Epidemic spreading under pathogen evolution

Xiyun Zhang
Jinan University https://orcid.org/0000-0002-7694-234X

Ruan Zhongyuan
Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou

Muhua Zheng
East China Normal University https://orcid.org/0000-0002-9811-529X

Jie Zhou
East China Normal University

Boccaletti Stefano
CNR

Baruch Barzel (baruchbarzel@gmail.com)
Bar-Ilan University https://orcid.org/0000-0001-8862-4384

Article

Keywords: pathogen evolution, epidemic, COVID-19, mutant strain

Posted Date: April 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-373402/v1

Version of Record: A version of this preprint was published at Nature Communications on October 20th, 2022. See the published version at https://doi.org/10.1038/s41467-022-34027-9.

Epidemic spreading under pathogen evolution

Xiyun Zhang,[1,*] Zhongyuan Ruan,[2] Muhua Zheng,[3] Jie Zhou,[4]

Stefano Boccaletti[5,6,7,8,†] & Baruch Barzel[9,10,†]

1. Department of Physics, Jinan University, Guangzhou, Guangdong 510632, China
2. Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China
3. School of Physics and Electronic Engineering, Jiangsu University, Zhenjiang, Jiangsu, 212013, China
4. School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China
5. CNR - Institute of Complex Systems, Via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy
6. Unmanned Systems Research Institute, Northwestern Polytechnical University, Xi'an 710072, China
7. Moscow Institute of Physics and Technology (National Research University), 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russian Federation
8. Universidad Rey Juan Carlos, Calle Tulipán s/n, 28933 Móstoles, Madrid, Spain
9. Department of Mathematics, Bar-Ilan University, Ramat-Gan, 5290002, Israel
10. Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, 5290002, Israel

* **Correspondence**: xiyunzhang@jnu.edu.cn

† **These Authors equally contributed to the manuscript**

Battling a widespread pandemic is an arms race between our mitigation efforts, *e.g.*, social distancing or vaccination, and the pathogen's evolving persistence. This is being observed firsthand during the current COVID-19 crisis, as novel mutations are constantly challenging our global vaccination race. To address this, we introduce here a general framework for epidemic spreading under pathogen evolution, which shows that mutations can fundamentally alter the projection of the spread. Specifically, we detect a new pandemic phase - the mutated phase - in which, despite the fact that the pathogen is initially non-pandemic ($R_0 < 1$), it may still spread due to the emergence of a critical mutation. The boundaries of this phase portray a balance between the epidemic and the evolutionary time-scales. If the mutation rate is too low, the pathogen prevalence decays prior to the appearance of a critical mutation. On the other hand, if mutations are too rapid, the pathogen evolution becomes volatile and, once again, it fails to spread. Between these two extremes, however, a broad range of conditions exists in which an initially sub-pandemic pathogen will eventually gain prevalence. This is especially relevant during vaccination, which creates, as it progresses, increasing selection pressure towards vaccine-resistance. To overcome this, we show that vaccination campaigns must be accompanied by fierce mitigation efforts, to suppress the potential rise of a resistant mutant strain.

Evolutionary time-scales are often considered to be vast, occurring gradually over the course of millions of years. However, if prevalent enough, a species may undergo even rare mutations at relatively short time-scales. This is especially relevant during the course of a widespread and prolonged pandemic.[1–6] The global spread ensures a sufficiently large pool of pathogens for mutations to occur, and on top of that, the long duration of the pandemic[7–13] affords the pathogens sufficient time to evolve.

