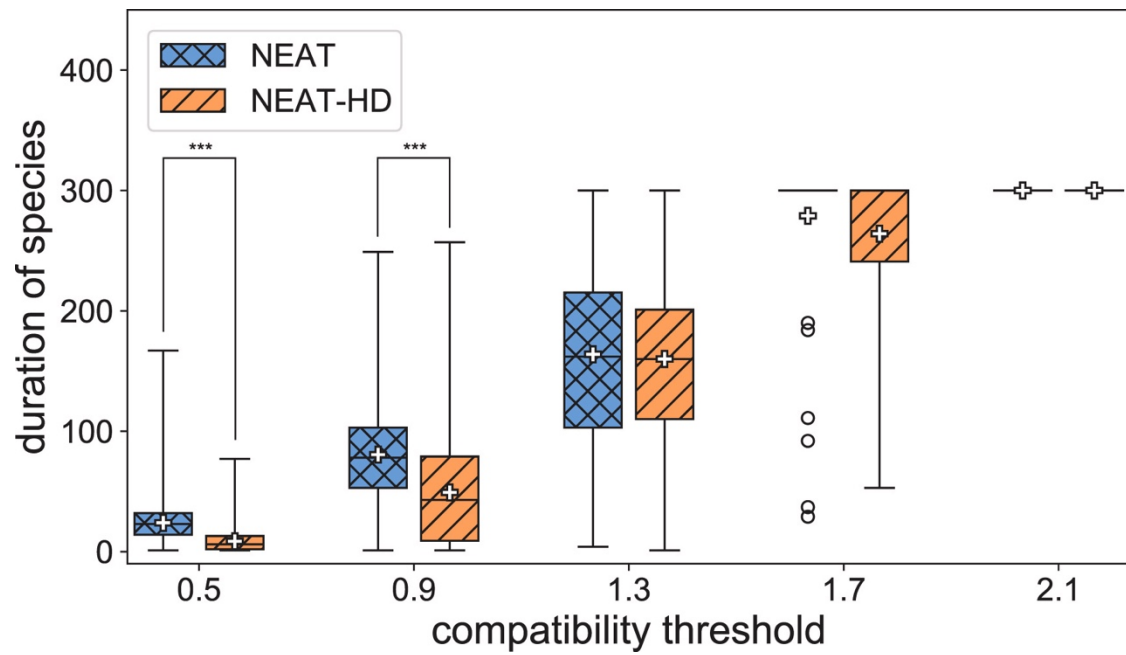
The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean. At (E), the species number was 1 through the generation and all simulations, expressing only as a line.

Figure 18: The relationship between the total number of the species and the number of species extinction counted through the run at each compatibility threshold.



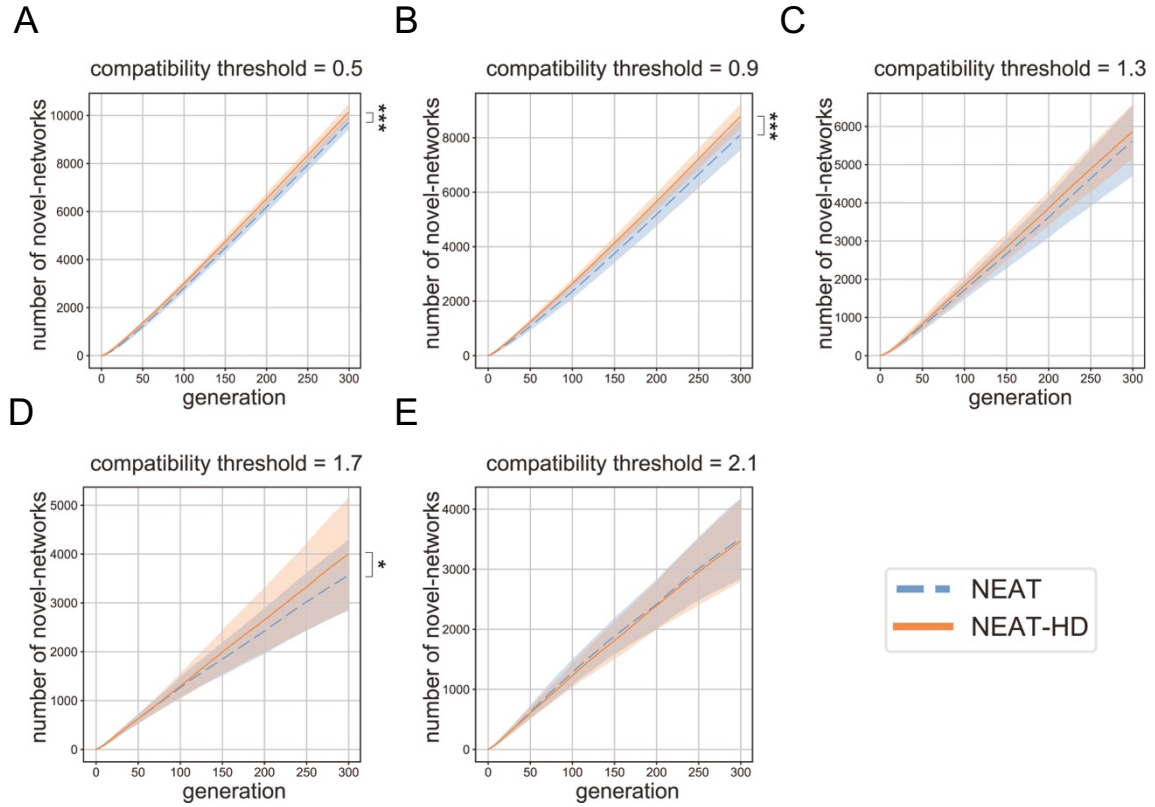
The orange circle represents NEAT-HD, and the blue cross represents NEAT. Dashed lines are the baseline, which means that all species formed through the generation were entirely extinct until the last generation.

Figure 19: Duration of species at each compatibility threshold.



The white cross in the box represents the mean value.

Figure 20: The number of novel networks through generation at each compatibility threshold.



The lines represent NEAT-HD, and dashed lines represent NEAT. The shaded regions indicate the standard deviation from the mean.

Supplementary Tables

Table 1: The parameters used in this study.

parameter	value
Generation	300
Population size	128
Generations to trigger stagnation	64
Number of repetition when evaluating individuals	2
Type of speciation ("neat" or "none")	"neat"
Percentage chance of applying second objective when using Multi Objective Optimization	0.0
Activation function for hidden nodes	5
Probability of adding connection	0.15
Probability of adding node	0.1
Probability of crossover	0.8
Probability to enable disabled connection	0.01
Probability of changing node activation function	0.0
Probability of mutating connection weight	0.8
Probability of enabling each initial connection	1.0
Percent of individuals to remove from parent pool before selection	0.1
Percent of individuals to pass on to next generation unchanged	0.1
How to scale individual rank ("exponential" or "linear")	"exp"
Number of competitors in each tournament	2
Species threshold	1.3
Weighting of non-matching genes in compatibility distance calculation	1
Weighting of weight differences in compatibility distance calculation	0.5

Table 2: The stabilities of both algorithms (NEAT and NEAT-HD) at each compatibility threshold in cart-pole swing-up task.

Compatibility threshold	0.5		0.9		1.3		1.7		2.1	
Type	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD
Generation=100	1.0	1.0	0.94	1.0	0.88	0.94	0.92	0.90	0.86	0.84
Generation=200	0.96	1.0	0.96	0.96	0.80	0.84	0.92	0.98	0.86	0.84

Table 3: The stabilities of both algorithms (NEAT and NEAT-HD) at each compatibility threshold in lunar lander task.

Compatibility threshold	0.5		0.9		1.3		1.7		2.1	
Type	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD	NEAT	NEAT-HD
Generation=100	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Generation=200	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Chapter2

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

This study in this chapter was published in J. Sunagawa, R. Komorizono, H. Park, WS. Hart, RN. Thompson, A. Makino, K. Tomonaga, S. Iwami, R. Yamaguchi, “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. (2023).

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 ε , 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 ε), $D(V_{total})$ is calculated by Eqs (5–6) based on the virus infection model [Eqs (1–4)] with other fixed parameters (see **Table A**). 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 θ 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 ε , 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) (**Fig. 2A and 2D**). 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) (**Fig. 2C and D**). The white dots in **Fig. 2A** and **2C** correspond to the coordinates of the combination of p and ε that maximize R_{TP} 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 ε 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 (**Fig. E and F**). 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 ε , 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 θ , and each contact history is made by the combination of each k and θ . We conducted a total of 500 independent iterations for each parameter set of (k, θ) and determined the evolutionary consequences of the median parameter set of p and ε .

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 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 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, ε , 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 (**Fig. 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**). 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 (**Fig. 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**). 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 ε . The RNA-dependent RNA polymerase encoded by RNA viruses lacks proofreading 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 ε determines the establishment of the persistent infection phenotype.

Note that we can also easily extend our interpretation of ε 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 ε . In other words, $p\varepsilon I(t)$ in our model represents “infectious virus production”, meaning the parameter ε 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., δ in Eq (3)) and the virus infection rate (i.e., β 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]:

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

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

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

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

where the variables $T(t)$, $I(t)$, $V_i(t)$ and $V_{non}(t)$ 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 β , δ , p , and cc 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 ε of the produced virus is infectious but $1 - \varepsilon$ is noninfectious (the value of ε represents the accuracy of replication). To mimic viral cytopathogenesis depending on the viral replication level, we defined $\delta = \delta_{max} \frac{p}{p+p_{50}}$, that is, an increasing Hill function of p . The parameters δ_{max} and p_{50} are the maximum value of δ and the viral production rate satisfying $\delta = \frac{\delta_{max}}{2}$, respectively. We fixed the parameter values (except for p and ε) and initial values to correspond to biologically reasonable values or ranges (see below and **Table A**). The sensitivity analysis on δ is provided in **Fig. G**.

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]:

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

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

where D_{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) (**Fig. H, I, and J**). 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 ρ 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 λ 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 θ are the shape parameter and the scale parameter, respectively. We note that k influences the skewness, and θ influences the variance of the distribution, respectively. The mean contact number through the history is the product of k and θ , 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, Fig. E, and F**). 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:

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

where $D(V_{total})$ 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 ε , respectively. We calculate R_{TP} in various patterns of (p, ε) , and thus the fitness landscape of R_{TP} is constructed as a function of p and ε .

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 ε . The genetic algorithm heuristically searches for the optimal R_{TP} with a group of solutions made up of combinations of p and ε (see **Alg.1** and

Table B 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 ε 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 ε , 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 ε with the highest R_{TP} . These iterations can be considered as the virus "evolving" its properties of p and ε 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** 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 (**Fig. EC and FC**), which corresponds to the characteristics of the acute infection phenotype.

Algorithm 1: Genetic Algorithm

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

$t \leftarrow 0$;

Initialize ($P(t = 0)$) with random p and ε ;

Evaluate ($P(t = 0)$);

while not termination do

$P(t)_e \leftarrow \text{Elitism}(P(t))$;

$P(t)_p \leftarrow \text{selectSolutions}(P(t))$;

$P(t)_c \leftarrow \text{crossover}(P(t)_e + P(t)_p)$;

Mutate ($P(t)_c$);

Evaluate ($P(t)_c$);

$P(t+1) \leftarrow P(t)_c$;

$t \leftarrow t + 1$;

end while

References

1. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S. Rapid evolution of RNA genomes. *Science*. 1982;215: 1577–1585.
2. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. *J Virol*. 2010;84: 9733–9748.
3. Llaureing AS, Frydman J, Andino R. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol*. 2013;11: 327–336.
4. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature*. 2006;439: 344–348.
5. 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. *Nat Commun*. 2016;7: 13787.
6. 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*. 2021;118: e2024815118.
7. 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: 991–999.
8. Woolhouse ME, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis*. 2005;11: 1842.
9. McCarthy MK, Morrison TE. Persistent RNA virus infections: do PAMPS drive chronic disease? *Curr Opin Virol*. 2017;23: 8–15.
10. Geoghegan JL, Senior AM, Di Giallonardo F, Holmes EC. Virological factors that increase the transmissibility of emerging human viruses. *Proc Natl Acad Sci*. 2016;113: 4170–4175.
11. Randall RE, Griffin DE. Within host RNA virus persistence: mechanisms and consequences. *Curr Opin Virol*. 2017;23: 35–42.

12. 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. *Commun Biol.* 2021;4: 508.
13. Arias A, Thorne L, Ghurburrun E, Bailey D, Goodfellow I. Norovirus polymerase fidelity contributes to viral transmission in vivo. *MSphere.* 2016;1: 10–1128.
14. Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, Gong P, et al. Coxsackievirus B3 mutator strains are attenuated in vivo. *Proc Natl Acad Sci.* 2012;109: E2294–E2303.
15. Trimpert J, Groenke N, Kunec D, Eschke K, He S, McMahon DP, et al. A proofreading-impaired herpesvirus generates populations with quasispecies-like structure. *Nat Microbiol.* 2019;4: 2175–2183.
16. 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. *Nat Ecol Evol.* 2020;4: 1156–1159.
17. Frickel J, Feulner PG, Karakoc E, Becks L. Population size changes and selection drive patterns of parallel evolution in a host–virus system. *Nat Commun.* 2018;9: 1706.
18. Rouzine IM, Weinberger AD, Weinberger LS. An evolutionary role for HIV latency in enhancing viral transmission. *Cell.* 2015;160: 1002–1012.
19. 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: e63537.
20. Shrestha S, Bjørnstad ON, King AA. Evolution of acuteness in pathogen metapopulations: conflicts between “classical” and invasion-persistence trade-offs. *Theor Ecol.* 2014;7: 299–311.
21. Sasaki A, Lion S, Boots M. Antigenic escape selects for the evolution of higher pathogen transmission and virulence. *Nat Ecol Evol.* 2022;6: 51–62.

22. Kinnunen PM, Palva A, Vaheri A, Vapalahti O. Epidemiology and host spectrum of Borna disease virus infections. *J Gen Virol.* 2013;94: 247–262.
23. Staeheli P, Sauder C, Hausmann J, Ehrensperger F, Schwemmler M. Epidemiology of Borna disease virus. *J Gen Virol.* 2000;81: 2123–2135.
24. Pauli G, Ludwig H. Increase of virus yields and releases of Borna disease virus from persistently infected cells. *Virus Res.* 1985;2: 29–33.
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. *Sci Rep.* 2020;10: 7257.
26. Ke TingYung KT, Liao WeiYu LW, Wu HungYi WH. A leaderless genome identified during persistent bovine coronavirus infection is associated with attenuation of gene expression. 2013.
27. Peterhans E, Bachofen C, Stalder H, Schweizer M. Cytopathic bovine viral diarrhoea viruses (BVDV): emerging pestiviruses doomed to extinction. *Vet Res.* 2010;41: 44.
28. Hamers C, Lecomte C, Kulcsar G, Lambot M, Pastoret P-P. Persistently infected cattle stabilise bovine viral diarrhoea virus leading to herd specific strains. *Vet Microbiol.* 1998;61: 177–182.
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: 205–213.
30. Vilcek S, Alenius S, Paton D, Mittelholzer C, Belak S. Genetic clustering of bovine viral diarrhoea viruses in cattle farms: genetic identification and analysis of viruses directly from cattle sera. *Vet J.* 1999;158: 33–38.
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:

93–102.

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: veaa105.
33. Gulbudak H, Cannataro VL, Tuncer N, Martcheva M. Vector-borne pathogen and host evolution in a structured immuno-epidemiological system. *Bull Math Biol.* 2017;79: 325–355.
34. King AA, Shrestha S, Harvill ET, Bjørnstad ON. Evolution of acute infections and the invasion-persistence trade-off. *Am Nat.* 2009;173: 446–455.
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 R Soc B Biol Sci.* 2013;280: 20122047.
36. Gunning CE, Wearing HJ. Probabilistic measures of persistence and extinction in measles (meta) populations. *Ecol Lett.* 2013;16: 985–994.
37. Brooke CB. Biological activities of 'noninfectious' influenza A virus particles. *Future Virol.* 2014;9: 41–51.
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: 10651–10660.
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 Res.* 1994;34: 281–289.
40. Schneider PA, Briese T, Zimmermann W, Ludwig H, Lipkin WI. Sequence conservation in field and experimental isolates of Borna disease virus. *J Virol.* 1994;68: 63–68.
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: e3000562.

42. Andino R, Domingo E. Viral quasispecies. *Virology.* 2015;479: 46–51.

43. Holland JJ, De La Torre JC, Steinhauer DA. RNA virus populations as quasispecies. *Curr Top Microbiol Immunol.* 1992;176: 1–20.
doi:10.1007/978-3-642-77011-1_1

44. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev.* 2012;76: 159–216.

45. Dolan PT, Whitfield ZJ, Andino R. Mechanisms and concepts in RNA virus population dynamics and evolution. *Annu Rev Virol.* 2018;5: 69–92.

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: 1603–1619.

47. Lenormand T, Roze D, Rousset F. Stochasticity in evolution. *Trends Ecol Evol.* 2009;24: 157–165. doi:10.1016/j.tree.2008.09.014

48. Briesse 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: 11486–11489.
doi:10.1073/pnas.89.23.11486

49. Cubitt B, Oldstone C, de la Torre JC. Sequence and genome organization of Borna disease virus. *J Virol.* 1994;68: 1382–1396.
doi:10.1128/JVI.68.3.1382-1396.1994

50. Villarreal LP, Defilippis VR, Gottlieb KA. Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology.* 2000;272: 1–6. doi:10.1006/viro.2000.0381

51. McNeill WH, McNeill W. *Plagues and peoples.* 1976.

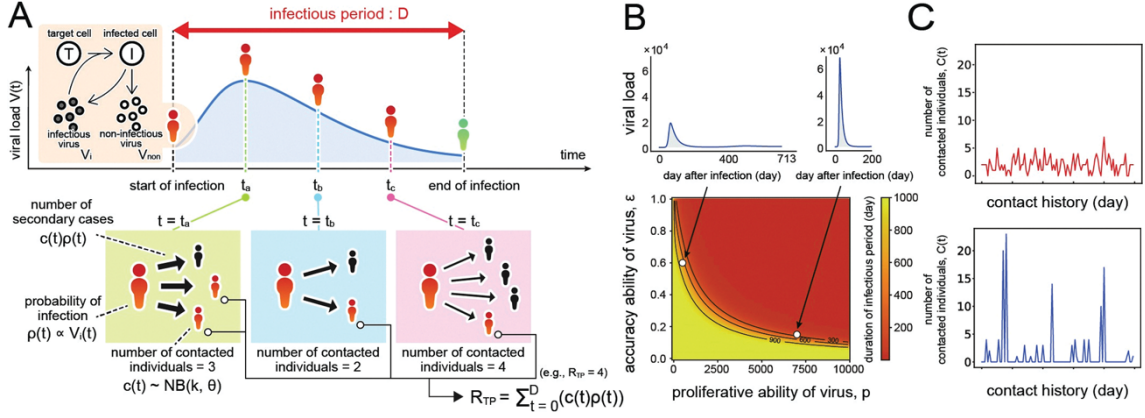
52. Michael B. A. Oldstone. *Viruses, Plagues, and History.* 1998.

53. Jared M. Diamond. *Guns, Germs, and Steel: The Fates of Human Societies*. 1999.
54. Koelle K, Martin MA, Antia R, Lopman B, Dean NE. The changing epidemiology of SARS-CoV-2. *Science*. 2022;375: 1116–1121. doi:10.1126/science.abm4915
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: 603–610. doi:10.1016/S1473-3099(22)00001-9
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. doi:10.1093/molbev/msac013
57. Nowak MA, Sigmund K. Evolutionary dynamics of biological games. *Science*. 2004;303: 793–799. doi:10.1126/science.1093411
58. Renaud F, de Meeüs T. A simple model of host-parasite evolutionary relationships. *Parasitism: compromise or conflict? J Theor Biol*. 1991;152: 319–327. doi:10.1016/s0022-5193(05)80197-3
59. Thompson RN, Gilligan CA, Cunniffe NJ. Detecting Presymptomatic Infection Is Necessary to Forecast Major Epidemics in the Earliest Stages of Infectious Disease Outbreaks. *PLoS Comput Biol*. 2016;12: e1004836. doi:10.1371/journal.pcbi.1004836
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. *J R Soc Interface*. 2020;17: 20200230. doi:10.1098/rsif.2020.0230
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: 18. doi:10.1186/1742-4690-9-18

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. *Sci Rep*. 2015;5: 12256. doi:10.1038/srep12256
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: 17441–17446. doi:10.1073/pnas.0708559104
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: 11–17. doi:10.7326/M14-2255
65. Lindén A, Mäntyniemi S. Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology*. 2011;92: 1414–1421. doi:10.1890/10-1831.1
66. Barabasi AL, Albert R. Emergence of scaling in random networks. *Science*. 1999;286: 509–512. doi:10.1126/science.286.5439.509
67. Watts DJ, Strogatz SH. Collective dynamics of “small-world” networks. *Nature*. 1998;393: 440–442. doi:10.1038/30918
68. John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence (Complex Adaptive Systems)*. 1992.

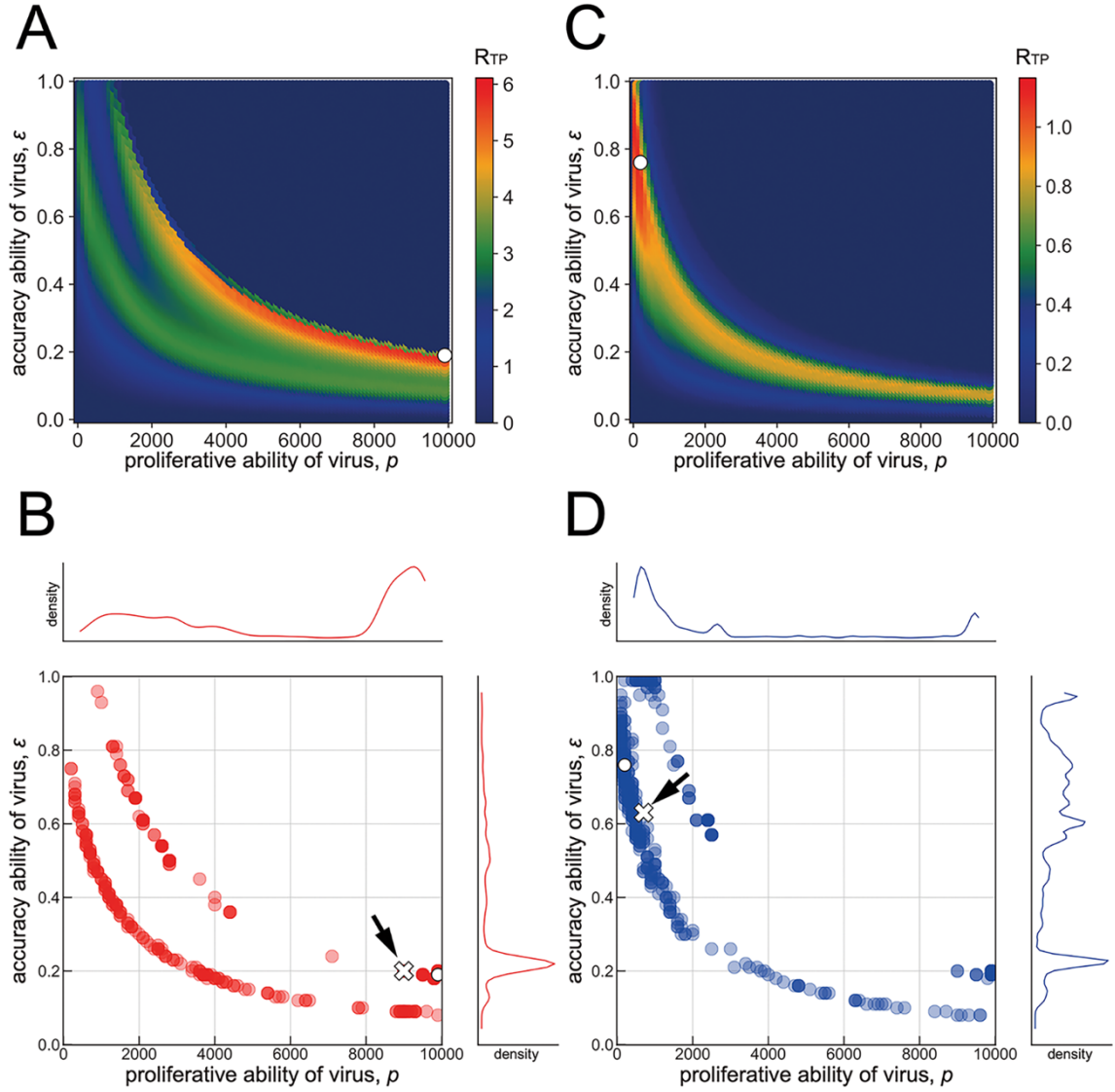
Figures

Figure 1: Schematic illustration of multi-level disease transmission.



A schematic of our model is depicted in (A). The probability of infection $\rho(t)$ 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 ε) 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.

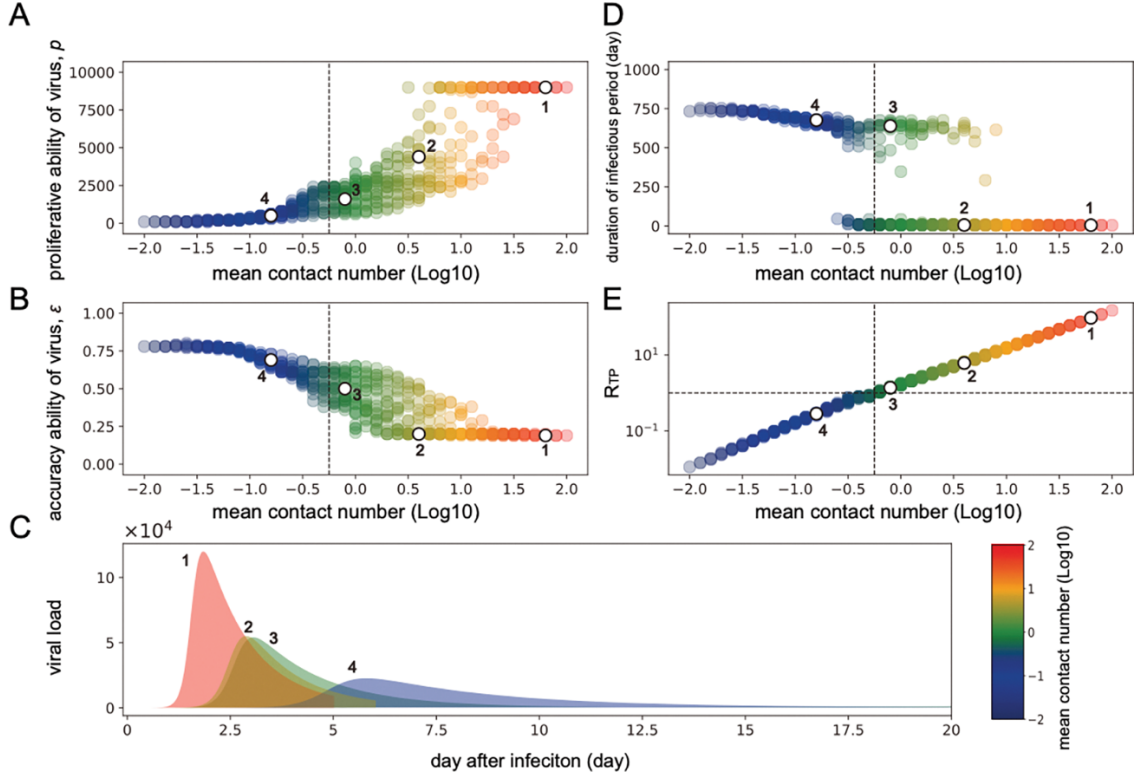
Figure 2: Virus evolution in a given environment.



The optimal parameter combinations (p and ε) 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 ε , 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 ε 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 ε) 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.

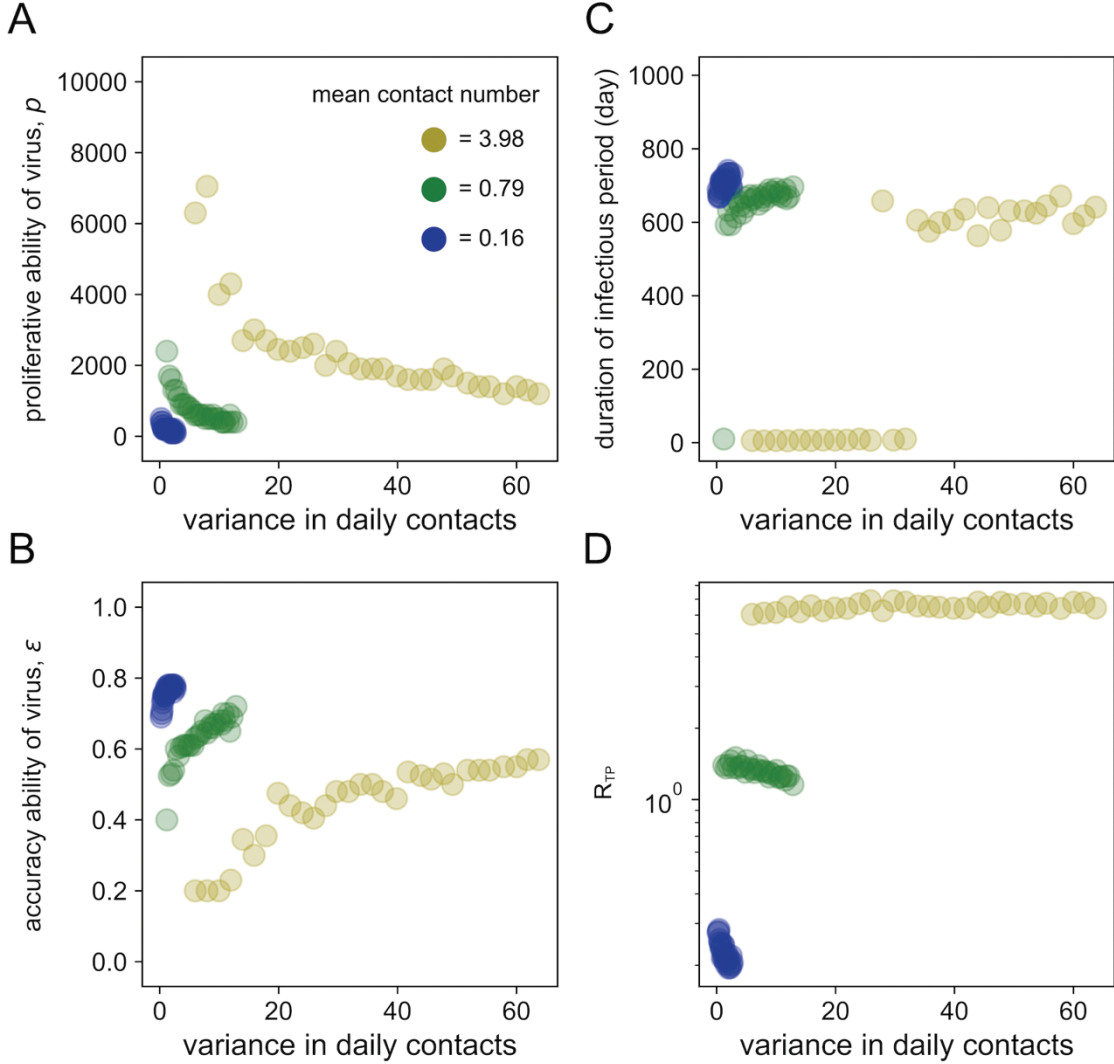
Figure 3: Virus evolution in various scenarios of contact history.



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 θ are conducted to determine the optimized parameter set of p and ε characterizing final virus infection dynamics (i.e., different viral infection phenotypes). The median of the optimized parameter set of p and ε 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 θ . The combination of p and ε 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).

Figure 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 θ 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 θ , and showed the mean values. For different variance in daily contact numbers shown in the horizontal axis, the optimal parameter combination of p and ε 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.

Supplementary Tables

Table A: Parameters used for virus infection dynamics in the main text

parameter	value
Number of uninfected target cells at initial time, $T(t = 0)$	1000
Number of infected target cells at initial time, $I(t = 0)$	0
Amount of infectious virus at initial time, $V_i(t = 0)$	1
Amount of noninfectious virus at initial time, $V_{\text{non}}(t = 0)$	0
Initial number of uninfected cells, λ	10
Rate of virus infection, β	0.0001
Death rate of noninfectious cells, μ	0.01
Death rate of infectious cells, δ	$\delta_{\text{max}} p / (p + p_{50})$
Maximum value of δ , δ_{max}	1
Viral production rate satisfying $\delta_{\text{max}}/2$, p_{50}	1000
Clearance rate of the virus, c	10
Proliferative ability of virus, p	[0, 100, 200, ..., 9900] (100 points in total)
Accuracy of virus replication, ε	[0.0, 0.01, 0.02, ..., 0.99] (100 points in total)
Maximum duration of the infectious period, D_{max}	1000
Viral load at which the duration is half of its maximum, V_{50}	1.0×10^6
Steepness at which duration decreases with increasing viral load, D_k	7.5
Basilar probability of infection immediately, b	9.8×10^{-8}
r-fold increase in infectiousness for each 10-fold increase in infectious viral load, r	25

Table B: Parameters used in GA Algorithm

parameter	value
The size of the population	100
The number of iterations	100
The number of elitism	2
The probability of selection	1.0
The probability of crossover	0.7
The probability of mutation	0.1
The pre-defined range of p in mutation	200
The pre-defined range of ϵ in mutation	0.2

Table C: Parameters used for the network models

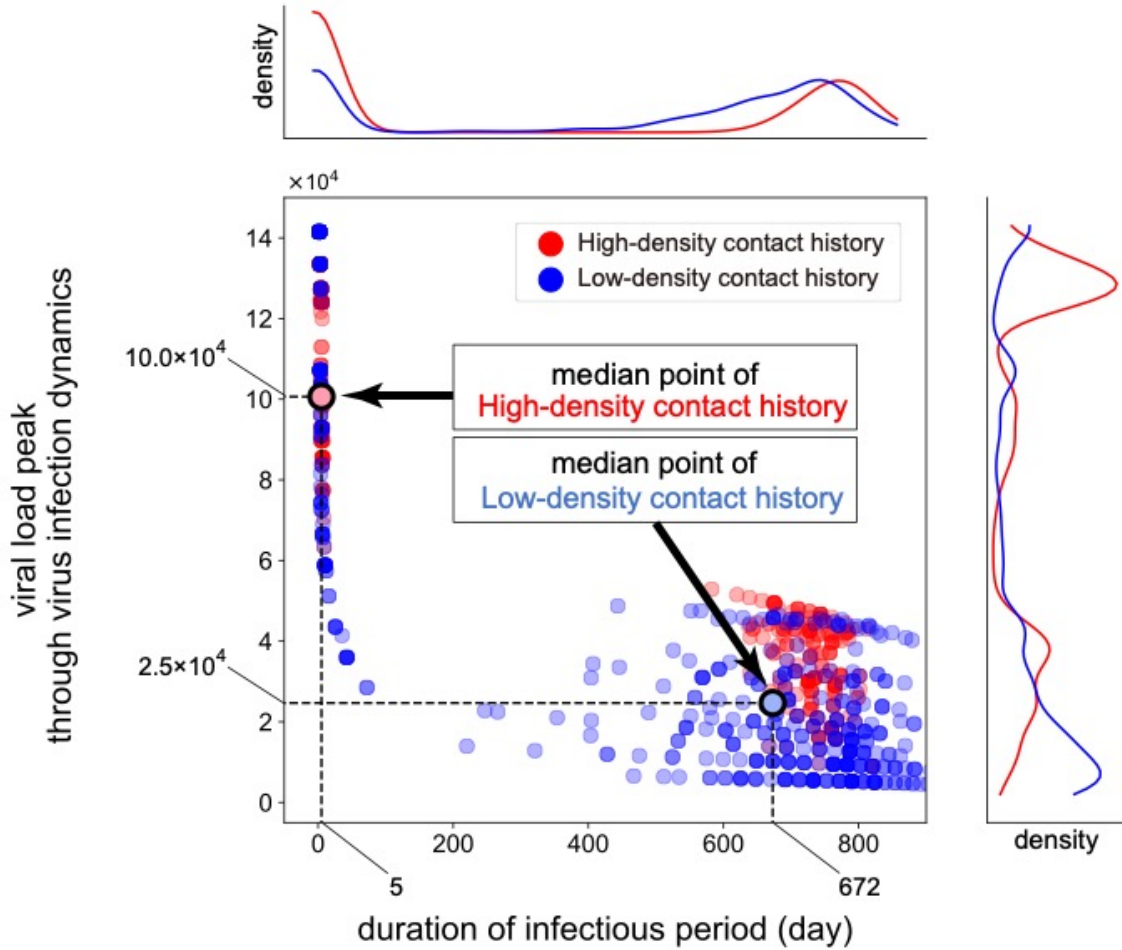
model	parameter	value
BA	number of nodes, n	100
	number of edges to attach from a new node to existing nodes, m	1
WS	number of nodes, n	100
	nearest nodes which are connected, k	2
	the probability of rewiring each edge, p	0.5

BA: Barabási-Albert scale-free network model

WS: Watts-Strogatz small world network model

Supplementary Figures

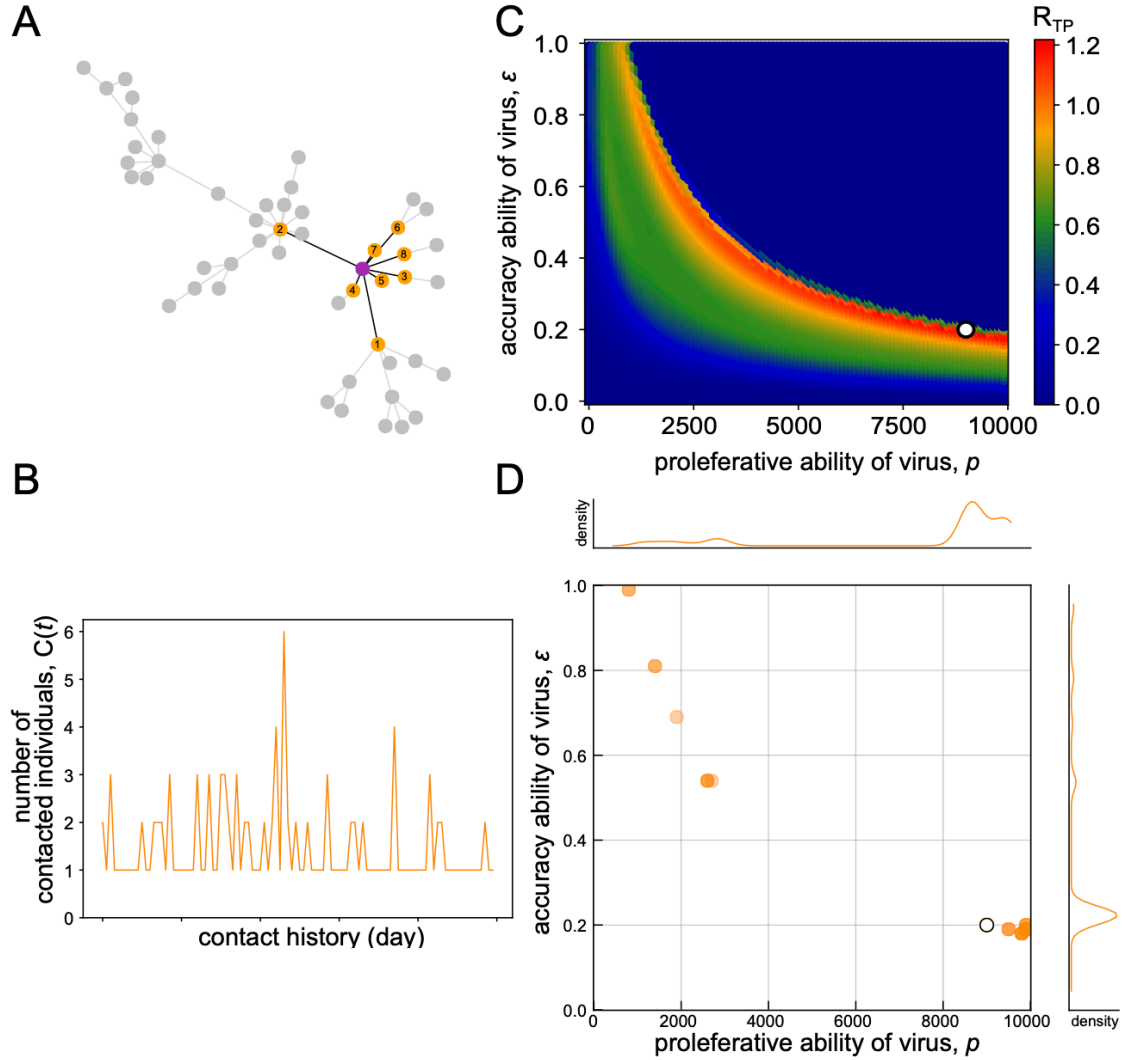
Figure D: Relation between duration of infectious period and peak viral load.



In each independent iteration, viruses evolve for the duration of the infectious period and the peak viral load depending on the contact history: high-density contact history (red) and low-density contact history (blue), respectively. The black arrows correspond to the median values representing the evolutionary endpoints. The parameter sets of shape and scale used are as follows: $(k, \theta) = (1000, 0.002)$ for high-density contact history, and $(k, \theta) = (0.12, 10)$ for low-density contact history, respectively. For both contact histories, the distributions of duration of

infectious period are illustrated as kernel plots above the main Figure. Also shown to the right of the main Figure are the distributions of viral load peak through virus infection dynamics. For the high-density contact history, virus evolution results in a short duration of infectious period (5 days) with a high viral load (1.0×10^5). In contrast, for low-density contact history, virus evolution results in a long duration of infectious period (672 days) with a low viral load (2.5×10^4). 500 simulations were run for each contact history parameter set.

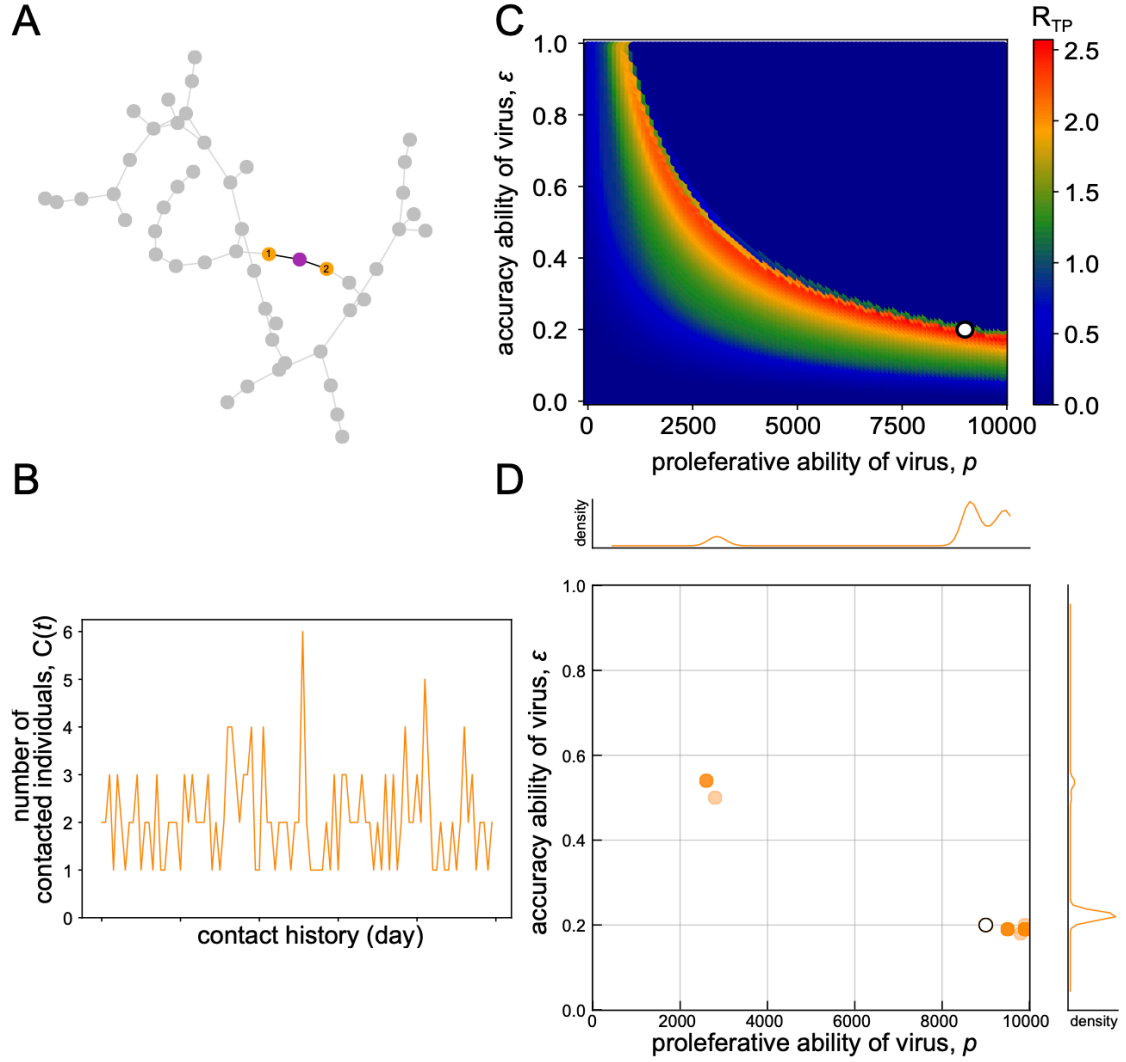
Figure E: Evolution of virus on a contact history generated from the Barabási-Albert scale-free network model.



An example graph of the BA model is illustrated in (A). Contact numbers are calculated by counting the number of nodes that are connected to a single node chosen at random (purple). In this example, the contact number is 8 (orange). Iterating this process generates a contact history. An example of a given contact history is plotted in (B) (only 100 days is shown). Optimal parameter combinations (p and ε) to increase R_{TP} are calculated for a given contact history generated from the BA model in (C). Color represents the degree of R_{TP} : blue corresponds to lower values of R_{TP} ,

and red corresponds to higher values of R_{TP} . The white dot represents the optimal point that increases R_{TP} in a single simulation (see the corresponding white dot in (D)). Using the same parameter set, 50 optima based on independent contact histories are calculated in (D). The kernel plots illustrated at the top and right of the Figure are the distribution of p and ε , respectively.

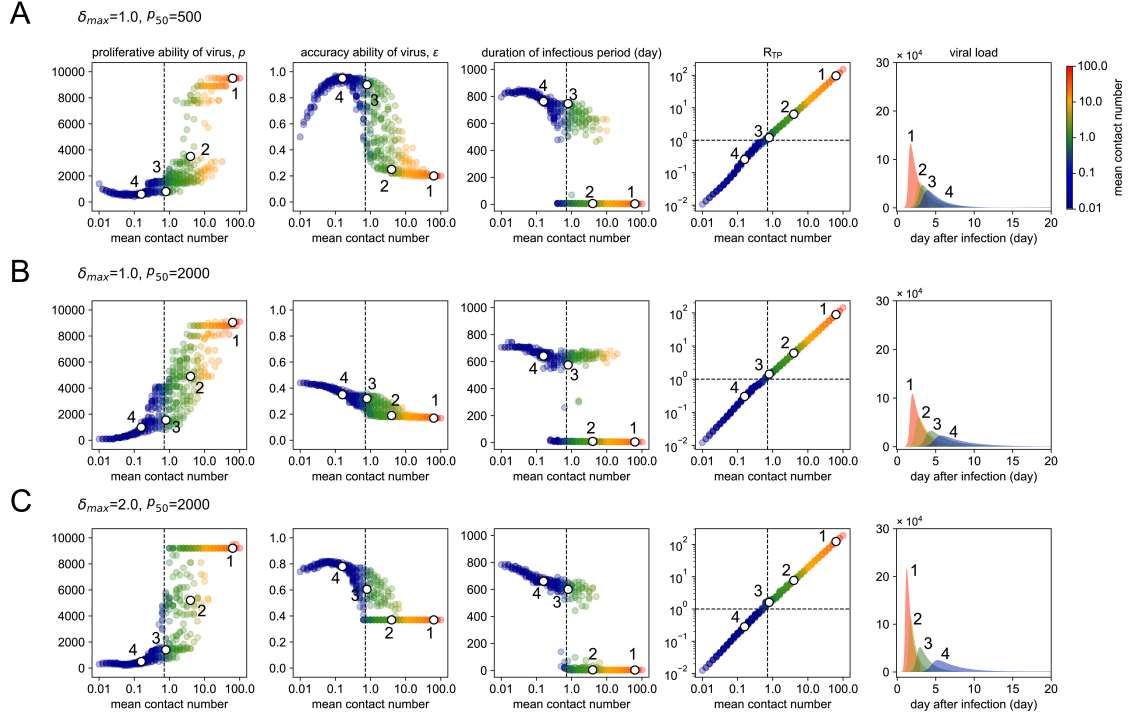
Figure F: Evolution of virus on a contact history generated from the Watts-Strogatz small world network model.