Such troubling scenario is currently unfolding in the case of COVID-19, where novel mutations of the SARS-CoV-2 virus continue to challenge our mitigation efforts.[14–18] They are, however, equally relevant in other infections, such as influenza A, forcing us to distribute a dedicated

vaccine in each yearly cycle.[19–24] Another notable example is norovirus, whose enhanced transmission, likely due to mutation, led to an observable spike in gastric flu patients in England and Wales from 1991 to 2006,[25] and finally, beyond viruses, artemisinin-resistance, a parasite mutation, rendered void the common treatment of malaria in Africa.[26,27]

The common approach for tracking the spread of evolving pathogens is to introduce several competing strains and extract their interacting contagion process.[28–32] This captures the patterns of spread of already evolved pathogens, overlooking the dynamics, and most importantly, the time-scales, of the *evolution* itself. Indeed, in an ongoing pandemic, mutations represent a gradual random process, in which an originally unfit pathogen mutates step-by-step via a series of small changes, until reaching a critical mutation that allows it to efficiently spread. Such process may take a significant amount of time, and, in some cases, the disease may taper off before such critical mutation has the opportunity to take over.

Another crucial aspect, absent when considering pre-mutated strains, is the fact that pathogen evolution is *responsive*. As we tighten our mitigation, either through prophylactic measures[33–35] or via pharmaceutical interventions, we induce a selective pressure for mitigation resistance. For example, if one enforces social distancing to push the reproduction rate R_0 below the pandemic threshold, the pathogen becomes naturally pressured towards higher transmissibility. Similarly, if one employs therapeutic treatment to expedite recovery, natural selection will push the pathogen to higher drug-persistence.

To address this, we introduce here an evolving pathogen model, which encompasses the delicate interplay between the pathogen's spread and its developing fitness. The evolution, a random walk in *fitness space*, is driven by the pathogen's mutation rate. At the same time the *natural selection*, in which the fitter strains proliferate, is pushed by the epidemiological parameters, characterizing how fast a mutated strain propagates. Together, we identify a rather broad set of conditions - the *mutated phase* - in which a non-pandemic pathogen will eventually reach an evolved pandemic state.

We find that besides the classic epidemiological parameters, *i.e.* infection/recovery rates, two additional components factor in - the mutation rate governing the evolutionary time-scales, and the number of infected individuals, which determines the likelihood of a critical mutation to occur within the relevant time-frame. Therefore, as opposed to classic pandemic transitions, which depend solely on the epidemiological parameters,[1–5] here the current prevalence $\rho(t)$ of the pathogen has direct impact on its anticipated spread. This has significant implications pertaining to our two main mitigation strategies • *Social distancing* suppresses the reproduction number R_0 to below the pandemic threshold.[7–13] However, if many have already been infected, *i.e.* $\rho(t)$ is large, then a stricter suppression may be required to avoid the emergence of a critical mutation. This indicates that the projection of the spread, and hence also its mitigation, depends on its present state $\rho(t)$ - a hysteresis phenomenon, unobserved in the classic modeling frameworks[36–41] • *Vaccination campaigns* create strong evolutionary pressure towards a vaccine resistant mutation, whose risk, once again, is directly related to the current pathogen prevalence. Hence, to succeed, we show that vaccine roll-out must be coupled with fierce suppression via social distancing.

Evolving pathogen model

Consider a social random network of N individuals linked through the adjacency matrix $A \equiv \{A_{ij}\}$ and with average degree $\overline{k}$. At $t = 0$ the network experiences an outbreak, which then spreads via the susceptible-infected-susceptible[42] (SIS) dynamics. In the classic SIS formulation, the projected spread is driven by two time-independent parameters: the recovery rate μ and the infection rate β, whose ratio $R_0 = \overline{k}\beta/\mu$, the *reproduction number*, determines the state of the system - pandemic ($R_0 \geq 1$) or healthy ($R_0 < 1$). Here, however, the pathogen is allowed to evolve, therefore these parameter may change over the course of the spread. This is captured by the individual recovery rate

$$\mu_i(t) = \frac{1}{F_i(t)}\mu, \tag{1}$$

where the *fitness* $F_i(t)$ stands for the level of mutation of the pathogen carried by individual i at time t, hence the unmutated pathogen has $F_i(0) = 1$. The above equation models the fact that (i) each individual i may carry a distinct version of the virus; (ii) this version may gradually change in time t due to mutations. The smaller is $\mu_i(t)$, the higher is the transmissibility of the pathogen, as described by the evolving reproduction number

$$R_i(t) = \frac{\overline{k}\beta}{\mu_i(t)} = R_0 F_i(t). \tag{2}$$

Indeed, a low rate of recovery $\mu_i(t)$ extends the duration of the infectious state, providing individual i with more opportunities to infect their peers. Hence, as the r.h.s. of (2) indicates, pathogens with increased $F_i(t)$ exhibit higher reproduction, and therefore spread more efficiently than their lower fitness competitors.