An example graph of the WS model is illustrated in (A). Contact numbers are calculated by counting the number of nodes that are connected to a single node chosen at random (purple). In this example, the contact number is 2 (orange). Iterating this process generates a contact history. An example of a given contact history is plotted in (B) (only 100 days is shown). Optimal parameter combinations (p and ε) to increase R_{TP} are calculated for a given contact history generated from the WS model in (C). Color represents the degree of R_{TP} : blue corresponds to lower values of R_{TP} ,

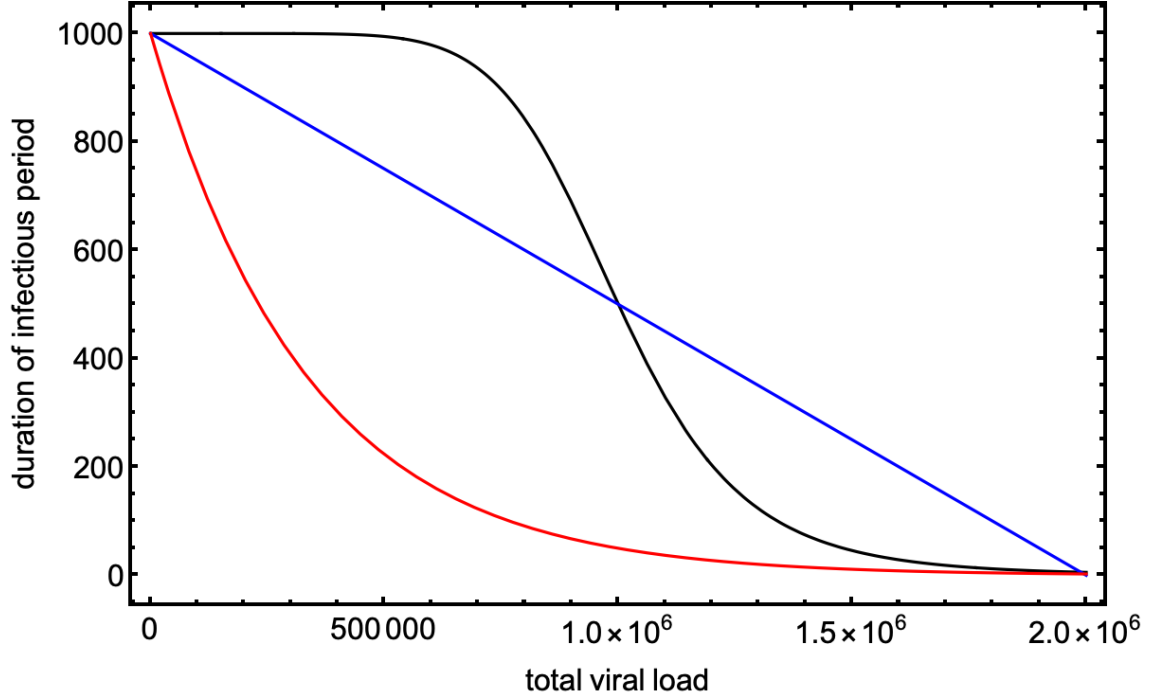
and red corresponds to higher values of R_{TP} . The white dot represents the optimal point that increases R_{TP} in a single simulation (see the corresponding white dot in (D)). Using the same parameter set, 50 optima based on independent contact histories are calculated in (D). The kernel plot illustrated at the top and right of the Figure are the distribution of p and ε , respectively.

Figure G: Virus evolution in various scenarios of contact history.



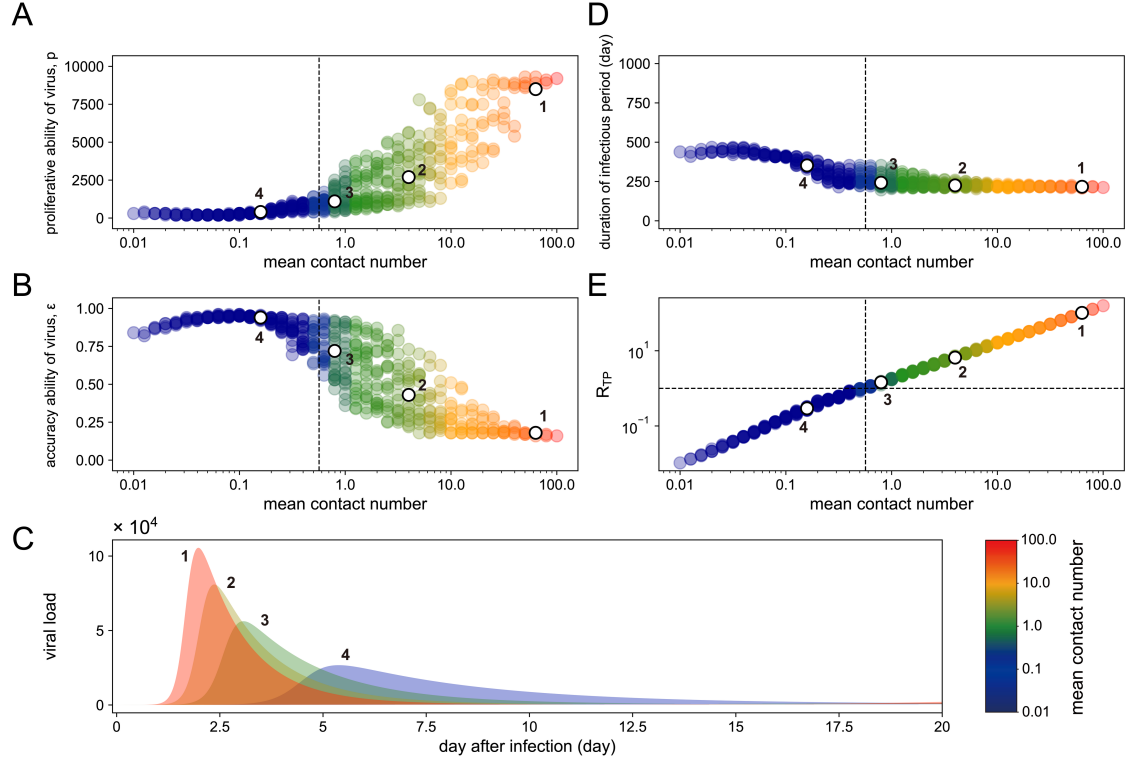
Each row represents a single parameter set of $(\delta_{max}$ and $p_{50})$, that is (A), (B), (C) for $(1.0, 500)$, $(1.0, 2000)$, and $(2.0, 2000)$, respectively. A total of 441 contact history patterns are explored for each parameter set of $(\delta_{max}$ and $p_{50})$ to determine the viral evolution to optimize R_{TP} in a given environmental scenario. We conducted a total of 500 independent iterations for each parameter combination of k and θ as well as in Fig.3 in the main text. Numbers 1, 2, 3, and 4 in the panels of viral load are the corresponding dynamics illustrated as white dots in the columns of p , ϵ , the duration of infectious period, and R_{TP} , respectively. In the panels of R_{TP} , the vertical dashed line separates the area of mean contact number by an index, $R_{TP} = 10^0 (= 1)$ described by the horizontal dashed lines. The corresponding dashed line is also illustrated in the panels of p , ϵ , and the duration of infectious period.

Figure H: Different function types in the duration of infectious period assumed to depend on the cumulative viral load.



The black line represents the sigmoidal relationship calculated by Eq.(5) shown in the main text, and the function shape is exactly the same as that used throughout the paper (i.e., using the parameters provided in Table A). For elucidating the monotonically decreasing relationship, we define the linear (blue line) and the exponential (red curve) relationship as follows: $D_{max} - mV_{total}$ and $D_{max}exp(-nV_{total})$, respectively. We have assumed that the maximum duration of the infectious period is the same (1000) among all the functions. For testing robustness in the results of the current model using the linear and exponential relationship, see the following Fig. I and J. Parameters include $m = 5.0 \times 10^{-4}$ and $n = 3.0 \times 10^{-6}$.

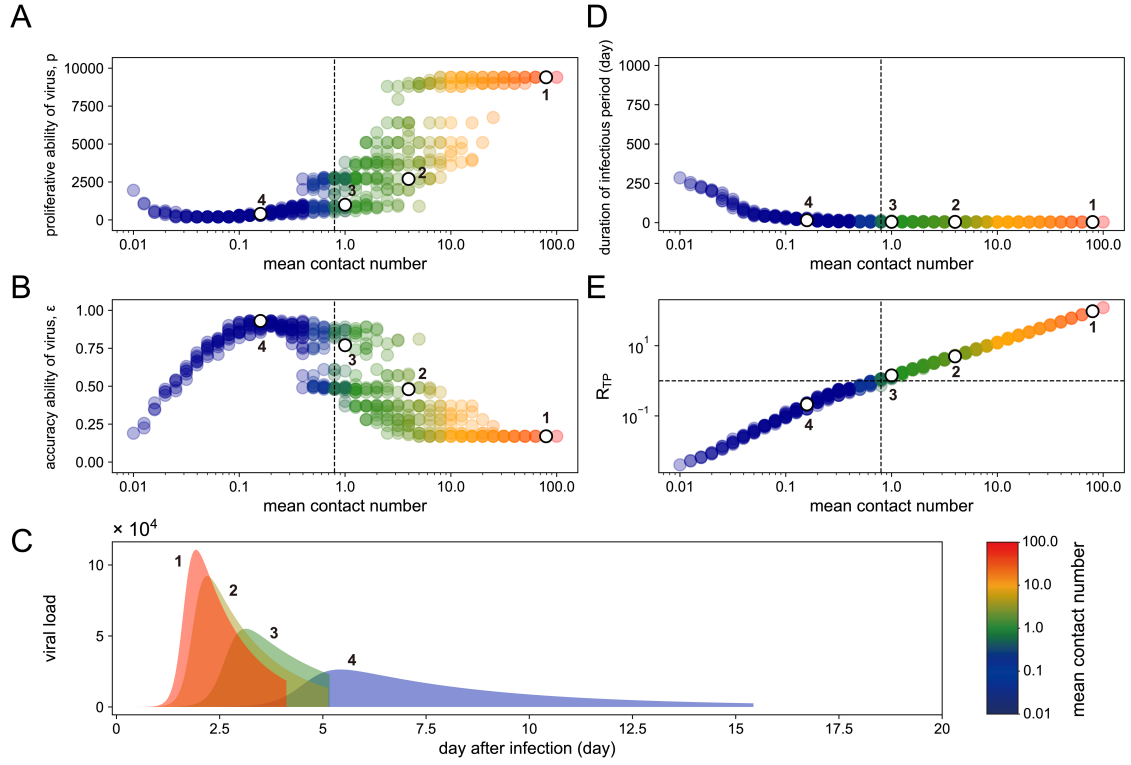
Figure I: Virus evolution in various scenarios of contact history where a monotonically decreasing linear relationship is assumed between the duration of the infectious period and the cumulative viral load.



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 θ are conducted to determine the optimized parameter set of p and ϵ characterizing final virus infection dynamics (i.e., different viral infection phenotypes). The median of the optimized parameter set of p and ϵ 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 θ . The combination of p and ϵ 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} by 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). Concerning the actual function shape for the duration of infectious period, see Fig. H.

Figure 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.



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 θ are conducted to determine the optimized parameter set of p and ε characterizing final virus infection dynamics (i.e., different viral infection phenotypes). The median of the optimized parameter set of p and ε 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 θ . The combination of p and ε 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} by 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), (D). Concerning the actual function shape for the duration of infectious period, see Fig. H.

Chapter3

Isolation may select for earlier and higher peak viral load but shorter duration in SARS-CoV-2 evolution

This study in this chapter was published in J. Sunagawa, H. Park, KS. Kim, R. Komorizono, S. Choi, LR. Torres, J. Woo, YD. Jeong, WS. Hart, RN. Thompson, K. Aihara, S. Iwami, R. Yamaguchi, “Isolation may select for earlier and higher peak viral load but shorter duration in SARS-CoV-2 evolution.” Nat Commun 14, 7395 (2023).

Introduction

The human impact of population densities and activities on the global environment has increased so dramatically that the current geological era has been termed the Anthropocene [1]. Human-mediated selection constitutes one of the most significant and pervasive selective pressures on Earth, changing at a pace that requires rapid evolution of adaptive responses by all organisms [2]. COVID-19-related restrictions and the resultant changes in human activities created a phenomenon termed “anthropause,” that is, a considerable global slowing of human activities and the effects of human activity on nature [3].

Human history has always been interwoven with viruses. Viruses debilitate many people and can have large-scale demographic effects on human populations according to immunity and prior disease exposure [4]. On the other hand, humans are an essential arena in which viruses evolve. Changes in human population size, immunity, and behavior based on public health policy can facilitate the rapid evolution of viruses [5].

We are facing the ongoing rapid emergence and adaptation of SARS-CoV-2. The virus was initially discovered in Wuhan, China, in late December 2019 (B lineage, strain Wuhan-Hu-1) and spread worldwide. Virus bearing the D614G substitution on the surface of the spike protein emerged in 2020 and became dominant among the circulating SARS-CoV2 variants [6]. The Alpha variant (B.1.1.7), first detected in the UK at the end of 2020, displayed a 43–90% higher reproduction number than pre-existing variants [7]. Interestingly, however, Alpha faded away with the rise of the more transmissible Delta variant (B.1.617.2) [8,9], which was first documented in India in October 2020. Then, the Omicron variant (originally B.1.1.529), which was first isolated in South Africa on October 2021, spread aggressively. A range of Omicron subvariants (BA.1, BA.2,

BA.4, and BA.5) eclipsed the Delta variant, and the Omicron variant became the predominant variant worldwide after February 2022 [10]. Currently, several new Omicron subvariants (BQ.1, XBB, and others) are emerging and are reported to be more transmissible and resistant to immunity generated by previous variants or vaccinations [11]. Collectively, these observations indicate that SARS-CoV-2 has continuously been evolving and variants have continued to emerge to replace existing viral strains worldwide, as observed with influenza [12].

An understanding of the epidemiological and clinical characteristics of current and future emerging infectious diseases is important for developing adaptive treatments, including antivirals and vaccinations, and screening and isolation strategies. Thus, evaluating and predicting how viral dynamics changes throughout infection through evolution is essential [5,13–16]. An unprecedentedly large volume of high-quality data have accumulated during the COVID-19 pandemic, and here we analyze the data on SARS-CoV-2 variants to benefit preparedness for future pandemics. As we discuss elsewhere [15–18], quantifying and comparing the timing and height of peak viral load and the duration of viral shedding among SARS-CoV-2 variants are of critical importance. In addition, understanding the driving forces behind viral evolution is also required, given the different selective pressures acting on the virus [5]. In general, the infection- and/or vaccine-induced immune response to antiviral drugs (i.e., pharmaceutical interventions: PIs) leads to virus evolution. The strongest evidence for selection based on human intervention is the rapid evolution of immune-escaping mutations in the Omicron variant, which often occur in parallel and can be predicted ahead of time [19]. In contrast, nonpharmaceutical interventions (NPIs), such as isolation, quarantine, social distancing, and wearing a face covering, efficiently prevent close contact, in particular, in symptomatic patients, given that most cases are the result of community transmission. NPIs have proven effective in reducing the spread of SARS-

CoV-2 in many contexts [20–25]. The introduction of strong selection pressure by NPIs and the evolutionary impact of this has been the focus of much attention [26–28], but how viral load dynamics are altered in vivo has not yet been explored. We were thus interested in the role of isolation as a strong NPI and explored the possible impact of such human behavioral changes on SARS-CoV-2 evolution.

So far, statistical analyses have shown that SARS-CoV-2 viral kinetics are mainly dominated by individual-level variation [29], but the kinetics may be partly determined by immunity and variant [30]. Here, to explore the isolation-driven viral evolution, we first quantified the viral dynamics from existing data on the viral load over the course of SARS-CoV-2 infection for the pre-Alpha (non-variants of interest/variants of concern [VOI/VOCs] types) [14,31], Alpha, and Delta variants (i.e., individual-level virus infection model) considering both individual and variant-specific variations in kinetics. Then, to explain and predict the evolutionary patterns of SARS-CoV-2 variants in terms of the time-series pattern of viral load, we developed a multi-level population dynamics model by coupling a population-level virus transmission model with the individual-level virus infection model. Under different clinical phenotypes for COVID-19 patients, defined here as the incubation period and the proportion of symptomatic infections, we evaluated how the time-series patterns of viral load and, therefore, SARS-CoV-2, evolve under isolation of infected individuals. We interestingly demonstrate that NPIs may select for earlier and higher peak viral load but shorter duration of infection in SARS-CoV-2. Although there are multiple ways to cause the recent evolutionary trajectory, the concepts and variables that affect the current transition can be anticipated. We discuss the potential evolution of SARS-CoV-2 as it adapts to maximize transmissibility in the presence of a human behavior change.

Results

Characterizing the time-series pattern of viral load for SARS-CoV-2 variants

We analyzed the viral load of SARS-CoV-2 for the pre-Alpha (non-VOI/VOCs) [14,31], Alpha [31,32], and Delta variants [31,32] using a simple mathematical model describing viral dynamics [15–18] (see **Table S1** and **METHODS**). To consider inter-individual variation in the patients' viral loads, we employed nonlinear mixed effect models to estimate the parameters. The estimated population parameters are shown in **Table S2**. The typical behavior of the mathematical model, Eqs. (4, 5), using these parameters is shown in **Fig. 1A** and **Fig. S1** for the pre-Alpha (black), Alpha (blue), and Delta variants (red), respectively. Note that the time of infection (i.e., $t = 0$) is estimated in our model fitting by using fixed initial values [33] (see **METHODS**). These quantified virus infection dynamics highlight how the properties of the virus differ variant by variant.

To characterize and compare the infection dynamics of these SARS-CoV-2 variants, we quantified the distributions of the peak viral load, peak time, and duration of viral shedding, D , (**Fig. 1B–D**). Here duration of viral shedding indicates the time for which the viral load is detectable. We found that peak time did not differ significantly between the pre-Alpha and Alpha variants (**Fig. 1C**). The peak viral load also was not significantly different between the Alpha and Delta variants (**Fig. 1B, D**). However, interestingly, we found that the peak viral loads for the Alpha and Delta variants were higher than for the pre-Alpha variants ($p = 2.9 \times 10^{-7}$ and 8.8×10^{-6} by the Mann-Whitney-Wilcoxon test, respectively) (**Table 1**). The peak time and duration of viral shedding for the Delta variant were shorter than those for the pre-Alpha and Alpha variants (for the peak time: p -values are less than 2.2×10^{-16} for the duration of viral shedding: $p =$

1.4×10^{-15}

and 1.1×10^{-6} by the Mann-Whitney-Wilcoxon test, respectively) (**Table 1**). Taken together, our results highlight that the duration of viral shedding was shortened and the peak viral load was increased and advanced as SARS-CoV-2 evolved. This implies that SARS-CoV-2 is evolving to a more “acute phenotype” [34], mainly characterized by an advanced peak time and a shorter duration of viral shedding compared with the pre-Alpha variants.

To further characterize the time-series pattern of viral load for SARS-CoV-2 variants during evolution of the virus (i.e., the Alpha variant emerged during the spread of the pre-Alpha variants, and the Delta variant emerged after the spread of the Alpha variant), we first investigated the relation between the duration of viral shedding, D , and the cumulative log-transformed viral load, V_{total} , defined as the area under the curve (i.e., AUC [$\log_{10}$ viral RNA copies/ml $\times$ days]) (**Fig. 1E**). We found V_{total} increased as D increased (i.e., positive correlation) (see **METHODS**). Interestingly, there was also a trend that V_{total} decreased as SARS-CoV-2 evolved: V_{total} for the Delta variant was smaller than that for the pre-Alpha and Alpha variants ($p = 1.1 \times 10^{-7}$ and 9.4×10^{-6} by the Mann-Whitney-Wilcoxon test, respectively). In addition, to understand the epidemiological consequences of these variants, we also calculated the transmission probability of each variant depending on the time after infection (see **Fig. 1F** and **METHODS**). We found that the transmission probability of the Alpha variant remained a high transmission probability for a longer period of time than the pre-Alpha variant. Furthermore, the transmission probability of the Delta variant peaked at 3.0 days after infection [9] but decreased to almost 0 by 8.0 days, suggesting that evolutionary process promotes effective and rapid transmission within human populations (see below for details).

Prediction on time-series pattern of viral load of SARS-CoV-2 evolution with NPIs

In the comparison of viral shedding dynamics, the Alpha variant showed a higher peak viral load than the pre-Alpha variants (**Fig. 1B**). Although the duration of viral shedding for the Alpha variants was less than that for the pre-Alpha variant (**Fig. 1D**), the transmission probability of the Alpha variants remained higher for a long time (**Fig. 1F**). Note that a similar difference between the pre-Alpha and Delta variants can be observed (**Fig. 1B, D, F**). These findings demonstrate that the Alpha-like and the Delta-like variants showed a high peak viral load but a relatively short duration of viral shedding as the virus evolved among human populations.

In contrast, in a comparison of the Alpha and Delta variants, the selective force to facilitate the evolution of the Delta-like variant is not trivial. This is because the peak viral load and the duration of viral shedding were similar between the Alpha and Delta variants (**Fig. 1B, D**), and therefore the highest transmission probability maintains almost the same level and length for these variants (**Fig. 1F**). The difference between the Alpha and Delta variants is the peak time, that is, the peak time of the Delta variant was significantly earlier than for the Alpha variant (**Fig. 1C**). Our hypothesis is that the Delta-like variants showing an earlier peak time but similar peak viral load and duration of viral shedding (i.e., acute phenotype) had a general advantage under human-mediated selection pressures (i.e., earlier infections result in more descendants). COVID-19 patients show both symptomatic and asymptomatic infection, and the incubation period and the proportion of asymptomatic patients are decreasing and increasing, respectively, as SARS-CoV-2 evolves [35,36]. To prevent the chain of transmission, patients showing symptoms are

usually isolated or hospitalized or subjected to other NPIs [37,38]. Symptomatic patients primarily generate secondary infections prior to symptom onset, that is, during the pre-symptomatic period. On the other hand, asymptomatic patients can be infectious during the whole duration of viral shedding. Below we show that a strong NPI, defined as the isolation of symptomatic infected individuals after symptom onset, could be a driving force for SARS-CoV-2 evolution. Note that although we can easily incorporate variant-specific virological characteristics such as infectivity per virus by changing the probability of transmission (see **METHODS**), we ignore these differences here for simplicity.

NPI-driven SARS-CoV-2 evolution

Using the probabilistic multi-level model, which considers both individual-level virus infection dynamics and population-scale transmission, as shown in Eqs. 1–3, 6, and the genetic algorithm (GA), which mimics virus phenotypic evolution [34], we explored how the time-series patterns of viral load evolved under NPIs. The schematic diagram of our multi-level modeling is described in **Fig. 2**, and a detailed explanation is provided in the **METHODS**.

Briefly, from the assumption that the virus evolves to increase its “transmission potential”, defined as R_{TP} (i.e., the total number of secondary cases generated throughout the infectious period), we calculate R_{TP} under various patterns of infection rate, β , and virus production rate, p . Thus, the transmissibility fitness landscape of R_{TP} is constructed as a function of (β, p) , given the incubation period (i.e., time from infection to symptom onset), T^* , and the proportion of symptomatic patients, f . Note that the transmissibility fitness landscape of R_{TP} is constructed as a function of only β and p , because the other parameters or initial values are estimated or determined (see **METHODS**). Under this analysis, we

considered the (additional) effect of isolation, which is only applicable to symptomatic individuals and which assumes that isolation always perfectly prevents the transmission chain (i.e., perfect isolation; $T^* < t$). In other words, the (baseline) effect of isolation, which is applicable to all individuals regardless of symptomatic or asymptomatic infection (e.g., wearing a face mask, social distancing), is assumed to be involved in the parameters in our multi-level model without loss of generality.

In the absence of symptomatic infection (i.e., all patients are asymptomatic: $f = 0$), we analyze SARS-CoV-2 evolutionary dynamics by exploring the optimal set of β and p , which characterizes the time-series pattern of viral load, with GA (see **Table S3, S4** and **Algorithm S1**). Without symptomatic infection, there is no effect of symptom-dependent isolation on the transmission chain at all. The optimal sets of (β, p) , the transmissibility fitness landscapes, the trajectories of R_{TP} along the course of GA, and its distribution are described in **Fig. S2A–C**, respectively. The time-series patterns of viral load with the optimal parameters and the corresponding timing of peak viral load (i.e., peak time) are described in **Fig. S2D, E**, respectively.

In contrast, in the absence of asymptomatic infection (i.e., all patients are symptomatic; the fraction of symptomatic infection is $f = 1$), all individuals lose their transmissibility by NPIs after symptom onset ($T^* < t$). Depending on the incubation period, T^* , the distributions of optimal sets of (β, p) are located in different parameter ranges, and the corresponding transmissibility fitness landscapes are altered (**Fig. S3A, B**). Because isolation can prevent the transmission chain before the focal infected individuals became highly infectious, given a small T^* such as $T^* = 1.0$ day (i.e., almost no transmissions occur during the pre-symptomatic period), R_{TP} with the optimal parameters is basically less than 1 (**Fig. S3B**). On the other hand, given a large T^* (e.g., $T^* = 10.0$

days), isolation is delayed, and therefore the transmissions are allowed during the pre-symptomatic period. Thus, R_{TP} increases as T^* increases (**Fig. S3B**). Note that the effect of isolation is weak for $T^* = 10.0$ because the viral load already drops to a very low level at the moment when isolation is applied to the symptomatic patients, and thus the transmission probability is low. In fact, the virus evolution shows similar patterns observed in the absence of symptomatic infection ($f = 0$) (**Fig. S2A–C** vs. **Fig. S3A–C** for $T^* = 10.0$, for example). The time-series patterns of viral load and the corresponding timing of peak viral load are also described in **Fig. S3D, E**, respectively. Interestingly, compared with the case where all patients are asymptomatic ($f = 0$), acute phenotypes showing an earlier peak but similar peak viral load and duration of viral shedding are selected (i.e., isolation-driven evolution). Thus, in the absence of asymptomatic infection, selection generated by isolation is maximized and drives convergent evolution of viral load dynamics (**Fig. S3D**). Additionally, because of smaller transmission potential (i.e., secondary [focal] infectious individuals may not be effectively generated and sometime less than one), isolation-driven evolution might be unlikely to occur under this situation.

In the realistic situation, there are both asymptomatic and symptomatic infections (i.e., $0 < f < 1$), and only symptomatic patients interfere with secondary infection by isolation after symptom onset. We next evaluated virus evolution under the different incubation period of T^* with the fixed fraction of $f = 0.3$ (i.e., 30% of infected individuals show symptoms). Note that there is substantial uncertainty in the value of f , and we conduct supplementary analyses for different values of f (see below). As described in **Fig. 3A**, the optimal sets of (β, p) are widely distributed for $T^* = 1.0$ and $T^* = 3.0$ days, but those are narrowed for $T^* = 6.0$ and 10.0 days. This is because the transmissibility fitness landscapes of viruses differ between asymptomatic and symptomatic patients in the case of relatively small incubation periods, whereas the difference is small for

large periods (**Fig. 3B**). We note that having a higher peak viral load while maintaining a rather long duration favors an increase in R_{TP} , and viral dynamics with this characteristic are achieved when β is low and p is relatively high, such as optimal sets of (β, p) for the asymptomatic case (or in the absence of isolation) in **Fig. 3B**.

In the search for the optimal parameter set by GA, depending on the type of focal infectious individual (who is stochastically chosen), directions of virus evolution stochastically change so that neither peak of the transmissibility fitness landscape can be unreachable (compare the colored dots in **Fig. 3A** with the white dots in **Fig. 3B** for $T^* = 1.0$ and 3.0) unless the evolutionary rate is extremely fast. Thus, an intermediate parameter set becomes the equilibrium state. Moreover, we calculated R_{TP} for each infected individual: the trajectories of R_{TP} with the optimal parameters vary around 1, especially with short incubation periods (**Fig. 3C**), because R_{TP} depends on the types of focal infectious individual and R_{TP} for the symptomatic individuals is basically smaller than that for the asymptomatic individuals (sometimes less than 1) because of isolation (**Fig. 3B**).

As observed in the absence of asymptomatic infection, the time-series patterns of viral load (**Fig. 3D**) with the optimal parameters (**Fig. 3E**) also depend largely on the incubation period T^* because of isolation. Of note, in the realistic situation with $0 < f < 1$, because (secondary) infectious individuals are mainly maintained via transmissions by asymptomatic individuals, isolation-driven evolution for relatively small incubation periods may occur, and the peak viral load is advanced. **Fig. 3D, E** also demonstrate that faster peak time is associated with larger value of the product of p and β . This is due to the combined effect of larger p and β values, resulting in both rapid and increased amount of virus production during an early phase of infection. For our sensitivity analysis,

we conducted the same simulations for $f = 0.5, 0.6, 0.7$ and obtained similar conclusions (see **Fig. S4**).

We have thus far assumed ‘perfect isolation’ in our results, where transmission is completely halted once patients become symptomatic. However, we understand that in practice, isolation is sometimes imperfect, leading to continued virus transmission. To account for this, we introduced a factor ξ , modifying the contact numbers post symptom onset ($C^*(t) = C(t) \times \xi$), with ξ representing incomplete isolation probabilities of 0.1 and 0.3. In general, this extension yielded consistent outcomes with our initial findings (**Fig. S5**). However, it indicated that incomplete isolation leads to a less pronounced evolution towards an earlier, higher peak in viral load dynamics at a small T^* (such as $T^* = 1$: it is unreasonably small in the context of COVID-19 and therefore we may exclude this situation thought) (**Figs. S5–1**). Even in cases where T^* is small, and it significantly drives virus evolution, the deterministic selective pressure is weakened due to the probabilistic failure of isolation. This situation is analogous to a scenario where all patients have asymptomatic infections as a result, making advancing the peak viral load no longer beneficial because of the short pre-symptomatic period. The trend becomes more prominent when the probabilistic failure of isolation occurs frequently (e.g., $\xi = 0.3$ in **Figs. S5–2**).

As we confirmed in **Figs. 3, Fig. S2, S3, and S4**, the higher proportion of symptomatic infection enhances the selection pressure for virus evolution via isolation. Taken together, these results suggest that while isolation inhibits the chain of transmission from symptomatic focal infectious individuals after symptom onset, in the realistic situation, SARS-CoV-2 evolves toward the acute phenotypes.

Effect of prior immunity on NPI-driven SARS-CoV-2 evolution

Considering another modeling assumption, our initial mathematical model did not account for the influence of COVID-19 vaccinations or prior infections. Given the widespread infection and vaccination leading to high-level population immunity worldwide, this factor could have a profound impact on the evolution of the virus. Some studies suggest that COVID-19 vaccinations may affect the SARS-CoV-2 infection dynamics [32,39]. Accordingly, we conducted further analysis to investigate the impact of prior immunity caused by vaccinations or prior infections on our findings.

We first reanalyzed SARS-CoV-2 infection dynamics in **Fig. S6**, including additional 68 participants (14 and 54 patients for the Alpha and the Delta variants, respectively), who were excluded due to prior infections, and obtained the consistent statistical trends in line with our findings described in **Fig. 1A**, reinforcing the robustness of our conclusions. Specifically, we observed that the duration of viral shedding keeps decreased in evolution. The peak time did not exhibit significant differences between the pre-Alpha and Alpha variants, and the peak viral load did not significantly differ between the Alpha and Delta variants. However, the peak viral loads for both the Alpha and Delta variants were higher compared to the pre-Alpha variants ($p = 0.002$ and 3.3×10^{-7} , respectively). Moreover, Delta variant exhibited a shorter peak time in comparison to both the pre-Alpha and Alpha variants ($p = 2.2 \times 10^{-16}$ for both cases) (**Fig. S6**). These confirmed that the statistical trend for SARS-CoV-2 evolution is remained regardless of the prior infections.

Furthermore, we extended our *in silico* analysis to incorporate a vaccinated population, considering vaccination rates set probabilistically at 50% and 80%. For those vaccinated, we increased the clearance rate (δ) by 1.5 times. This assumption for clearance rate shortened the viral shedding period by about 3 days, and this is larger than the previously reported shortening period of 2 days [32]. Thus, a 1.5-fold increase in clearance rate

is sufficient to account for an increase in clearance rate caused by the vaccination. The incorporation of this stochastic environment led to a modest increase in the variance of optimal virus parameters. However, it is essential to note that these changes in the model maintained our main conclusion. The mean values derived remained consistent with our original findings: viruses were selected for had an earlier peak and higher viral load dynamics, but a shorter duration of infection (**Fig. S7**). The selective pressure for increased transmissibility shapes the viral load dynamics, and isolation measures are likely to be a critical driver of these evolutionary transitions even considering the prior immunity.

Validating SARS-CoV-2 evolution with Omicron variants

The Omicron variant (originally B.1.1.529) was first isolated in South Africa in October 2021 [10,40]. After that time, of the identified Omicron subvariants, the BA.1 subvariant rapidly spread worldwide and became the most prevalent SARS-CoV-2 variant in many countries (**Fig. 4A**). Here, we consider simply the situation that BA.1 emerged during the spread of the Delta variant, which was the most predominant variant worldwide before BA.1, under NPIs.

To validate our hypothesis that isolation may be a selection pressure to further drive the evolution of viral phenotypes, we additionally used the longitudinal viral load of SARS-CoV-2 from the BA.1 subvariant in 49 infected patients [29] (**Table S1**), and similarly analyzed those data together with the same data for the Delta variant (see **METHODS** in detail). The individual viral loads for the Delta variant (red) and BA.1 subvariant (green) are described in **Fig. S8**, and the estimated population parameters are shown in Table S5. We found that the time-series pattern of viral load for the BA.1 subvariant differed from that for the Delta variant as described in **Fig. 4B**. We further characterized and compared the

distributions of the peak viral load, peak time, and duration of viral shedding in **Fig. 4C**. Interestingly, while we found almost the same peak time for the BA.1 subvariant viral load with a similar duration of viral shedding ($p = 0.09$ and 0.48 by the Mann-Whitney-Wilcoxon test, respectively), the comparison of peak viral load suggested a lower peak for the BA.1 subvariant than for the Delta variant ($p = 1.3 \times 10^{-6}$ by the Mann-Whitney-Wilcoxon test). On the other hand, as we expected, the BA.1 subvariant showed an earlier peak time and shorter viral shedding compared with the pre-Alpha and Alpha variants (see **Fig. 1A–D** vs. **Fig. 4B, C**). Altogether, these results imply that SARS-CoV-2 evolution toward an earlier peak viral load was maintained but slowed down. Interestingly, similar evolutionary trends were confirmed in a non-human primate model infected with different VOCs of SARS-CoV-2 [41], obviously without NPIs in these cases. We discuss a mechanism behind those convergent phenotypic changes in detail later (see Discussion). Our quantitative empirical clinical data analysis based on our hypothesis could support that SARS-CoV-2 variants have evolved their acute phenotypes via a human-mediated selection pressure such as isolation.

Discussion

Understanding the entire arc of a virus's impacts on an outbreak requires knowledge of how the virus and human populations interact and placing the relationship within not only immunological but societal and cultural contexts. Here we quantified the infection dynamics of the SARS-CoV-2 variants from the patients' longitudinal viral load data. In the process of evolutionary transitions among the pre-Alpha, Alpha, Delta, and Omicron variants, we found the durations of viral shedding were shortened, except for the Omicron variant, and the peak viral load was increased (from the pre-Alpha to Alpha variants) and advanced (from the Alpha to Delta variants) (**Fig. 1A–D**). Increasing peak viral load (unless individuals become so sick that they do not circulate) is obviously advantageous for virus evolution because it effectively increases transmission probability (**Fig. 1F**). In contrast, we interestingly demonstrated that the advantage of the advancing peak viral load may depend on clinical phenotypes (i.e., the incubation periods: T^* and the proportion of symptomatic: f) for virus variants and the community environment (i.e., with/without isolation) (**Fig. 3, Figs. S2 and S3**). Hence, using quantitative empirical data analysis and the multi-level population dynamics model, we suggested that isolation potentially plays the role of a human-mediated selection pressure for SARS-CoV-2 evolution.

A higher proportion of asymptomatic infection and a longer incubation period than for other acute respiratory viruses are notable features of COVID-19 and are major factors increasing the basic/effective reproduction number [42,43]. The proportion is reported to be around 30% [44,45], and the incubation period is 2 to 12 days [46]. Because transmission via symptomatic individuals is usually prevented by isolation (and other PIs/NPIs), SARS-CoV-2 is mainly transmitted during the asymptomatic and pre-symptomatic periods, implying that advancing the

peak viral load can effectively generate more secondary infections. As we explained in **Fig. 3**, under isolation, early peak viral loads such as for the Delta and Omicron variants are beneficial to variants having a shorter incubation period, because the effect of isolation on reducing transmission probability is mitigated and therefore R_{TP} increases (see **Fig. S9**). Thus, in addition to PIs, isolation could be one of the driving forces promoting virus evolution as the Delta variant dominated after the spread of Alpha. In fact, broken down by VOCs, the incubation period of COVID-19 was 6.65 days for the pre-Alpha, 5.00 days for the Alpha, 4.50 days for the Beta, 4.41 days for the Delta, and 3.42 days for the Omicron variants [36,47]. The incubation period decreased as SARS-CoV-2 mutated, implying the selection pressures by NPIs enhanced virus evolution, as we confirmed in **Fig. 3**.

In contrast, when comparing the recently emerged Omicron variant with other variants, it has been reported that the Omicron variant is evolving to show a higher proportion of asymptomatic infection [35]. Whereas a shorter incubation period enhances an earlier peak viral load as discussed above, a higher proportion of asymptomatic infection impairs the driving force of virus evolution, meaning a simple one-directional virus evolution (i.e., earlier and higher peak viral load) does not occur. Interestingly, these trends are observed for the subvariant BA.1 in **Fig. 4B, C**, that is, decreasing the peak viral load. From the perspective of virus evolution, which maximizes the basic/effective reproduction number (or an average of the transmission potential for symptomatic and asymptomatic infection R_{TP} in this study, for example), the evolution of a higher proportion of asymptomatic infection is always favored because transmissions are not limited when those infected individuals are not isolated. This means that once the higher proportion of asymptomatic infection evolves, other viral phenotypic changes designed to avoid isolation will be suppressed. In other words, viral phenotypes showing

slower and/or lower peak viral load can evolve (see **Fig. S9**). This perspective grounded on the assumption that the evolution of virus phenotype and the evolution of clinical phenotype may often be decoupled. While we believe that this assumption can be accepted [48,49], we also acknowledge that it should be noted that a consensus applicable to all cases has not been established [48,50,51].

A reasonable expectation during the COVID-19 pandemic is that the virus could evolve to develop increased transmissibility, reflecting adaptations to propagation in the new human host [5] and the rapidly changing environments by PIs and NPIs. Taken together, as described in **Fig. S9**, we found there are at least two antagonistic forces enhancing and mitigating the earlier and higher peak viral load evolution via isolation during the process of SARS-CoV-2 evolution from the pre-Alpha to the Omicron variants via the Alpha and Delta variants: decreasing the incubation period and increasing the proportion of asymptomatic infection, respectively. Although both forces may increase transmissibility, the balance between these two forces mitigates or stops the evolution of SARS-CoV-2 phenotypes defined as the time-series patterns of viral load. There is a possibility that the strength of NPIs per se could be weakened worldwide as interest in COVID-19 drops off, and therefore the mitigated selection pressure by NPIs allows the observed trend of the lower peak viral load of the BA.1 subvariant (which provides benefits at the host individual level such as evading host immunity) in **Fig. 4B, C**. However, the current framework of our multi-level model cannot explain the evolution of a “stealth” phenotype showing a decreased peak viral load with a high proportion of asymptomatic infection. A further detailed analysis is required once numerous datasets of different Omicron subvariants, such as BA.2, BA.4, BA.5, BQ.1, and XBB, become available.

Our study acknowledges certain limitations. Firstly, we assumed

that the incubation period, T^* , and the proportion of symptomatic infected persons, f , are given as “environment” parameters, meaning we ignored the relation between virus evolution and clinical phenotypes. We expect that this is a reasonable assumption because recent SARS-CoV-2 human challenge clearly showed no quantitative correlation between patients’ time-series pattern of viral load and symptoms [49]. Second, coevolutionary dynamics between hosts and viruses may result in long-term non-equilibrium evolutionary dynamics. A recent theoretical study revealed that viruses that often show antigenic escape from host immunity evolve through non-equilibrium dynamics [52], indicating that transmissibility fitness R_{TP} is no longer the appropriate measure to understand patterns of virus evolution. Although the current framework does not consider such coevolutionary dynamics explicitly, we successfully captured transmission in viral load dynamics throughout the pandemic, which suggests that in our case transmissibility potential is a useful proxy. This may be because SARS-CoV-2 mutates slowly enough (on average, SARS-CoV-2 evolves 3-4 times slower than influenza virus [53]) to consider the long-term evolutionary equilibrium as in the classic context of endemic diseases. Third, our empirical data analysis and the multi-level simulation model do not explicitly account for the influence of population-level immune status on the evolution of SARS-CoV-2. Our additional analysis (**Figs. S6 and S7**), taking into consideration individuals with prior immunity, can alleviate concerns that individual host immune status significantly affects infection dynamics. However, the collective immune status at the population level can also shape the environment in which the virus evolves [54,55]. While strictly demonstrating the influence of population-level immune status on viral evolution through empirical data analysis may pose challenges, we maintain that our research offers significant insights into the potential interplay between human behavior changes and viral evolution.

The last but not least caveat for our modeling is the possibility that the evolutionary speed of SARS-CoV-2 and other potential selection pressures could generate the same evolutionary pattern presented here. The speed of evolution of a viral pathogen depends not only on the background mutation rate but also the virus generation time, the duration of infection, the number of variants that develop during the infection of an individual, etc [5]. If mutations occur too slowly, the virus prevalence decays prior to the appearance of a beneficial mutation that makes transmissibility higher. On the other hand, if mutations occur too rapidly, the pathogen evolution becomes volatile and, once again, the virus fails to spread [56]. Thus, mutation-selection balance is not negligible when discussing the evolutionary and epidemic outcomes in the future. We assumed that isolation truncated transmissibility time-series so that an earlier peak was selected for in a plausible way. However, there may be another selection pressure to increase the number of secondary infections by increasing the initial slope of viral load dynamics, which also leads to an earlier peak. In general, which functional mutations can occur, how frequently they can occur, and how favorable they are compared with other strains should all be combined to predict the future trajectory of COVID-19.

In conclusion, using longitudinal viral load data, we demonstrated that an earlier and higher peak viral load but a shorter duration were selected for during SARS-CoV-2 evolution. We hypothesized that this evolution of SARS-CoV-2 infection dynamics might be associated with non-pharmacological interventions. However, it was challenging to strictly substantiate this hypothesis with current available data. Thus, we designed a probabilistic multi-level model, and our simulations using this model can inform hypotheses about the mechanisms that underlie virus evolution under isolation according to clinical phenotypes. Because NPIs are among the best ways of controlling a pandemic when a vaccine is not yet available, isolation as a strong NPI might be the first choice for a primary

mitigation strategy for current and future emerging infectious disease. The impact of human behavior change on patterns of virus evolution must be considered when evaluating future scenarios of COVID-19. It would also be essential to investigate evolution considering economically preferred interventions, such as social distancing and mask-wearing, as well as isolation and combinations of these measures [57]. Although several types of COVID-19 vaccine are now available, many breakthrough infections are still being observed. Therefore, NPIs still play an important role in controlling the COVID-19 pandemic. It is also expected that new variants of concern after Omicron will become dominant. Thus, assuming that new variants and other pathogens will emerge in the future, it is important to consider the evolution of pandemic viruses when shaping public health strategies.

Methods

Study data

We searched for longitudinal viral load data from COVID-19 patients categorized with pre-Alpha, Alpha, Delta, and Omicron variants in PubMed and Google Scholar. Data meeting the following criteria were used to estimate parameters of the viral dynamics model: (1) viral loads were assessed at two time points or more and (2) samples were collected from the upper respiratory tract, such as from the nose or pharynx. Because all the data were extracted from published papers, ethics approval was not required in this study.

Four papers were identified that met the inclusion criteria. In total, viral load data from 86 patients infected with pre-Alpha variants (non-VOI/VOCs types), 59 Alpha variant patients, 80 Delta variant patients, and 49 Omicron (BA.1) variant patients were used for parameter estimation of the viral dynamics model. Three studies were from the USA and the other was from the UK (**Table S1**). First, viral load data from two previously published papers were used [14,31]. All patients' SARS-CoV-2 infections were by pre-Alpha variants. One case was reported from the USA and the other was from the UK. The original number of patients was 68 from the USA and 49 from the UK. Among those, data for 29 patients from the USA and 2 from the UK were excluded because of insufficient data points. For the second dataset consisting of the Alpha variant, two published papers were used [31,32]. For those studies, data for 22 patients from the USA and 37 from the UK were accepted; data from 14 patients in the USA and 2 in the UK were not included because of insufficient data points. Third, data for 64 patients from the USA [29,32] and 16 from the UK [31] were used for the Delta variant analysis. Originally, 26 additional data were in the USA dataset, but these were not used because there were two data points or

fewer. Lastly, data from 49 patients collected in the USA [29] were used for the Omicron variant analysis.

Modeling SARS-CoV-2 infection dynamics

To describe SARS-CoV-2 dynamics among susceptible target cells, we used a simple mathematical model for virus infection dynamics to define viral load for each patient:

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

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

$$\frac{dV(t)}{dt} = pI(t) - cV(t), \quad (3)$$

The variables $T(t)$, $I(t)$, and $V(t)$ are the number of uninfected target cells, virus-producing infected cells, and the amount of virus at time t since infection, respectively. The parameters β , δ , p , and c represent the rate constant for virus infection, the death rate of infected cells, the viral production rate, and the clearance rate of the virus, respectively.

To estimate the parameter values from the longitudinal viral load data of COVID-19 patients, we further simplified Eqs. (1–3). Since the clearance rate of the virus ($c = 20$ is fixed here) is typically much larger than the death rate of the infected cells in vivo, we made a quasi-steady state (QSS) assumption, $dV(t)/dt = 0$, and obtained $V(t) = pI(t)/c$, which yields $I(t) = cV(t)/p$. Then, substituting this equation into Eq. (2), we have

$$\frac{dV(t)}{dt} = \frac{p\beta}{c} T(t)V(t) - \delta V(t)$$

Furthermore, we replaced $T(t)$ by the proportion of target cells remaining at time t , that is, $f(t) = T(t)/T(0)$, where $T(0)$ is the initial number of uninfected target cells. Accordingly, we obtained the following simplified mathematical model [15–18], which we employed to analyze the viral load data in this study:

$$\frac{df(t)}{dt} = -\beta f(t)V(t), \quad (4)$$

$$\frac{dI(t)}{dt} = \gamma f(t)V(t) - \delta V(t), \quad (5)$$

where $\gamma = p\beta T(0)/c$ corresponds to the maximum viral replication rate under the assumption that target cells are continuously depleted during the course of infection. Thus, $f(t)$ is equal to or less than 1 and continuously declines.

Nonlinear mixed effect model and test of the model fit

In our analyses, the variable $V(t)$ in Eq. (5) corresponds to the viral load for SARS-CoV-2. We separately analyzed two viral load datasets including the pre-Alpha, Alpha, and Delta variants and the Delta and Omicron variants because of different phases of virus evolution (see above). Note that the data for the Delta variant are common. We fixed values $V(t)$ and $T(t)$ at the time of infection as 10^{-2} and 1.33×10^5 , respectively [33]. To fit the patient's viral load data, we used MONOLIX 2021R2 (www.lixoft.com), implement maximum likelihood estimation of parameters in nonlinear mixed effect model. The nonlinear mixed effect model allows a fixed effect as well as interpatient variability. This method estimates each parameter $\theta_i (= \theta \times e^{\eta_i})$ for each individual where θ is a fixed effect, and η_i is a random effect, and which obeys a Gaussian distribution with mean 0 and standard deviation Ω . Here we used lognormal distributions as prior distributions of parameters to guarantee the

positiveness (i.e., negative values do not biologically make sense). In parameter estimation, since time 0 on the original data set is the peak viral load time, we also estimated time from infection to peak viral load (corresponding to T_p in **Table S2** and **S5**) along with other parameters. We used the SARS-CoV-2 variants as a categorical covariate in estimating the parameters β , δ , γ (as well as $p = \gamma c / \beta T(0)$), and T_p , which provide the lowest AIC. Applying a stochastic approximation expectation-approximation algorithm and empirical Bayes method, the population parameters and individual parameters were estimated, respectively. The estimated fixed parameters and the initial values are listed in **Table S2** for the pre-Alpha, Alpha, and Delta variants, and in **Table S5** for the Delta and Omicron variants, respectively. The viral load curve using the best-fit parameter estimates for each individual is shown with the data in **Figs. S1** and **S8**.

Multi-level modeling for virus evolution

We explored how the time-series patterns of viral load, including the duration of viral shedding and the amount and timing of peak viral load of COVID-19 patients, which are all characterized by SARS-CoV-2 infection dynamics, change during the process of SARS-CoV-2 evolution. To this purpose, we here developed a probabilistic multi-level population dynamics model by coupling a population-level virus transmission model and the individual-level virus infection model (e.g., ref. [34]).

First, as an individual-level virus infection model, we employed Eqs. (1–3) instead of Eqs. (4, 5). This is because all parameters in Eqs. (1–3) can be estimated from the data fitting of Eqs. (4, 5) to the longitudinal viral load of SARS-CoV-2 for each variant. We interestingly found a

positive correlation between the estimated death rate of infected cells, δ , and the viral production rate p (see **Fig. S10**). To mimic viral cytopathogenesis depending on the viral replication level, we defined $\delta = \delta_{max}p/(p + p_{50})$, that is, an increasing Hill function of p . The parameters δ_{max} and p are the maximum value of δ and the viral production rate satisfying $\delta = \delta_{max}/2$, respectively, and are estimated as summarized in **Table S3**. The dynamics of viral load calculated by this individual-level virus infection model will be translated into a probability of transmission and used in the following population-level model.

Next, as the population-level virus transmission model, we calculated the total number of secondary cases generated throughout the infectious period (corresponding to the duration of viral shedding, D) as R_{TP} , called “transmission potential”. One simple assumption is that the viral population will eventually be dominated by the virus with the largest 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, $C(t)$) and the transmission probability, $\rho(t)$, and therefore R_{TP} is defined as below (and explain its formulation below) (see **Fig. 2**):

$$R_{TP} = \sum_{t=0}^D (C(t)\rho(t)), \quad (6)$$

To mimic a daily contact history in Eq. (6), we assume the number of contacts on any day is drawn from a negative binomial distribution. Since a negative binomial distribution is identical to a Poisson distribution where the mean parameter λ 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 [58]:

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

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

where k and θ are the shape parameter and the scale parameter,

respectively (see **Table S3**). We note that k influences the skewness, and θ influences the variance of the distribution, respectively. The mean contact number through the history is the product of k and θ , that is, $k\theta$.

Once fixing the daily contact number, R_{TP} (i.e., Eq. (6)) is dependent on $\rho(t)$ and D , meaning the parameters appearing in Eqs. (1–3). To evaluate the daily transmissibility of an infected individual through the duration of viral shedding, we assumed that a probability of transmission (per contact with a susceptible individual), $\rho(t)$, depends on viral load. As reported in [59], we simply employed the step function as the range of viral load; thus, the transmission probability $\rho(t)$ at time t is described by

$$\rho(t) = \begin{cases} 0.00, & \text{if } \log_{10} V(t) < 5, \\ 0.12, & \text{if } 5 < \log_{10} V(t) < 7, \\ 0.15, & \text{if } 7 < \log_{10} V(t) < 10, \\ 0.24, & \text{if } 10 < \log_{10} V(t), \end{cases}$$

For example, $\rho(t)$ for SARS-CoV-2 variants are calculated based on Eq. (3) in **Fig. 1F**.

In addition, since there is a positive correlation between the duration of viral shedding (D) and the cumulative log-transformed viral load (V_{total}) for COVID-19 patients observed in **Fig. 1E**, we defined

$$D = \frac{D_{max} V_{total}^{D_K}}{V_{50}^{D_K} + V_{total}^{D_K}}, \quad (7)$$

$$V_{total} = \int_0^{T_{DL}} \log_{10} V(t) dt, \quad (8)$$

where T_{DL} is the time for the viral load to reach the detection limit ($DL = 1$ is fixed as an arbitrary value). The parameters D_{max} , V_{50} and D_K are the maximum duration of viral shedding, the total viral load at which the duration is half of its maximum, and the steepness at which duration increases with increasing viral load, respectively, and are estimated as summarized in **Table S3**.

We hereafter focus on the dependence of the rate constant for virus infection β and the viral production rate p in Eqs. (1–3), since the death rate of infected cells δ is function of p and the other parameters or initial values are fixed or common as explained above. We calculate R_{TP} in various patterns of (β, p) by Eqs. (1–3, 6), and thus the transmissibility fitness landscape of R_{TP} is constructed as a function of β and p .

Evolution of virus to increase transmission potential

When calculating the transmission potential R_{TP} , the focal infectious individuals are assigned one of two properties: symptomatic or asymptomatic, which is randomly assigned by a constant probability f or $1 - f$. We assume f to be independent of the viral load dynamics, because there is no significant difference in viral load between asymptomatic and symptomatic infection as discussed in [60,61]. We here simply assumed the transmission chains from the focal infectious individuals are perfectly inhibited by isolation after symptom onset. That is, if the focal infectious individual is symptomatic, we stopped calculating the cumulative number of secondary cases at symptom onset. Symptom onset starts after T^* days of infection, which represents the incubation period. Otherwise, R_{TP} is calculated using the whole duration of viral shedding of the focal asymptomatic infectious individual.

We analyze the virus evolutionary dynamics on the transmissibility potential landscape. The set of the value of (β, p) are explored in the sense of maximizing R_{TP} , using a genetic algorithm, GA [62] (see **Algorithm S1 and Table S4** for parameters used in GA). Initially, GA starts with an arbitrary set of parameters (β, p) and initial values (Initialize), and calculates R_{TP} as transmissibility fitness (Evaluate). Over successive iterations, the values of (β, p) with higher R_{TP} are selected to produce the next solutions (selectSolutions). Also, the two highest-

performing solutions are always chosen for the next iteration (Elitism). To find better values of (β, p) , 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), which corresponds to the rate of evolutionary change of the virus. Finally, by varying T^* and f , we explore a better set of values (β, p) that maximizes R_{TP} . Note that the set of values at the end of the GA run is not necessarily a global optimum. These iterations can be considered as the virus “evolving” its properties of β and p toward an optimal solution so that the virus can increase its own fitness in the evolutionary history.

Data availability

All viral load data that support the findings of this study are available at Zenodo [63] (<https://zenodo.org/records/10031081>).

Code availability

All analyses were performed with the software Python (version 3.10) and R (version 4.3.0). The analysis using nonlinear mixed effects model (including the algorithms for parameters estimation, such as Stochastic Approximation Expectation Maximization and Markov Chain Monte Carlo) was performed on MONOLIX 2021R2 (www.lixoft.com). Our code is publicly available at Zenodo [63] (<https://zenodo.org/records/10031081>)

References

1. Lewis SL, Maslin MA. Defining the Anthropocene. *Nature*. 2015;519: 171–180. doi:10.1038/nature14258
2. Palumbi SR. Humans as the World's Greatest Evolutionary Force. *Science*. 2001;293: 1786–1790. doi:10.1126/science.293.5536.1786
3. 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. *Nat Ecol Evol*. 2020;4: 1156–1159.
4. Athni TS, Shocket MS, Couper LI, Nova N, Caldwell IR, Caldwell JM, et al. The influence of vector-borne disease on human history: socio-ecological mechanisms. Chase J, editor. *Ecol Lett*. 2021;24: 829–846. doi:10.1111/ele.13675
5. Telenti A, Arvin A, Corey L, Corti D, Diamond MS, García-Sastre A, et al. After the pandemic: perspectives on the future trajectory of COVID-19. *Nature*. 2021;596: 495–504. doi:10.1038/s41586-021-03792-w
6. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*. 2020;182: 812-827.e19. doi:10.1016/j.cell.2020.06.043
7. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021;372: eabg3055. doi:10.1126/science.abg3055
8. Yang W, Shaman J. Development of a model-inference system for estimating epidemiological characteristics of SARS-CoV-2 variants of concern. *Nat Commun*. 2021;12: 5573. doi:10.1038/s41467-021-25913-9
9. 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: 603–610.

doi:10.1016/S1473-3099(22)00001-9

10. Elliott P, Bodinier B, Eales O, Wang H, Haw D, Elliott J, et al. Rapid increase in Omicron infections in England during December 2021: REACT-1 study. *Science*. 2022;375: 1406–1411.
doi:10.1126/science.abn8347
11. Qu P, Evans JP, Faraone J, Zheng Y-M, Carlin C, Anghelina M, et al. Distinct Neutralizing Antibody Escape of SARS-CoV-2 Omicron Subvariants BQ.1, BQ.1.1, BA.4.6, BF.7 and BA.2.75.2. *Microbiology*; 2022. doi:10.1101/2022.10.19.512891
12. Koelle K, Cobey S, Grenfell B, Pascual M. Epochal Evolution Shapes the Phylodynamics of Interpandemic Influenza A (H3N2) in Humans. *Science*. 2006;314: 1898–1903. doi:10.1126/science.1132745
13. Li J, Lai S, Gao GF, Shi W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature*. 2021;600: 408–418.
doi:10.1038/s41586-021-04188-6
14. Kissler SM, Fauver JR, Mack C, Olesen SW, Tai C, Shiue KY, et al. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. Riley S, editor. *PLOS Biol*. 2021;19: e3001333. doi:10.1371/journal.pbio.3001333
15. Jeong YD, Ejima K, Kim KS, Joohyeon W, Iwanami S, Fujita Y, et al. Designing isolation guidelines for COVID-19 patients with rapid antigen tests. *Nat Commun*. 2022;13: 4910. doi:10.1038/s41467-022-32663-9
16. Jeong YD, Ejima K, Kim KS, Iwanami S, Bento AI, Fujita Y, et al. Revisiting the guidelines for ending isolation for COVID-19 patients. *eLife*. 2021;10: e69340. doi:10.7554/eLife.69340
17. Kim KS, Ejima K, Iwanami S, Fujita Y, Ohashi H, Koizumi Y, et al. A quantitative model used to compare within-host SARS-CoV-2, MERS-CoV, and SARS-CoV dynamics provides insights into the pathogenesis and treatment of SARS-CoV-2. Sugden B, editor. *PLOS Biol*. 2021;19:

e3001128. doi:10.1371/journal.pbio.3001128

18. Iwanami S, Ejima K, Kim KS, Noshita K, Fujita Y, Miyazaki T, et al. Detection of significant antiviral drug effects on COVID-19 with reasonable sample sizes in randomized controlled trials: A modeling study. Kretzschmar MEE, editor. PLOS Med. 2021;18: e1003660. doi:10.1371/journal.pmed.1003660
19. Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature. 2022;602: 657–663. doi:10.1038/s41586-021-04385-3
20. Tian H, Liu Y, Li Y, Wu C-H, Chen B, Kraemer MUG, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science. 2020;368: 638–642. doi:10.1126/science.abb6105
21. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature. 2020;585: 410–413. doi:10.1038/s41586-020-2293-x
22. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020;368: 395–400. doi:10.1126/science.aba9757
23. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature. 2020;584: 262–267. doi:10.1038/s41586-020-2404-8
24. Leung NHL, Chu DKW, Shiu EYC, Chan K-H, McDevitt JJ, Hau BJP, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med. 2020;26: 676–680. doi:10.1038/s41591-020-0843-2
25. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health. 2020;5: e279–e288. doi:10.1016/S2468-2667(20)30090-6

26. Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. *Proc Natl Acad Sci.* 2020;117: 30547–30553. doi:10.1073/pnas.2013182117
27. Ali ST, Wang L, Lau EHY, Xu X-K, Du Z, Wu Y, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science.* 2020;369: 1106–1109. doi:10.1126/science.abc9004
28. Gurevich Y, Ram Y, Hadany L. Modeling the evolution of SARS-CoV-2 under non-pharmaceutical interventions and testing. *Evol Med Public Health.* 2022;10: 179–188. doi:10.1093/emph/eoac013
29. Hay JA, Kissler SM, Fauver JR, Mack C, Tai CG, Samant RM, et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: A retrospective cohort study. *eLife.* 2022;11: e81849. doi:10.7554/eLife.81849
30. Puhach O, Adea K, Hulo N, Sattonnet P, Genecand C, Iten A, et al. Infectious viral load in unvaccinated and vaccinated individuals infected with ancestral, Delta or Omicron SARS-CoV-2. *Nat Med.* 2022;28: 1491–1500. doi:10.1038/s41591-022-01816-0
31. Singanayagam A, Hakki S, Dunning J, Madon KJ, Crone MA, Koycheva A, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *Lancet Infect Dis.* 2022;22: 183–195. doi:10.1016/S1473-3099(21)00648-4
32. Kissler SM, Fauver JR, Mack C, Tai CG, Breban MI, Watkins AE, et al. Viral Dynamics of SARS-CoV-2 Variants in Vaccinated and Unvaccinated Persons. *N Engl J Med.* 2021;385: 2489–2491. doi:10.1056/NEJMc2102507
33. Gonçalves A, Bertrand J, Ke R, Comets E, De Lamballerie X, Malvy D, et al. Timing of Antiviral Treatment Initiation is Critical to Reduce SARS-CoV-2 Viral Load. *CPT Pharmacomet Syst Pharmacol.* 2020;9: 509–514. doi:10.1002/psp4.12543

34. Sunagawa J, Komorizono R, Park H, Hart WS, Thompson RN, Makino A, et al. Contact-number-driven virus evolution: A multi-level modeling framework for the evolution of acute or persistent RNA virus infection. Wertheim JO, editor. *PLOS Comput Biol*. 2023;19: e1011173. doi:10.1371/journal.pcbi.1011173
35. Garrett N, Tapley A, Andriesen J, Seocharan I, Fisher LH, Bunts L, et al. High Rate of Asymptomatic Carriage Associated with Variant Strain Omicron. *Infectious Diseases (except HIV/AIDS)*; 2021. doi:10.1101/2021.12.20.21268130
36. Wu Y, Kang L, Guo Z, Liu J, Liu M, Liang W. Incubation Period of COVID-19 Caused by Unique SARS-CoV-2 Strains: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 2022;5: e2228008. doi:10.1001/jamanetworkopen.2022.28008
37. Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect Dis*. 2021;21: 793–802. doi:10.1016/S1473-3099(21)00143-2
38. Sonabend R, Whittles LK, Imai N, Perez-Guzman PN, Knock ES, Rawson T, et al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *The Lancet*. 2021;398: 1825–1835. doi:10.1016/S0140-6736(21)02276-5
39. Levine-Tiefenbrun M, Yelin I, Katz R, Herzel E, Golan Z, Schreiber L, et al. Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. *Nat Med*. 2021;27: 790–792. doi:10.1038/s41591-021-01316-7
40. Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *The Lancet*. 2021;398: 2126–2128. doi:10.1016/S0140-6736(21)02758-6
41. Marc A, Marlin R, Donati F, Prague M, Kerioui M, Hérate C, et al. Impact of variants of concern on SARS-CoV-2 viral dynamics in non-human primates. *Microbiology*; 2022. doi:10.1101/2022.11.09.515748

42. The COVID-19 Genomics UK (COG-UK) consortium, Volz E, Mishra S, Chand M, Barrett JC, Johnson R, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*. 2021;593: 266–269. doi:10.1038/s41586-021-03470-x
43. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. 2020;584: 257–261. doi:10.1038/s41586-020-2405-7
44. Oran DP, Topol EJ. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic: A Systematic Review. *Ann Intern Med*. 2021;174: 655–662. doi:10.7326/M20-6976
45. Sah P, Fitzpatrick MC, Zimmer CF, Abdollahi E, Juden-Kelly L, Moghadas SM, et al. Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. *Proc Natl Acad Sci*. 2021;118: e2109229118. doi:10.1073/pnas.2109229118
46. Ejima K, Kim KS, Ludema C, Bento AI, Iwanami S, Fujita Y, et al. Estimation of the incubation period of COVID-19 using viral load data. *Epidemics*. 2021;35: 100454. doi:10.1016/j.epidem.2021.100454
47. Grant R, Charmet T, Schaeffer L, Galmiche S, Madec Y, Von Platen C, et al. Impact of SARS-CoV-2 Delta variant on incubation, transmission settings and vaccine effectiveness: Results from a nationwide case-control study in France. *Lancet Reg Health - Eur*. 2022;13: 100278. doi:10.1016/j.lanepe.2021.100278
48. Rochman N, Wolf Y, V. Koonin E. Evolution of human respiratory virus epidemics. *F1000Research*. 2021;10: 447. doi:10.12688/f1000research.53392.2
49. Killingley B, Mann AJ, Kalinova M, Boyers A, Goonawardane N, Zhou J, et al. Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. *Nat Med*. 2022;28: 1031–1041. doi:10.1038/s41591-022-01780-9

50. Acevedo MA, Dillemath FP, Flick AJ, Faldyn MJ, Elder BD. Virulence-driven trade-offs in disease transmission: A meta-analysis*. *Evolution*. 2019;73: 636–647. doi:10.1111/evo.13692
51. Jalali N, Brustad HK, Frigessi A, MacDonald EA, Meijerink H, Feruglio SL, et al. Increased household transmission and immune escape of the SARS-CoV-2 Omicron compared to Delta variants. *Nat Commun*. 2022;13: 5706. doi:10.1038/s41467-022-33233-9
52. Sasaki A, Lion S, Boots M. Antigenic escape selects for the evolution of higher pathogen transmission and virulence. *Nat Ecol Evol*. 2022;6: 51–62.
53. Bar-On YM, Flamholz A, Phillips R, Milo R. SARS-CoV-2 (COVID-19) by the numbers. *eLife*. 2020;9: e57309. doi:10.7554/eLife.57309
54. Rochman N, Wolf Y, V. Koonin E. Substantial impact of post-vaccination contacts on cumulative infections during viral epidemics. *F1000Research*. 2021;10: 315. doi:10.12688/f1000research.52341.2
55. Rochman ND, Faure G, Wolf YI, Freddolino L, Zhang F, Koonin EV. Epistasis at the SARS-CoV-2 Receptor-Binding Domain Interface and the Propitiously Boring Implications for Vaccine Escape. Diamond MS, editor. *mBio*. 2022;13: e00135-22. doi:10.1128/mbio.00135-22
56. Zhang X, Ruan Z, Zheng M, Zhou J, Boccaletti S, Barzel B. Epidemic spreading under mutually independent intra- and inter-host pathogen evolution. *Nat Commun*. 2022;13: 6218. doi:10.1038/s41467-022-34027-9
57. Newbold SC, Ashworth M, Finnoff D, Shogren JF, Thunström L. Physical distancing versus testing with self-isolation for controlling an emerging epidemic. *Sci Rep*. 2023;13: 8185. doi:10.1038/s41598-023-35083-x
58. Ito Y, Remion A, Tauzin A, Ejima K, Nakaoka S, Iwasa Y, et al. Number of infection events per cell during HIV-1 cell-free infection. *Sci*

Rep. 2017;7: 6559. doi:10.1038/s41598-017-03954-9

59. Marks M, Millat-Martinez P, Ouchi D, Roberts CH, Alemany A, Corbacho-Monné M, et al. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *Lancet Infect Dis.* 2021;21: 629–636. doi:10.1016/S1473-3099(20)30985-3

60. Zuin M, Gentili V, Cervellati C, Rizzo R, Zuliani G. Viral Load Difference between Symptomatic and Asymptomatic COVID-19 Patients: Systematic Review and Meta-Analysis. *Infect Dis Rep.* 2021;13: 645–653. doi:10.3390/idr13030061

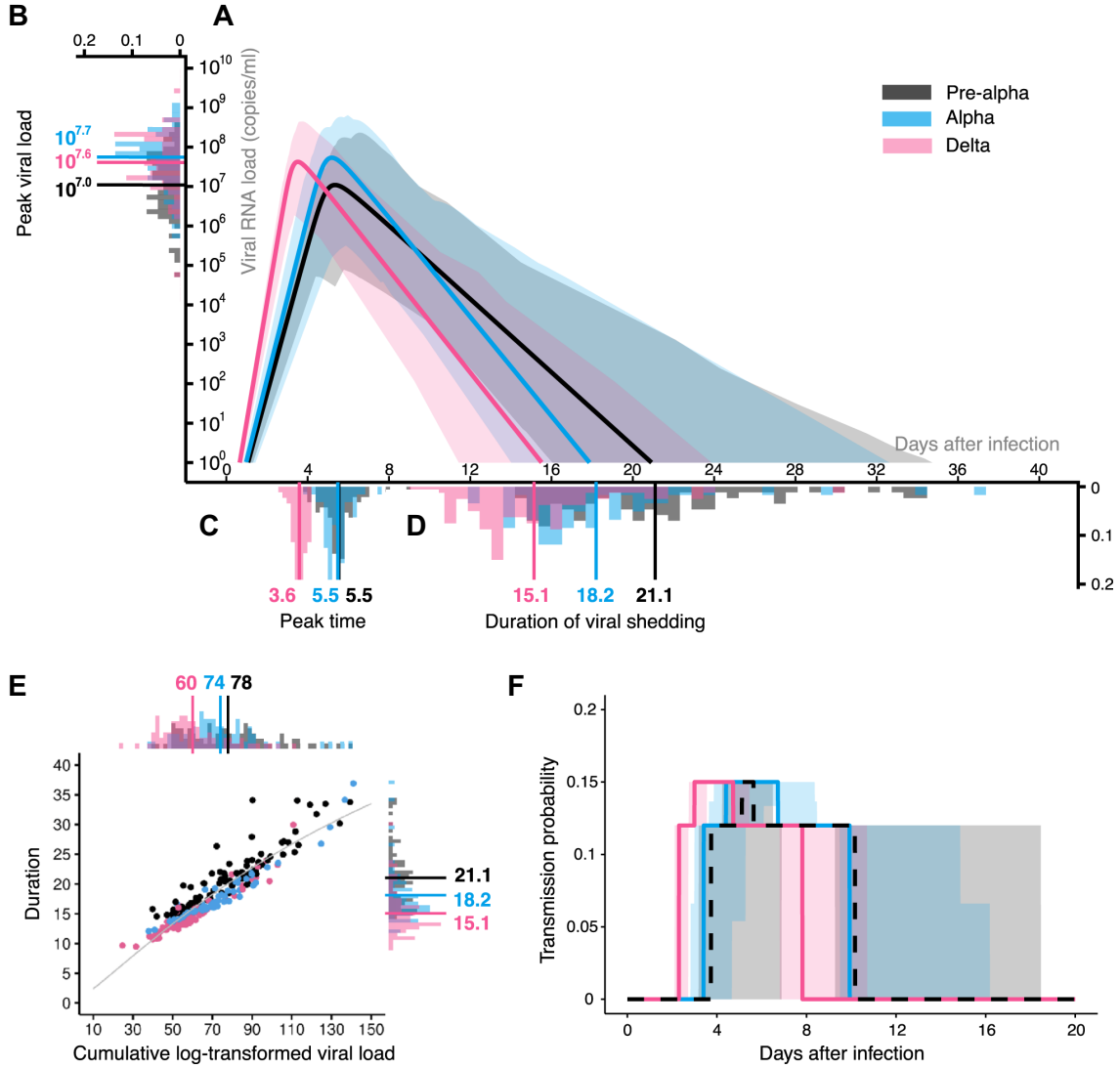
61. Lee S, Kim T, Lee E, Lee C, Kim H, Rhee H, et al. Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *JAMA Intern Med.* 2020;180: 1447. doi:10.1001/jamainternmed.2020.3862

62. JH Holland. *Adaptation in Natural and Artificial Systems.* 1992.

63. Sunagawa, J. & Park, H. Isolation may select for earlier and higher peak viral load but shorter duration in SARS-CoV-2 evolution. Available: <https://zenodo.org/records/10031081>

Figures

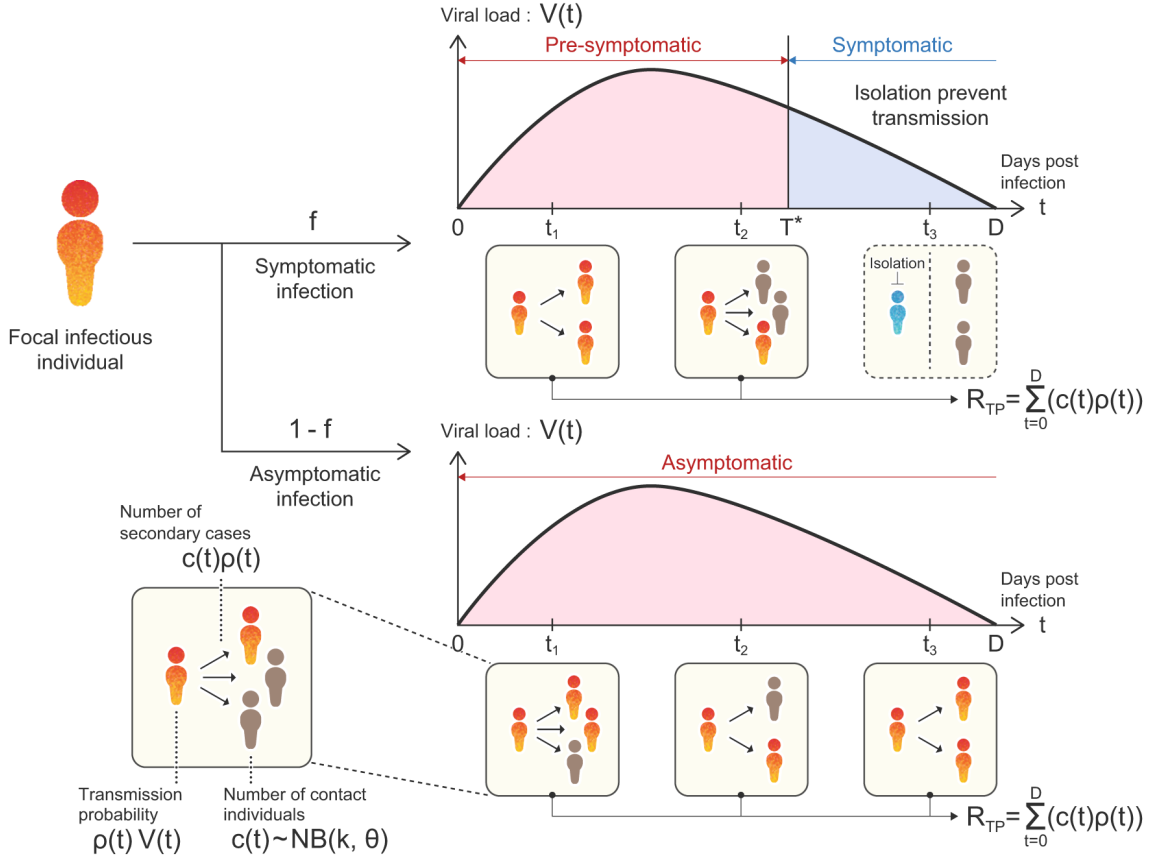
Figure 1: Quantification of evolving SARS-CoV-2 infection dynamics.



Viral load data from upper respiratory specimens (i.e., nose and pharynx) of 86 patients infected with the pre-Alpha (non-VOI/VOCs) variant, 59 Alpha variant patients, and 80 Delta variant patients were used for parameter estimation of the viral dynamics model. (A) The inferred viral dynamics for pre-Alpha (black), Alpha (blue), and Delta (red) variants are plotted. The solid curves correspond to the solution of Eqs. (4, 5) using the best-fit population parameters, and the error bands represented as shadow regions indicate the 95% interquartile range of the inferred viral load

interval for all individuals at each time point. The distribution of the peak viral load, peak time, and duration of viral shedding for each patient are also shown in (B), (C), and (D), respectively. Solid bars are their mean values. (E) The relation between the duration of viral shedding and the cumulative log-transformed viral load is shown. The gray curve is the best-fitting curve for Eq. (7). The distributions illustrated alongside the top and right parts are for the cumulative viral load and the duration of viral shedding, respectively, and the bars are their mean values. (F) The time-dependent transmission probability of each variant is calculated. In all panels, pre-Alpha, Alpha, and Delta variants are colored in black, blue, and red, respectively.

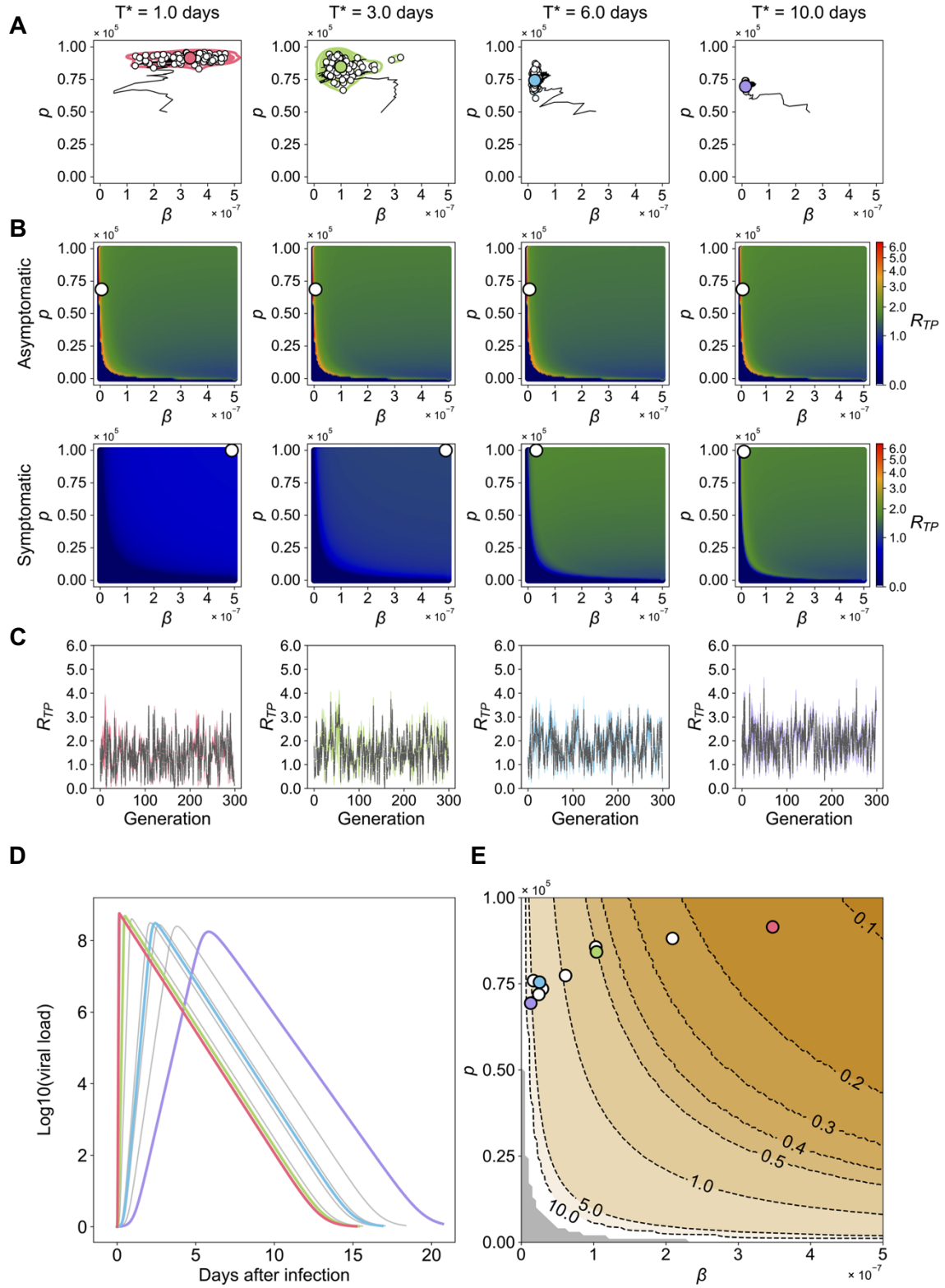
Figure 2: Schematic of multi-level disease transmission.



A schematic of the multi-level population dynamics model is depicted. Virus transmission occurs from an infected individual to susceptible individuals depending on the infected individual's infectivity, which depends on their viral load. At each time step (i.e., day), the focal infected individual has contact with multiple susceptible individuals. Here the contact numbers are assumed to follow a negative binomial distribution, which does not depend on the viral load. The shape parameters k and θ for this binomial distribution are presented in Table S3. The sum of newly infected individuals (i.e., $[c(t): \text{number of contacted individuals per day (everybody is susceptible)}] \times [\rho(t): \text{transmission probability per contacted individual}]$) during the infectious period is calculated as R_{TP} , called the “transmission potential”. Note that the focal infectious individuals are assigned one of two properties: symptomatic or asymptomatic. Whether being symptomatic or asymptomatic is randomly assigned by a constant probability f or $1-f$, respectively. It is assumed that the transmission

chains from the focal symptomatic infectious individuals are perfectly inhibited by isolation after symptom onset, $T^* < t$. Otherwise, R_{TP} is calculated using the whole duration of viral shedding of the focal asymptomatic infectious individual. Thus, R_{TP} for the symptomatic and asymptomatic individuals are different here.

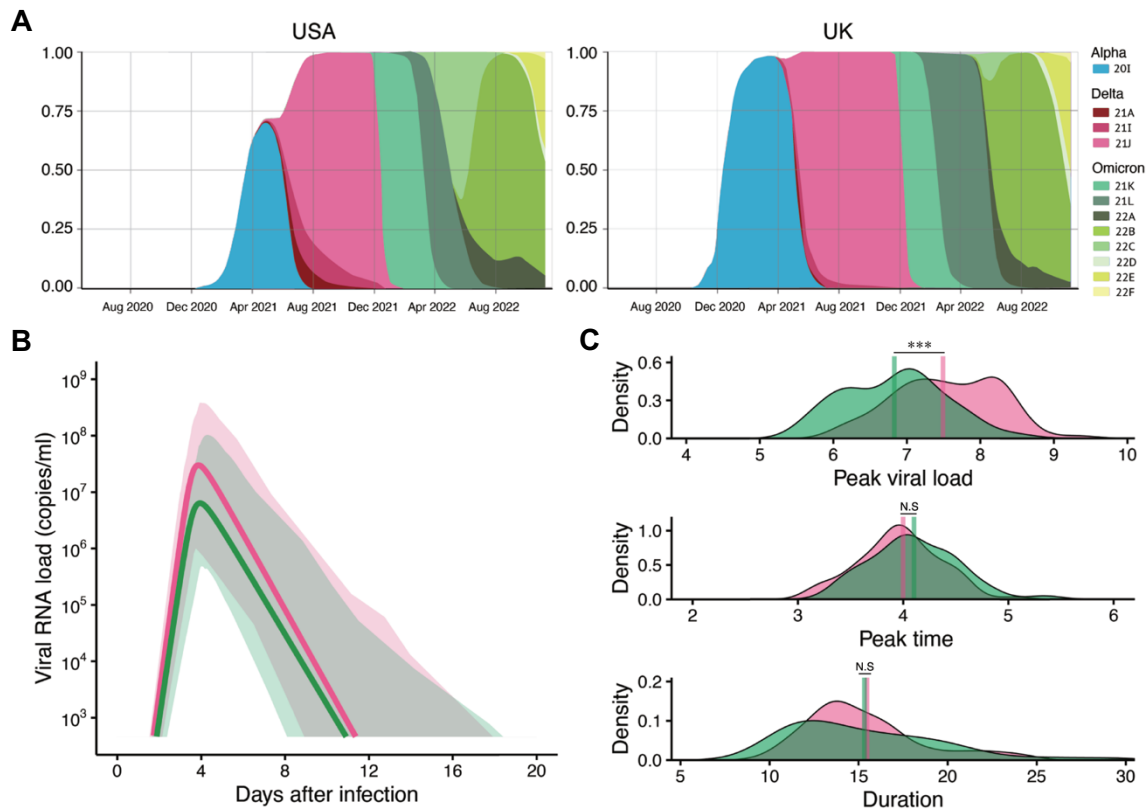
Figure 3: SARS-CoV-2 evolution in silico ($f = 0.3$).



(A) Genetic algorithm (GA) exploring the evolutionary trajectories on the

(β, p) plane until the generation of 300 is applied, depending on different values of the incubation period, T^* . Here β and p are the infection rate and the virus production rate, respectively. All symptomatic individuals lose their transmissibility by isolation after symptom onset ($T^* < t$). The white dots represent the endpoint of 100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in each panel is the mean value of the white dots, which represents the set of evolutionary outcomes of (β, p) under the parameters we used. The colors used here are independent to them in Fig. 1. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness landscapes aggregated solely from the asymptomatic (top row) and symptomatic (bottom row) individuals are described, respectively, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . Note the asymptomatic individuals have larger transmissibility fitness than symptomatic cases on average, regardless of T^* values. (C) The trajectories of R_{TP} along the course of GA with different T^* are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) with different T^* , which were obtained in (A), are shown. (E) The contour-plot for the timing of peak viral load (i.e., peak time) is shown. Each curve is colored accordingly. The gray region is the parameter range satisfying $R_{TP} < 1$.

Figure 4: Quantification of SARS-CoV-2 BA.1 subvariant infection dynamics.



(A) An overview of frequency changes in SARS-CoV-2 variants based on data from CoVariants.org is shown. The left and right panels show the trend in the United States and the United Kingdom, respectively, as examples (because the data we used here were from the USA and UK). See the labels for each color curve corresponding to each VOC. (B) The estimated viral dynamics using our model and publicly available data for the BA.1 subvariant (green) are described along with the Delta variant (red) using the best-fit population parameters. The solid curves correspond to the solution of Eqs. (4, 5) using the best-fit population parameters, and the error bands represented as shadow regions indicate the 95% interquartile range of the inferred viral load interval for all individuals at each time point. (C) The distributions of the peak viral load, peak time, and duration of viral shedding for each patient are shown with their mean values (vertical solid bars). A statistically significant difference between two groups was found only for peak viral load (p-values by the two-sided Mann-Whitney-

Wilcoxon test are 1.3×10^{-6} for peak viral load, 0.09 for peak time, and 0.48 for duration).

Supplementary Tables

Table 1: Summary of viral load properties for SARS-CoV-2 variants

variant	peak viral load	peak time	duration of viral shedding	Cumulative log-transformed viral load(V_{total})
pre-alpha	$10^{7.0a}(10^{4.8} - 10^{8.6})^b$	5.5(3.9 – 9.1)	21.1(14.6 – 34.1)	78.3(39.8 – 139.3)
alpha	$10^{7.7}(10^{5.8} - 10^{9.1})^c$	5.5(3.9 – 10.7)	18.2(12.1 – 36.9) ^c	74.4(37.8 – 141.1)
delta	$10^{7.6}(10^{4.8} - 10^{9.4})^c$	3.6(2.7 – 4.2) ^c	15.1(9.5 – 30.0) ^c	60.4(24.7 – 110.8) ^c

^aMean value.

^bMin-Max value.

^cStatistically different from pre-Alpha variant ($p = 0.01$ by the Mann-Whitney-Wilcoxon test)

Table S1: Summary of the study data

country	number of cases	specimens	detection limit	references
pre-alpha				
USA-1	39	Upper respiratory tract	2.66 $\log_{10}$ (copies/ml)	(1)
UK	47	Upper respiratory tract	1.49 $\log_{10}$ (copies/ml)	(2)
alpha				
USA-2	22	Upper respiratory tract	2.66 $\log_{10}$ (copies/ml)	(3)
UK	37	Upper respiratory tract	1.49 $\log_{10}$ (copies/ml)	(2)
delta				
USA-2	10	Upper respiratory tract	2.66 $\log_{10}$ (copies/ml)	(3)
USA-3	54	Upper respiratory tract	2.66 $\log_{10}$ (copies/ml)	(4)
UK	16	Upper respiratory tract	1.49 $\log_{10}$ (copies/ml)	(2)
omicron				
USA-3	49	Upper respiratory tract	2.66 $\log_{10}$ (copies/ml)	(4)

Table S2: Estimated fixed parameters for SARS-CoV-2 pre-Alpha, Alpha, and Delta variants

parameters	symbol	unit	pre-alpha	alpha	delta
Maximum rate constant for viral replication	γ	day ⁻¹	5.34	6.31 [#]	8.51 [#]
Rate constant for virus infection	β	(copies/mL) ⁻¹ day ⁻¹	2.33 × 10 ⁻⁷	4.74 × 10 ^{-8#}	1.03 × 10 ^{-7#}
Death rate of infected cells	δ	day ⁻¹	1.11	1.57 [#]	1.53 [#]
Death rate of infected cells	T_p	days	7.49	6.49 [#]	4.81 [#]

[#]Statistically different from pre-Alpha variant (the Wald test).

Table S3: Fixed parameters for multi-level population dynamics model

parameters	symbol	value	reference
The number of uninfected target cells at initial time	$T(0)$	133000	(5)
The clearance rate of the virus	c	20	(6)
The fraction of uninfected target cells compared with that at initial time	$f(0)$	1	(6)
The amount of the virus at initial time	$V(0)$	0.01	(5)
The maximum value of δ	δ_{max}	1.57	Estimated from data
The viral production rate satisfying $\delta = \delta_{max}/2$	p_{50}	100	Estimated from data
The maximum duration of viral shedding	D_{max}	80	Estimated from data
The total viral load at which the duration is half of its maximum	V_{50}	198.58	Estimated from data
The steepness at which duration increases with increasing viral load	Dk	1.17	Estimated from data
The shape parameter of gamma distribution	k	1.5	Estimated from data
The scale parameter of gamma distribution	θ	1.5	Estimated from data
The maximum value of β	β_{max}	5.0×10^{-7}	Estimated from data
The maximum value of p	p_{max}	1.0×10^5	Estimated from data

Table S4: Fixed parameters for genetic algorithm (GA)

parameters	value
The population size	100
The number of iterations	300
The number of elitism	2
The probability of selection	1.0
The probability of crossover	0.2
The probability of mutation	0.7
The predefined range of β in mutation	0.5×10^{-7}
The predefined range of p in mutation	0.1×10^5

Table S5: Estimated fixed parameters for SARS-CoV-2 Delta and Omicron variants

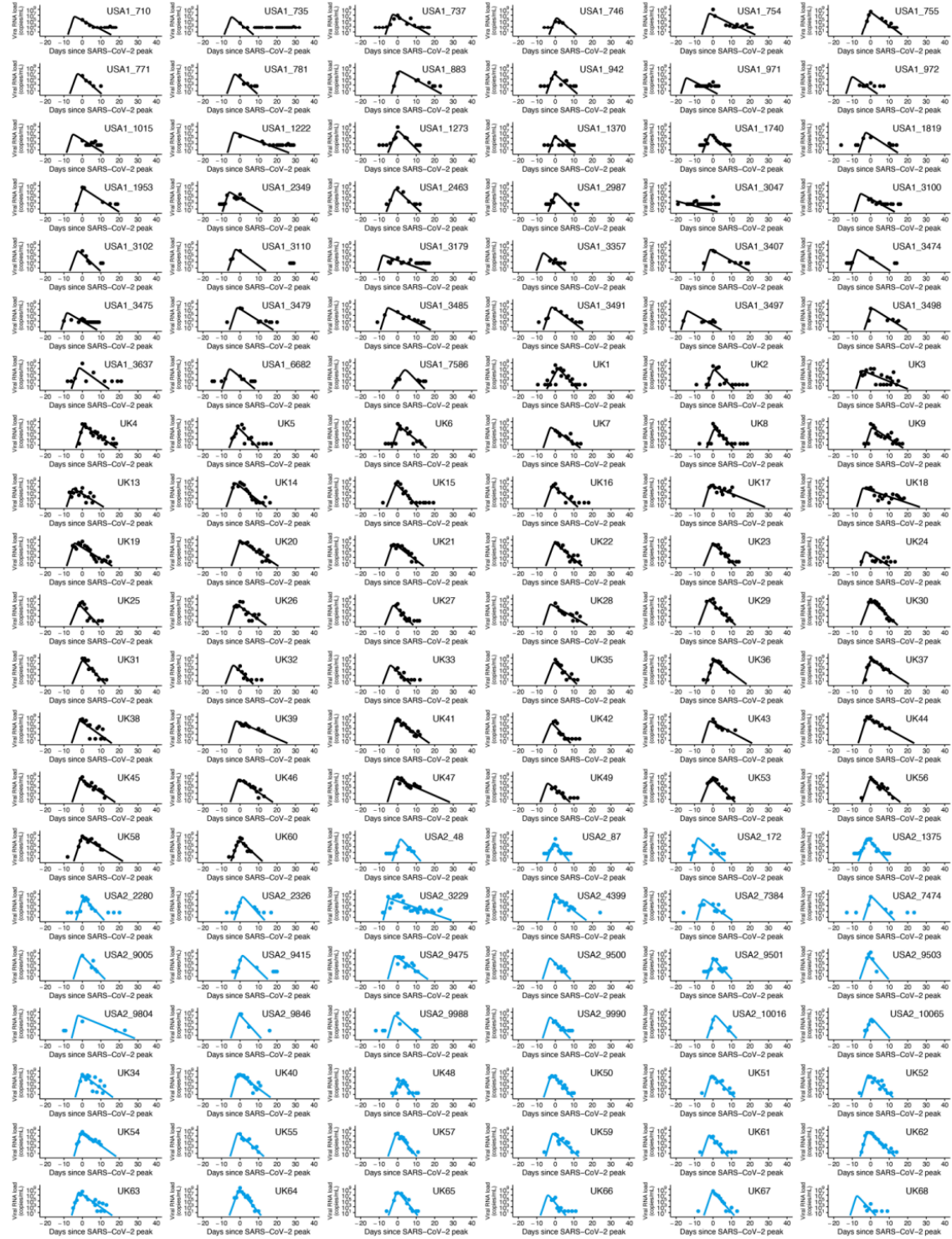
parameters	symbol	unit	delta	omicron
Maximum rate constant for viral replication	γ	day ⁻¹	7.83	7.21 [#]
Rate constant for virus infection	β	(copies/mL) ⁻¹ day ⁻¹	1.23×10^{-7}	5.38×10^{-7} [#]
Death rate of infected cells	δ	day ⁻¹	1.60	1.49
Death rate of infected cells	T_p	days	4.96	5.46

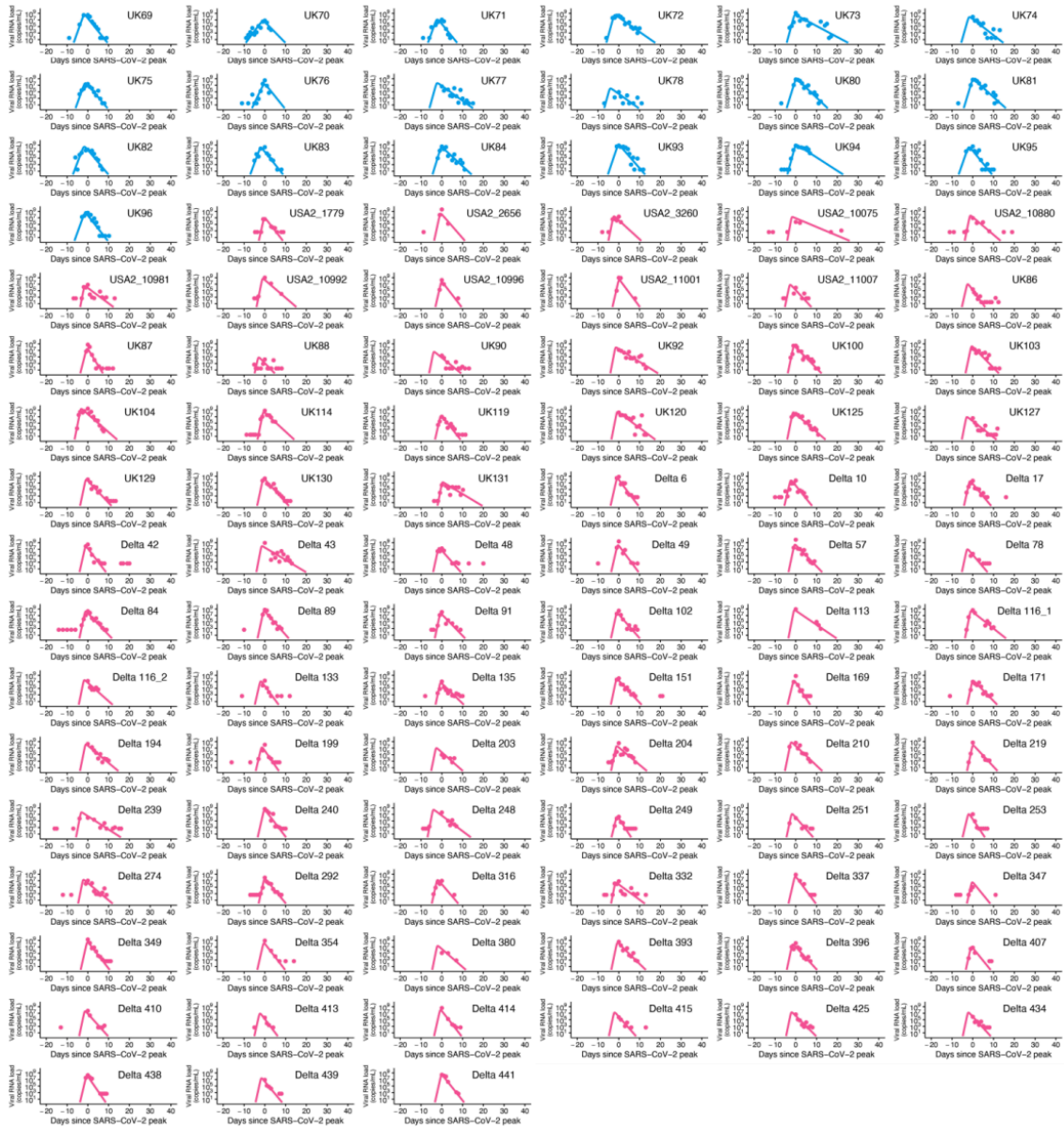
[#]Statistically different from pre-Alpha variant (the Wald test).

Algorithm S1. Procedure of genetic algorithm (GA)

genetic algorithm
Input: s_i , solution i of combinations of β and p
Input: $P(t = i) = \{s_1, s_2, \dots, s_j\}$, population of solutions in iteration i
$t \leftarrow 0$;
Initialize ($P(t = 0)$) with random β and p ;
Evaluate ($P(t = 0)$);
while not termination do
The scale parameter of gamma distribution
$P(t)_e \leftarrow \text{Elitism}(P(t))$;
$P(t)_p \leftarrow \text{selectSolutions}(P(t))$;
$P(t)_c \leftarrow \text{crossover}(P(t)_e + P(t)_p)$;
Mutate ($P(t)_c$);
Evaluate ($P(t)_c$);
$P(t+1) \leftarrow P(t)_c$;
$t \leftarrow t + 1$;
end while

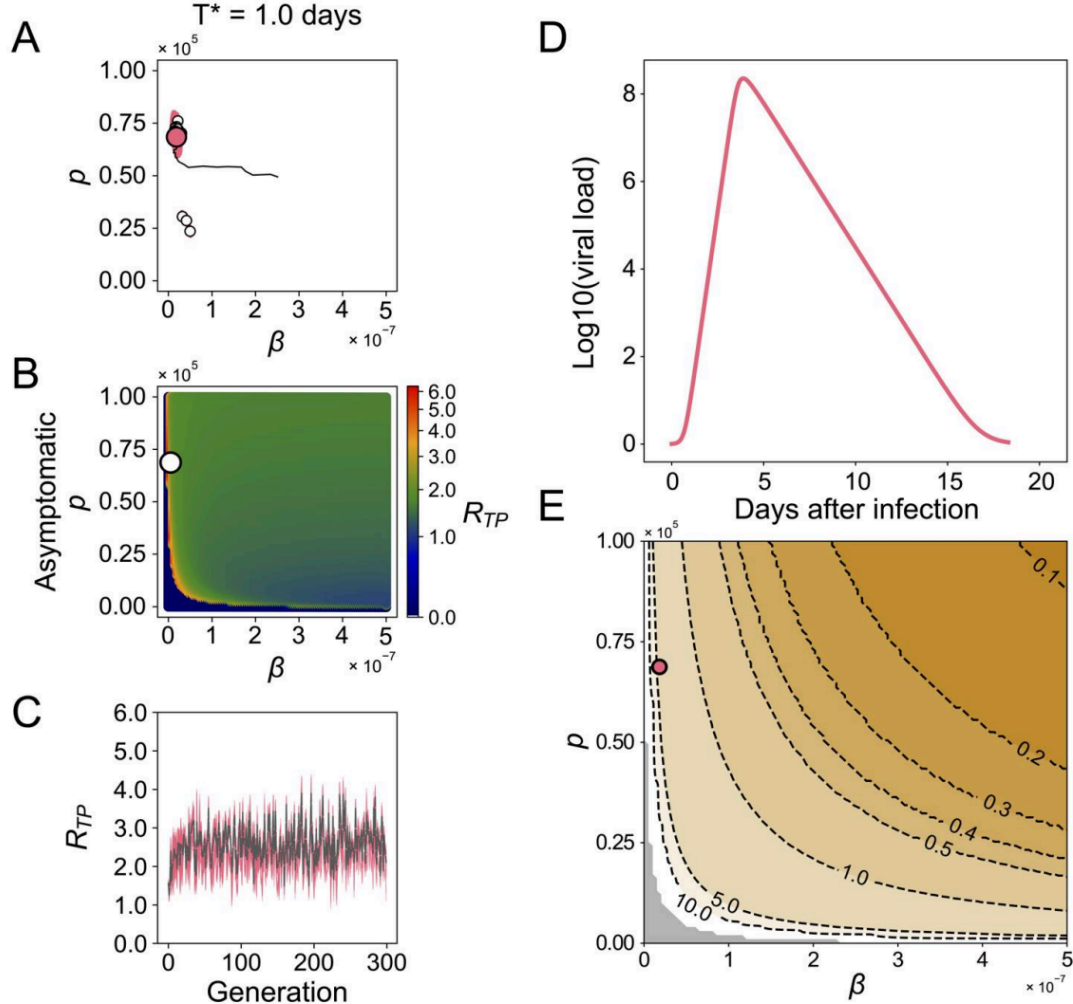
Figure S1: Viral load trajectory for individual patients infected with pre-Alpha, Alpha, and Delta variants.





The estimated viral load for each individual patient (solid lines) along with the observed data (closed dots) are depicted using the best-fit parameter estimates. Pre-Alpha, Alpha, and Delta variants are shown in black, blue, and red, respectively.

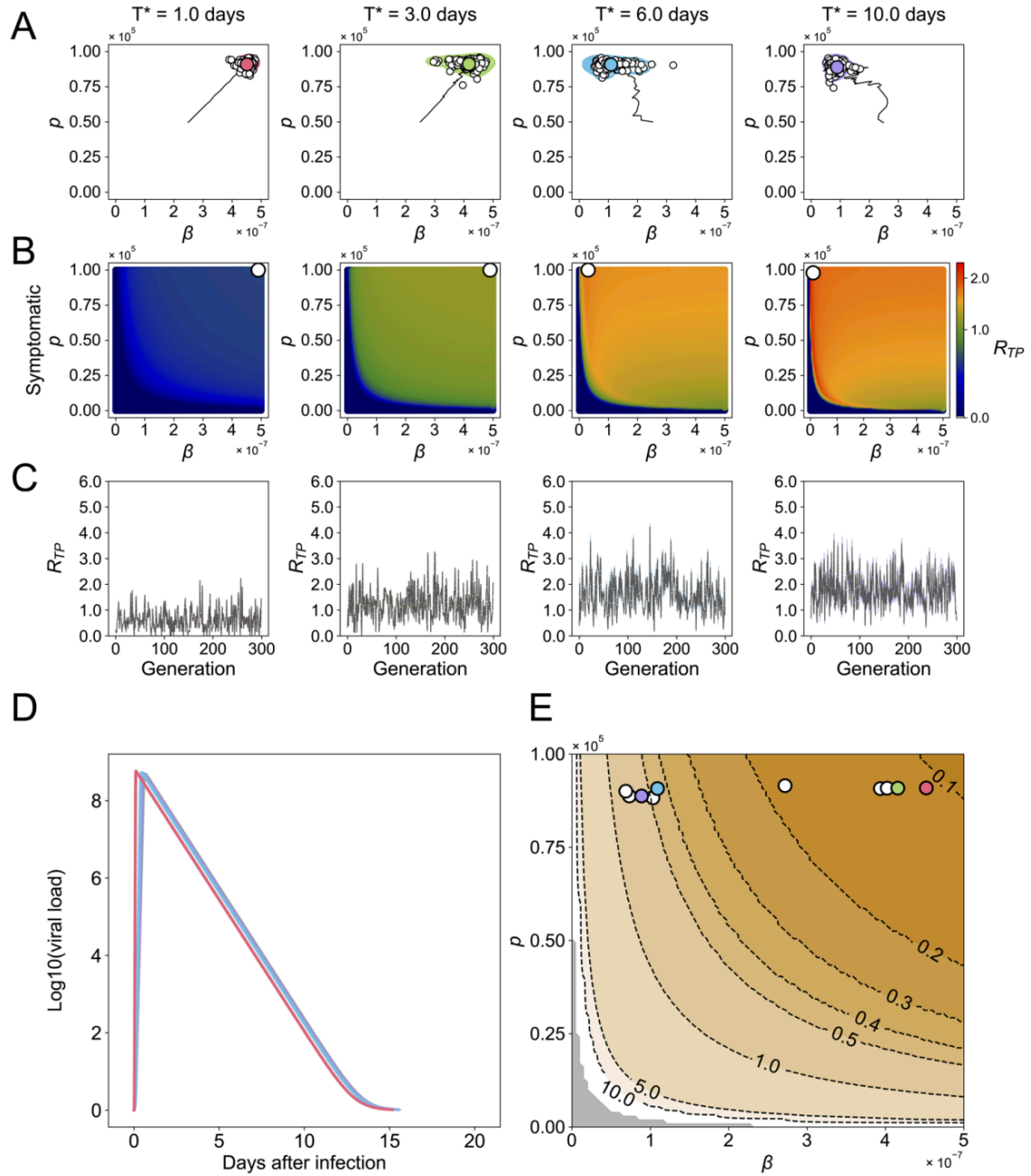
Figure S2: SARS-CoV-2 evolution in silico without symptomatic infection ($f = 0$ and $T^* = 1$).



(A) Genetic algorithm (GA) exploring the evolutionary trajectories on the (β, p) plane until the generation of 300 is applied. Our choice of the incubation period, T^* , does not affect the calculation because no asymptomatic individuals change their transmissibility by isolation (i.e., no individual is applied to isolation). The white dots represent the endpoint of 100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in the panel is the mean value of the white dots, which represent the optimal set of (β, p) under the parameters we used. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness

landscape aggregated from asymptomatic individuals (i.e., no symptomatic individuals due to $f = 0$) is described, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . (C) The trajectories of R_{TP} along the course of GA are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) , which were obtained in (A), are shown. (E) The counter plot for the timing of peak viral load (i.e., peak time) is shown. The gray region is the parameter range satisfying $R_{TP} < 1$.

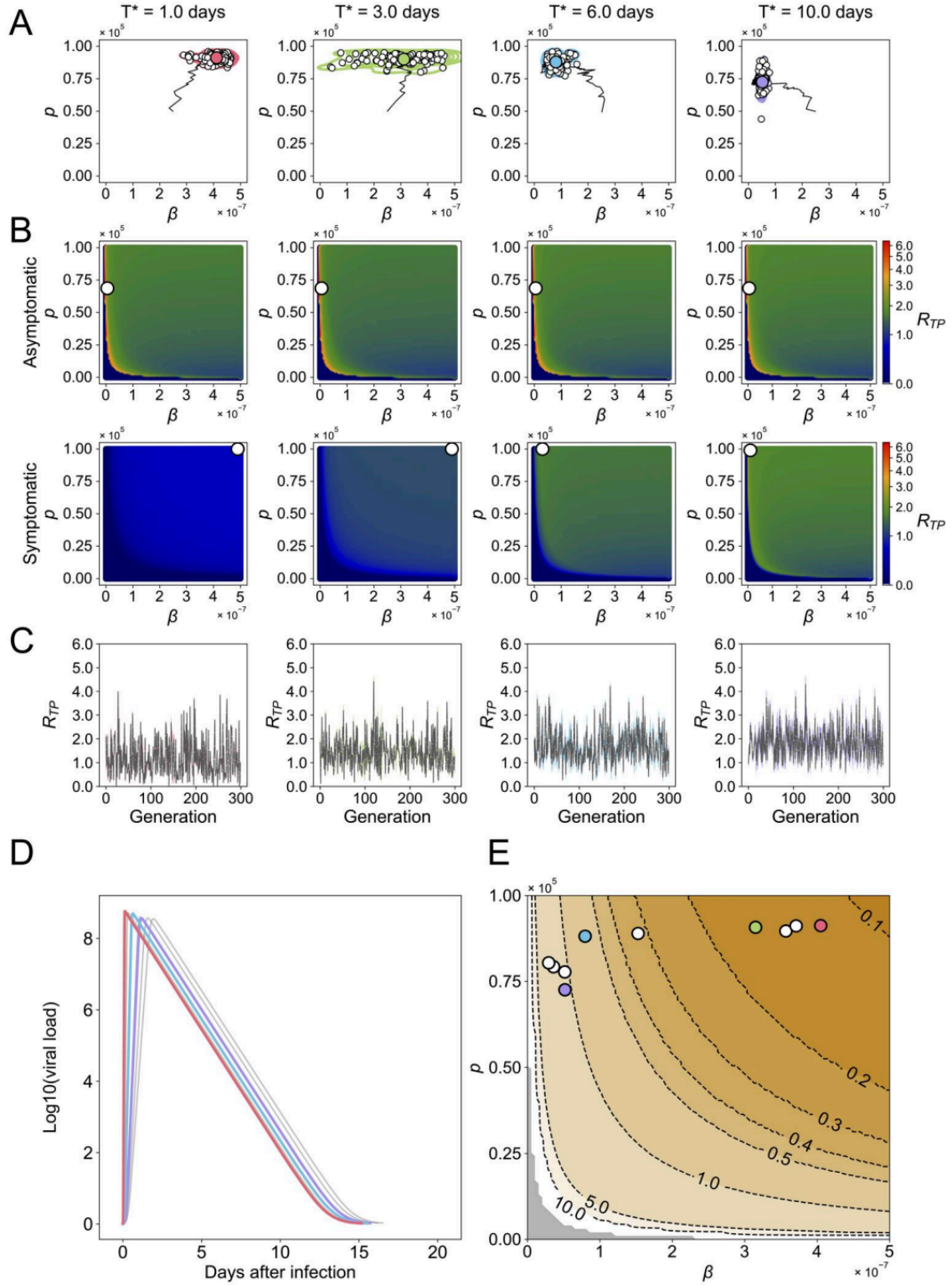
Figure S3: SARS-CoV-2 evolution in silico without asymptomatic infection ($f = 1$).



(A) GA exploring the evolutionary trajectories on the (β, p) plane until the generation of 300 is applied, depending on different values of the incubation period, T^* . All individuals lose their transmissibility by isolation after the symptom onset ($T^* < t$). The white dots represent the endpoint of

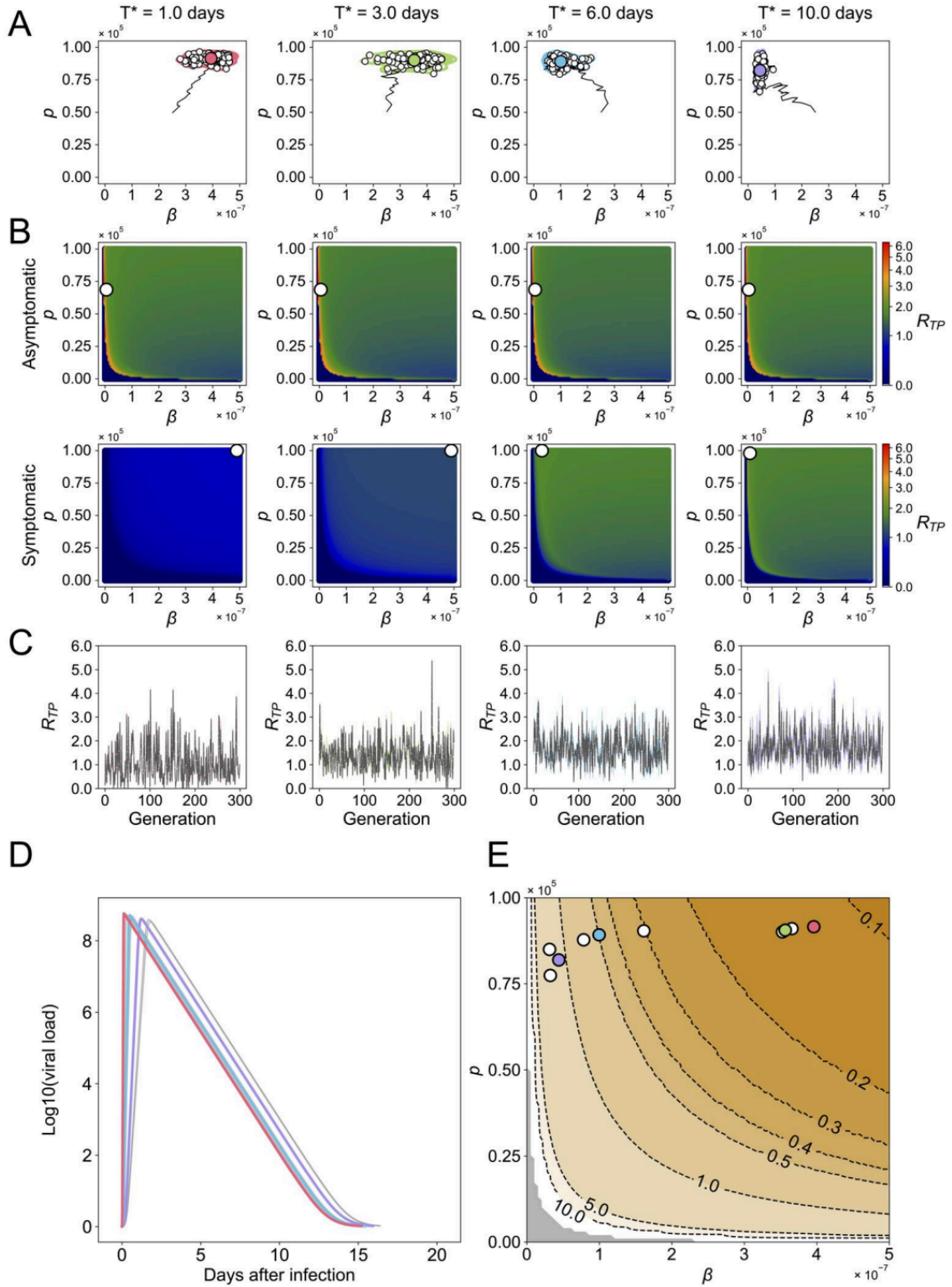
100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in each panel is the mean value of the white dots, which represents the optimal set of (β, p) under the parameters we used. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness landscapes aggregated from symptomatic individuals (i.e., no asymptomatic individuals due to $f = 0$) are described, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . (C) The trajectories of R_{TP} along the course of GA with different T^* are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) with different T^* , which were obtained in (A), are shown. Each curve is colored accordingly. (E) The counter plot for the timing of peak viral load is shown. Each dot is colored accordingly. The gray region is the parameter range satisfying $R_{TP} < 1$

Figure S4-1: Sensitivity analysis of the proportion of symptomatic infection ($f = 0.5$).



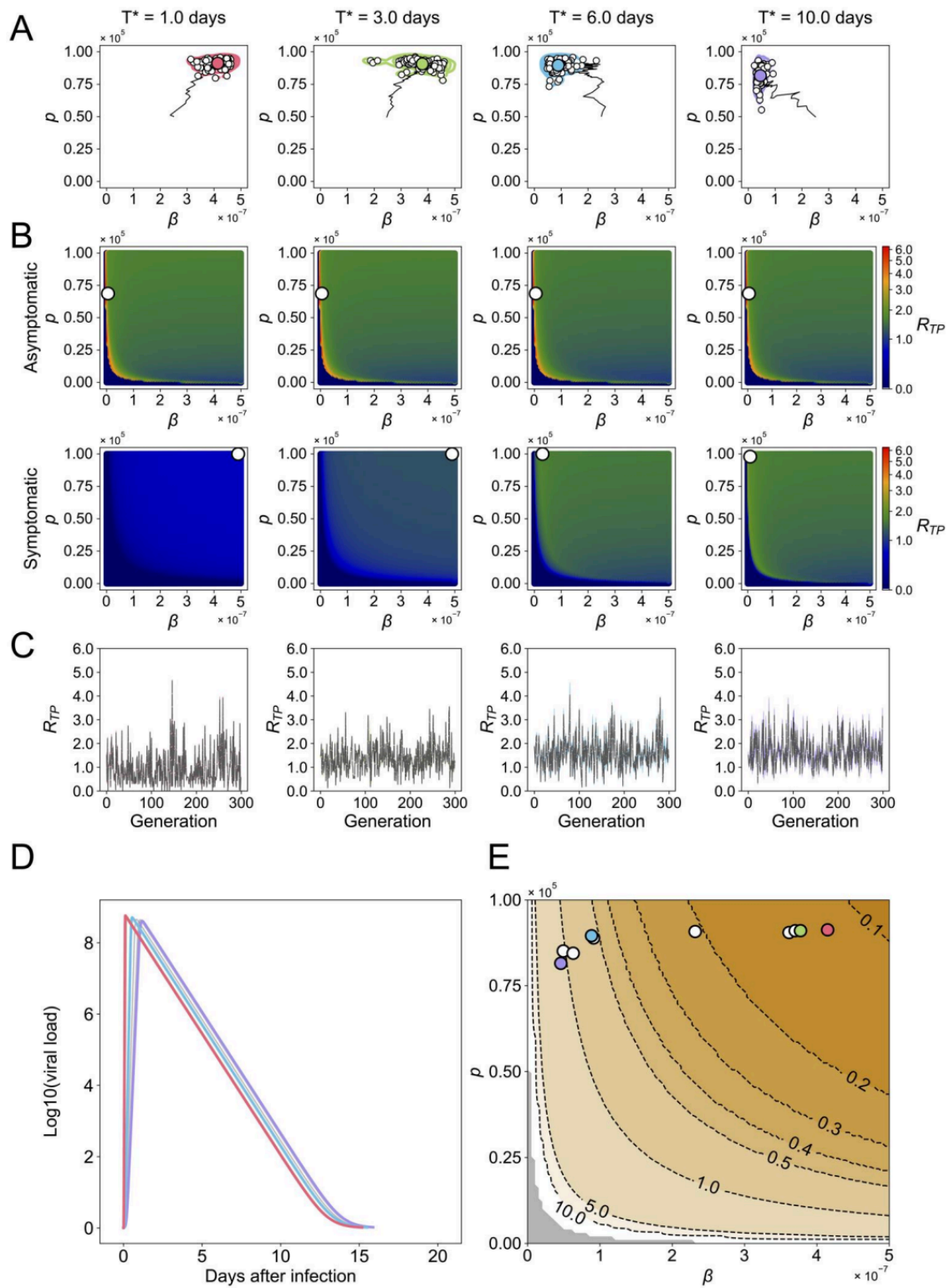
(A) Genetic algorithm (GA) exploring the evolutionary trajectories on the (β, p) plane until the generation of 300 is applied, depending on different values of the incubation period, T^* . All individuals lose their transmissibility by NPIs after symptom onset ($T^* < t$). The white dots represent the endpoint of 100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in each panel is the mean value of the white dots, which represent the optimal set of (β, p) under the parameters we used. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness landscape aggregated solely from the asymptomatic (top row) and symptomatic (bottom row) individuals, respectively, are described, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . (C) The trajectories of R_{TP} along the course of GA with different T^* are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) with different T^* , which were obtained in (A), are shown. Each curve is colored accordingly. (E) The counter plot for the timing of peak viral load is shown. Each dot is colored accordingly. The gray region is the parameter range satisfying $R_{TP} < 1$.

Figure S4-2: Sensitivity analysis of the proportion of symptomatic infection ($f = 0.6$).



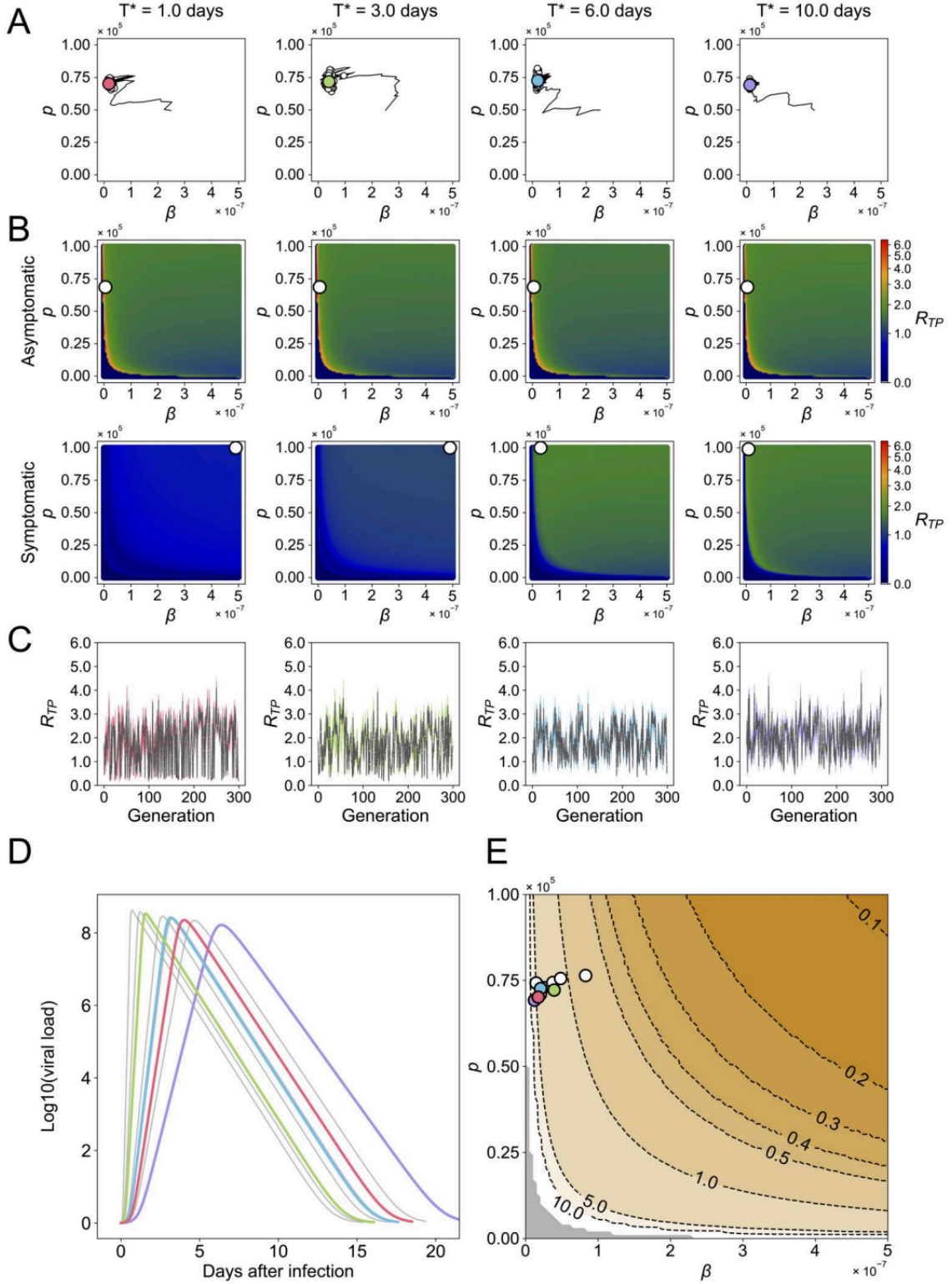
Same explanation for (A-E) is applied as in Figure S4-1.

Figure S4-3: Sensitivity analysis of the proportion of symptomatic infection ($f = 0.7$).



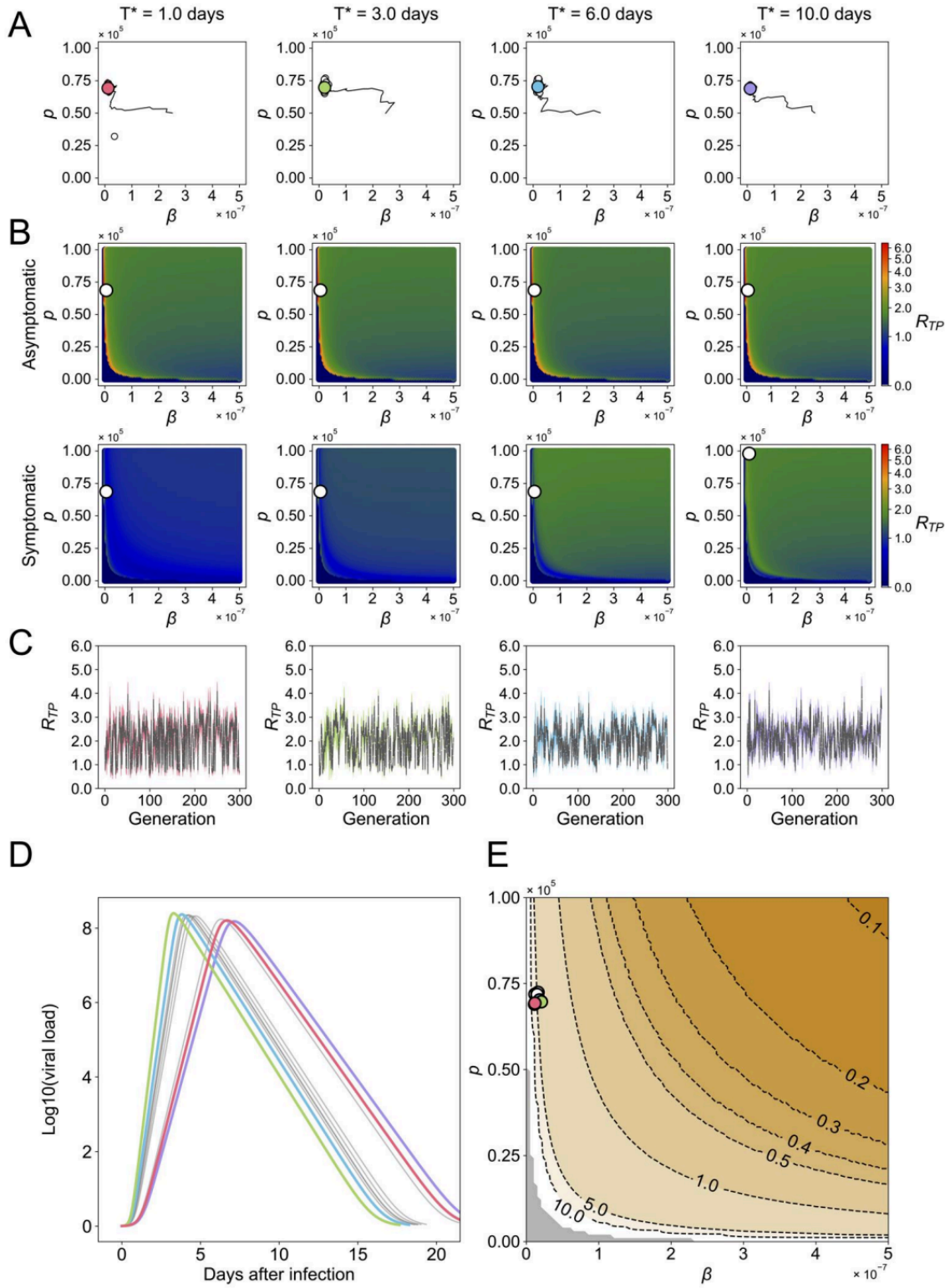
Same explanation for (A-E) is applied as in Figure S4-1.

Figure S5-1: Sensitivity analysis of the impact of Incomplete Isolation on Viral Load Dynamics ($f = 0.3$ and $\xi = 0.1$).



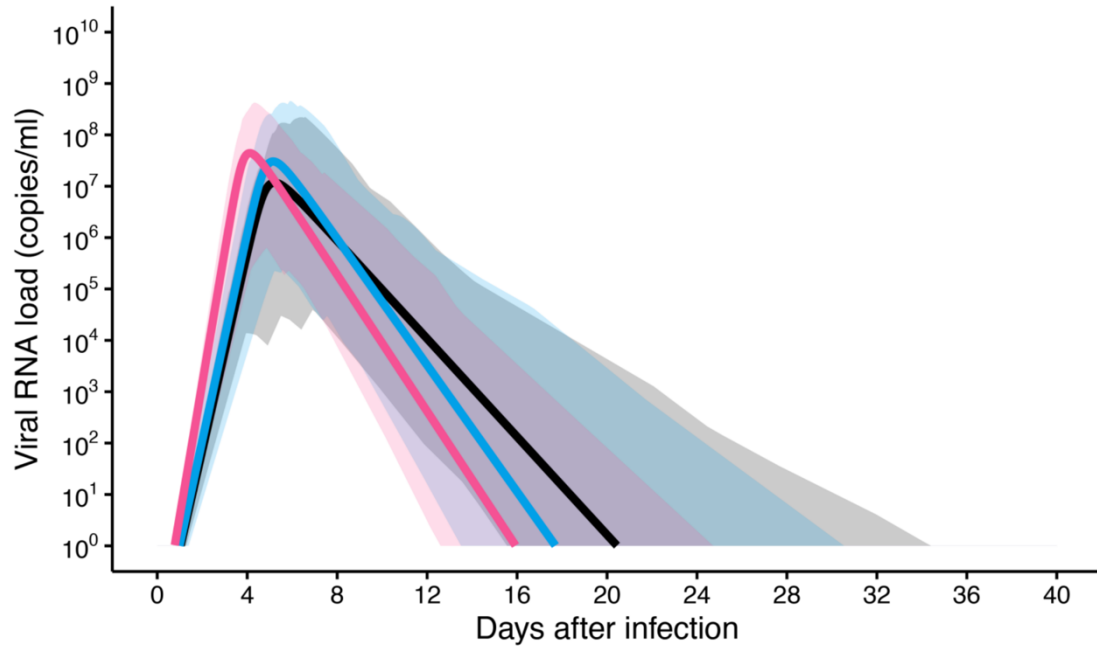
(A) Genetic algorithm (GA) exploring the evolutionary trajectories on the (β, p) plane until the generation of 300 is applied, depending on different values of the incubation period, T^* . All individuals lose their transmissibility by NPIs after symptom onset ($T^* < t$). The white dots represent the endpoint of 100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in each panel is the mean value of the white dots, which represent the optimal set of (β, p) under the parameters we used. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness landscape aggregated solely from the asymptomatic (top row) and symptomatic (bottom row) individuals, respectively, are described, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . (C) The trajectories of R_{TP} along the course of GA with different T^* are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) with different T^* , which were obtained in (A), are shown. Each curve is colored accordingly. (E) The counter plot for the timing of peak viral load is shown. Each dot is colored accordingly. The gray region is the parameter range satisfying $R_{TP} < 1$.

Figure S5-2: Sensitivity analysis of the impact of Incomplete Isolation on Viral Load Dynamics ($f = 0.3$ and $\xi = 0.3$).



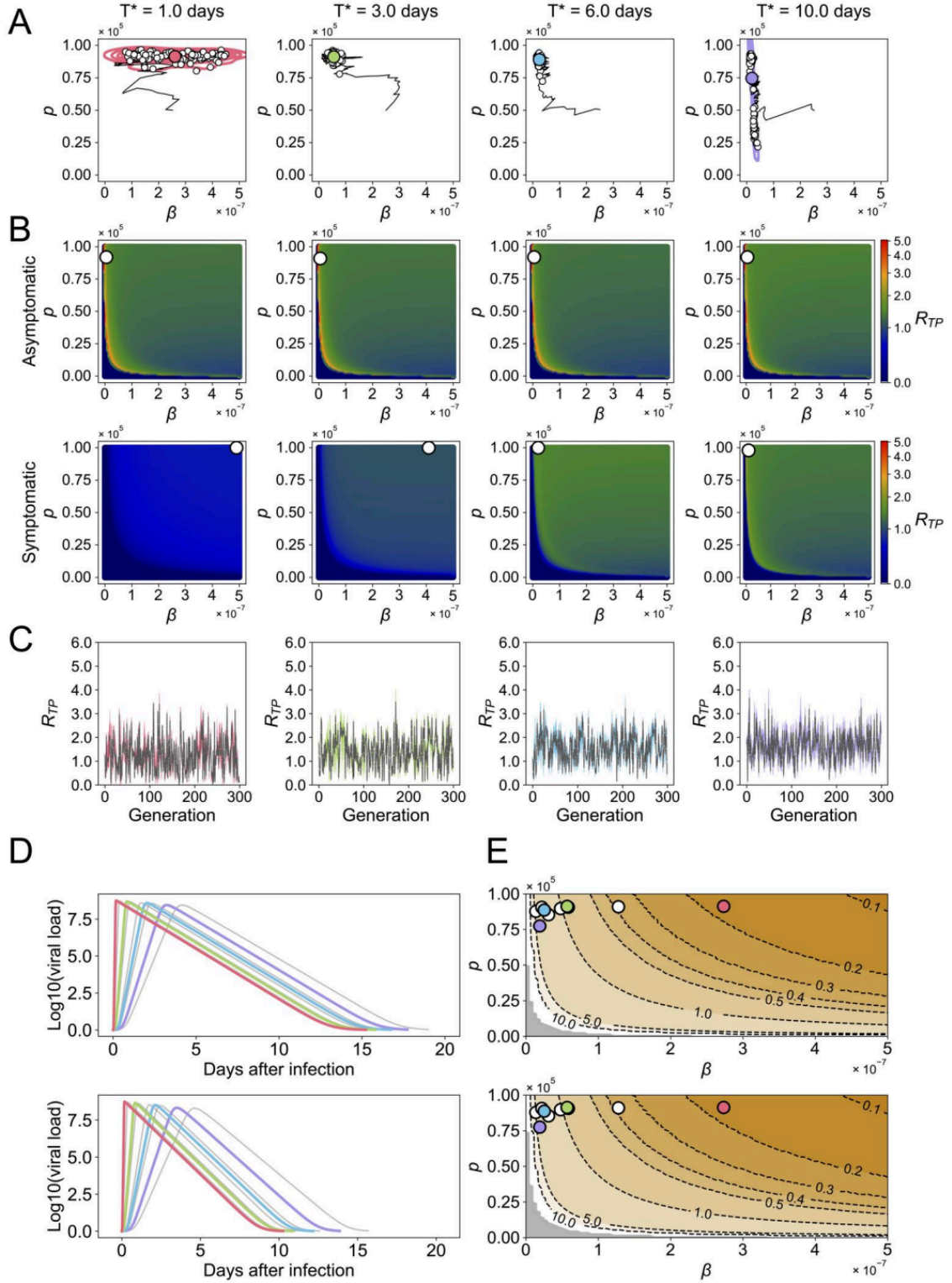
Same explanation for (A-E) is applied as in Figure S5-1.

Figure S6: Quantification of SARS-CoV-2 infection dynamics with patients who have prior immunity.



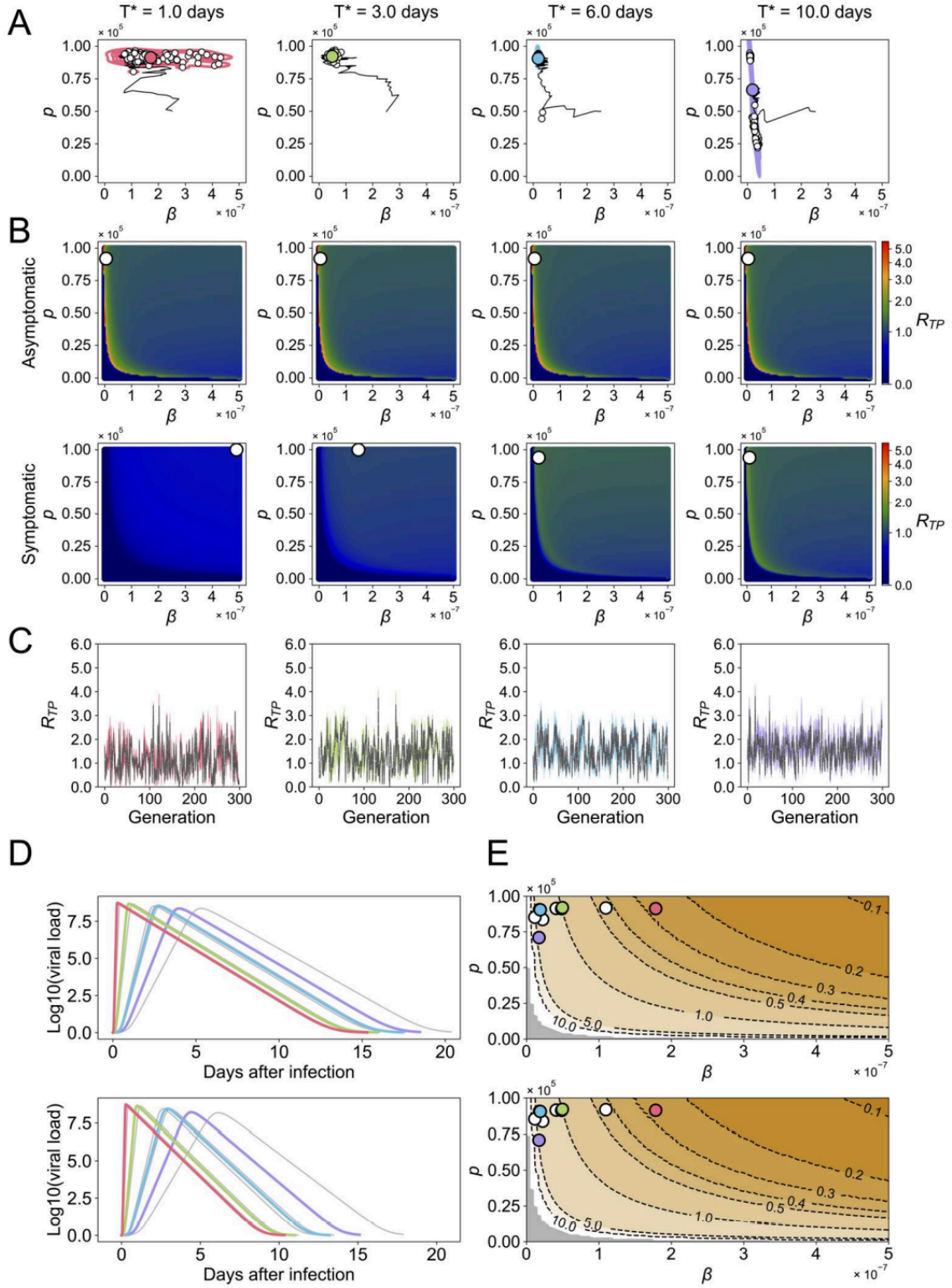
The inferred viral dynamics of 293 COVID-19 patients with and without prior immunity. It contains 225 patients without prior immunity (86 with pre-Alpha variant, 59 with Alpha variant and 80 with Delta variant), and 68 patients with prior immunity by vaccination or pre-infection (14 with Alpha variant and 54 with Delta variant). The solid curves correspond to the solution of Eqs.(4-5) using the best-fit population parameters, and the shadow regions are the 95% interquartile range of the predictive interval using the estimated parameters for each patient. The black, blue, and red colors indicate the pre-Alpha, Alpha, and Delta variants, respectively.

Figure S7-1: Sensitivity analysis of effects of vaccination rates on optimal viral parameters ($f = 0.3$ and vaccination rate = 0.5).



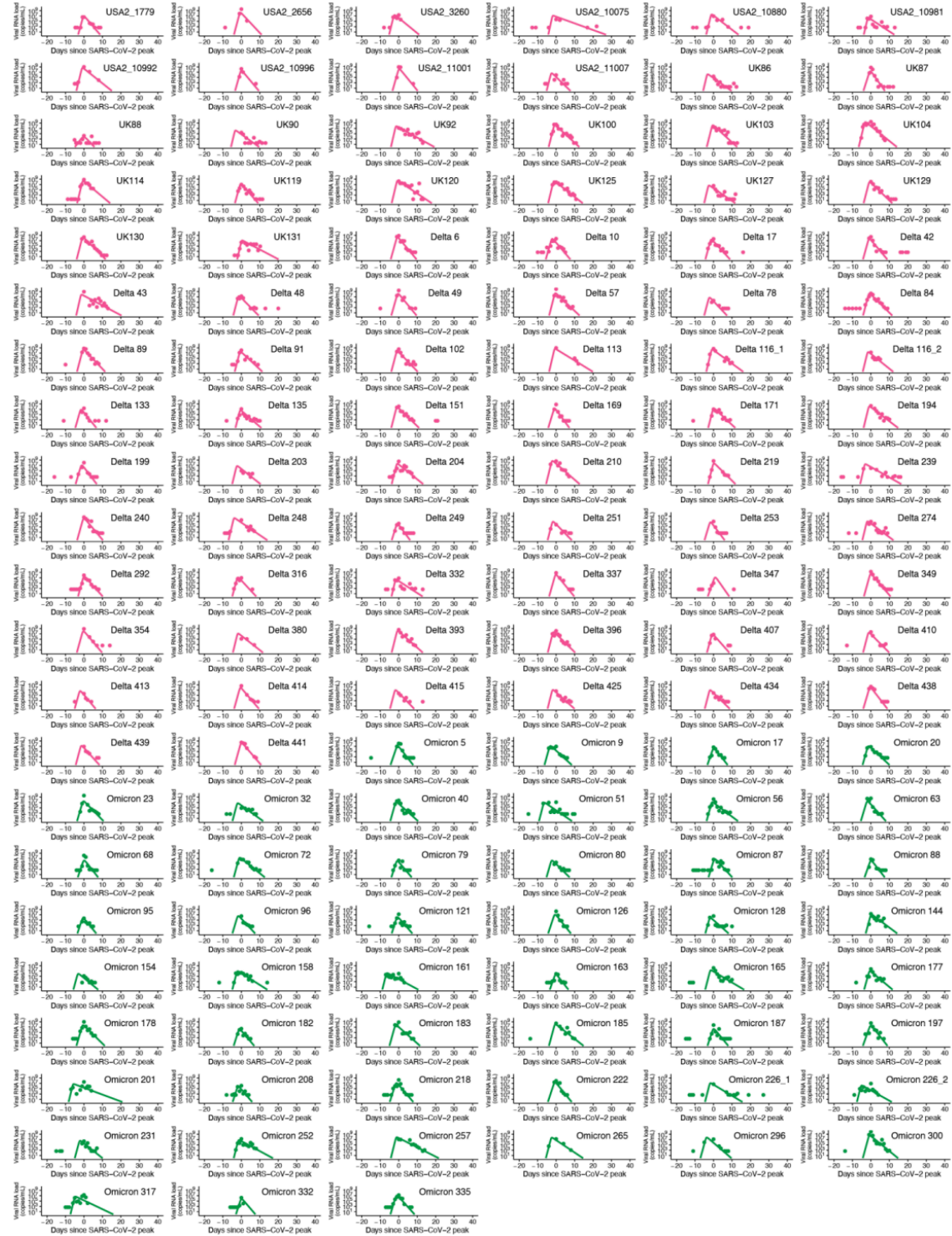
(A) Genetic algorithm (GA) exploring the evolutionary trajectories on the (β, p) plane until the generation of 300 is applied, depending on different values of the incubation period, T^* . All individuals lose their transmissibility by NPIs after symptom onset ($T^* < t$). The white dots represent the endpoint of 100 independent simulation runs, and the contour lines are the kernel density estimation of their distribution. The colored dot in each panel is the mean value of the white dots, which represent the optimal set of (β, p) under the parameters we used. The black line is the mean trajectory of the GA through 300 generations. (B) The mean transmissibility fitness landscape aggregated solely from the asymptomatic (top row) and symptomatic (bottom row) individuals, respectively, are described, using 100 runs of GA. The white dot represents the maximum value of the mean transmissibility fitness, R_{TP} . (C) The trajectories of R_{TP} along the course of GA with different T^* are calculated. The gray dotted lines are the mean trajectory over 100 trials of colored lines. (D) The time-series patterns of viral load with the optimal parameters of (β, p) with different T^* , which were obtained in (A), are shown. Each curve is colored accordingly. The upper row represents the case for unvaccinated individuals, while the lower row demonstrates the case for vaccinated individuals (with δ increased by 1.5 times). (E) The counter plot for the timing of peak viral load is shown. Each dot is colored accordingly. The gray region is the parameter range satisfying $R_{TP} < 1$. The upper and lower rows are the same as in (D).

Figure S7-2: Sensitivity analysis of effects of vaccination rates on optimal viral parameters ($f = 0.3$ and vaccination rate = 0.8).



Same explanation for (A-E) is applied as in Figure S6-1.

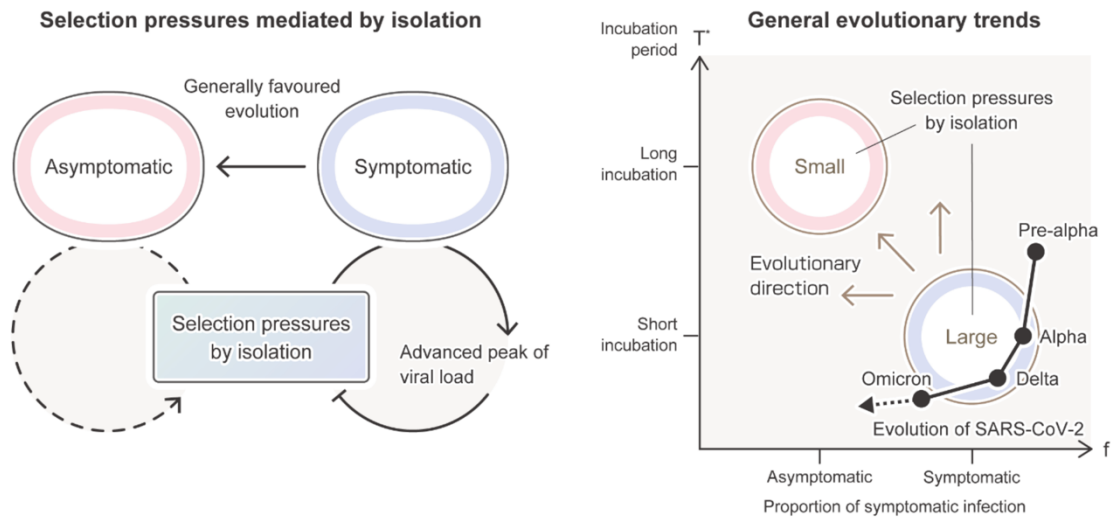
Figure S8: Viral load trajectory for individual patients infected with Delta and Omicron variants.



The estimated viral load for each individual patient (solid lines) along with the observed data (closed dots) are depicted using the best-fit parameter

estimates. Delta and Omicron variants are shown in red and green, respectively. The data for the delta variant are same as in Figure S1.

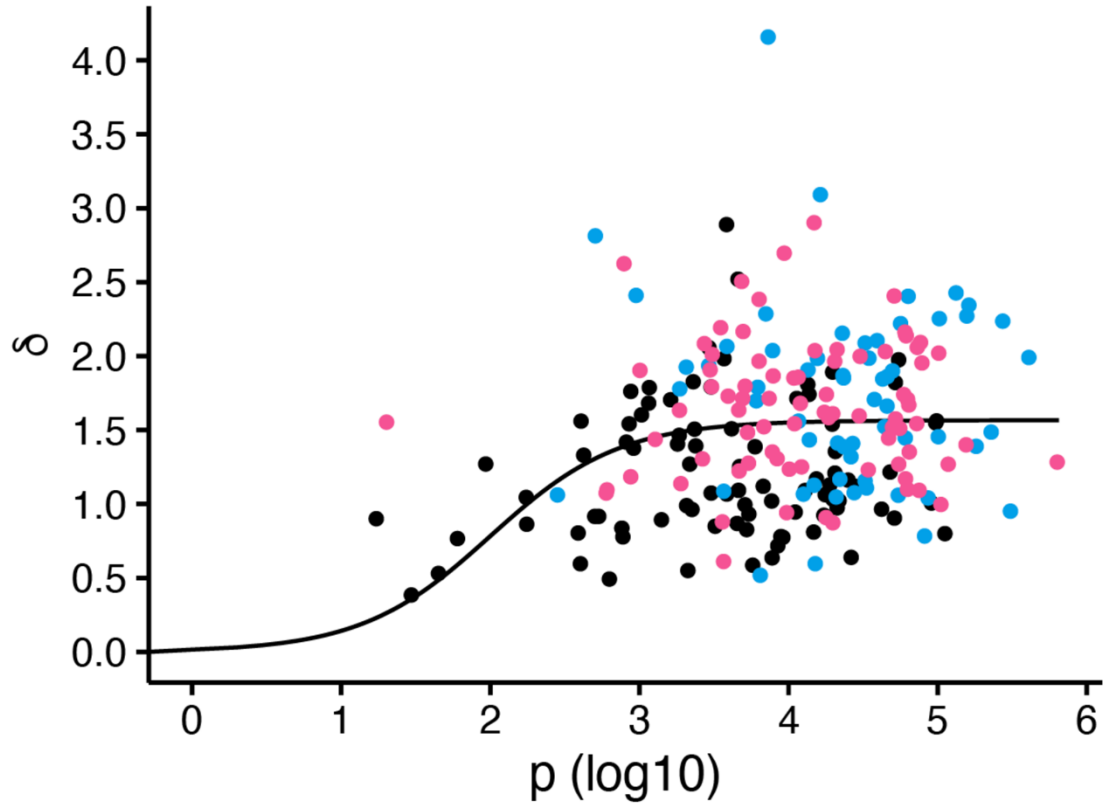
Figure S9: Possible mechanisms of adaptive virus response under isolation.



Selection pressures mediated by isolation (left panel) and general evolutionary trends (right panel) are described. The solid circular arrow in the left panel represents that isolation can be a substantial selective force only when a focal patient is symptomatic. In this case, the advanced viral peak evolves to mitigate the selection pressure. However, asymptomatic infection is always favored from a perspective of virus evolution that maximizes the number of secondary infectious individuals because isolation does not limit transmissions. Thus, the dashed circular arrow in the left panel shows that no selective force is generated by isolation when the focal patient is asymptomatic. In the right panel, the possible general evolutionary trends show that a more extended incubation period (i.e., larger T^*) and/or a higher proportion of asymptomatic infection (i.e., smaller f) may generally evolve to mitigate the selection pressure of isolation. The black trajectory represents evolutionary transitions among the pre-Alpha, Alpha, Delta, and Omicron variants of SARS-CoV-2 in our analysis. The incubation period decreases as SARS-CoV-2 mutates, implying the selection pressure by isolation, and therefore the peak viral load is advanced as observed in the Delta variant. In addition, a high proportion of asymptomatic infections has recently evolved. It is also noteworthy that once the higher proportion of asymptomatic infection

evolves, other viral phenotypic evolution, such as slower and/or lower peak viral load, may be allowed to evolve as observed in the Omicron variant. These multiple evolutionary outcomes may be counterbalanced during viral evolution under human behavior change (see the main text for further details).

Figure S10: Relation between viral production and death rate of infected cells.



Each dot shows the estimated values of p and δ for each individual for infection with the pre-Alpha, Alpha, and Delta strains. The solid line is the best-fitting curve for $\delta = \delta_{\max} p / (p + p_{50})$, whose root mean square error (RMSE) is 0.4. Pre-Alpha, Alpha, and Delta variants are colored in black, blue, and red, respectively.