Mutation may also impact the transmissibility of the pathogen directly by altering the value of the infection rate β, *e.g.*, by evolving a more infectious strain. However, in the SIS framework, the relevant parameter is not μ nor β, but their ratio, as provided by $R_i(t)$.[8,11,12,43,44] Therefore, for simplicity, in (1) we only track the pathogen evolution through $\mu_i(t)$, and its subsequent $R_i(t)$, setting β stationary. To complement this analysis, in Supplementary Section 1 we examine the case of β-mutations.

The spread is driven by the infection, recovery and mutation processes. The process of *infection* between a pair of individuals i and j is modeled by

$$S_i + I_j \xrightarrow{A_{ij}\beta} I_i + I_j \tag{3}$$

$$F_i(t) = F_j(t), \tag{4}$$

in which a susceptible (S) individual i interacts with their infected (I) neighbor j ($A_{ij} = 1$) at rate β. This leads to both individuals becoming infected. The newly infected individual i inherits j's pathogen, and hence in (4) we set i's fitness at the time of infection equal to that of j. Both fitness parameters, $F_i(t)$ and $F_j(t)$ may later change via mutation. Next, we consider the process of *recovery*

$$I_i \xrightarrow{\mu_i(t)} S_i, \tag{5}$$

in which an infected individual I_i transitions to S_i at the evolved recovery rate $\mu_i(t)$ of (1). Finally, the process of *mutation* follows

$$\begin{cases} F_i(0) = 1 \\ F_i(t+1) = \max\left(F_i(t) + \delta_i(t)\ ,\ 0\right) \end{cases}, \tag{6}$$

capturing a random walk with variable step size $\delta_i(t)$, *i.e.* a sequence of random shifts in fitness, caused by small discrepancies in the pathogen's reproduction. Note that $F_i(t)$ is prohibited from becoming negative, as, indeed, a below zero fitness in (2) is meaningless. The case where $F_i(t)$ does approach zero corresponds to $\mu_i(t) \to \infty$ in (1), a limit in which recovery is instantaneous, and hence the pathogen is *unfit* for reproduction. Such strains will be rapidly eliminated from the pathogen pool.

The magnitude of each mutation step is extracted from a zero-mean normal distribution, namely $\delta_i(t) \sim \mathcal{N}(0, \sigma^2)$. Consequently, in the limit where $\sigma = 0$, we have $\delta_i(t) = 0$ at all times, mutations are suppressed, and Eqs. (3) - (6) converge to the classic SIS model, with $R_i(t) = R_0$, a constant reproduction number. In contrast, as σ is increased, significant mutations become more frequent and the pathogens rapidly evolve. We therefore vary σ to control the *mutation rate* of the pathogens.

Taken together, our modeling framework accounts for the dynamics of infection and recovery (SIS) under the effect of pathogen mutation. As the spread progresses, pathogens evolve via Eq. (6), blindly altering their epidemiological parameters at random. Natural selection, however, will favor the positive mutations, in which $\delta_i(t) > 0$. Indeed, such mutations lead to higher fitness, reducing the recovery rate $\mu_i(t)$, and consequently increasing $R_i(t)$. Such pathogens, with increased $R_i(t)$, will proliferate more rapidly, and will eventually dominate the population.

Critical mutation. Consider an outbreak of a pathogen with $R_0 < 1$, *i.e.* below the epidemic threshold. This can be either due to the pathogen's initial sub-pandemic parameters, or a result of mitigation, *e.g.*, social distancing to reduce β. In the classic SIS formulation, such pathogen with fail to penetrate the network. However, in the presence of mutations ($\sigma > 0$) the pathogen may potentially undergo selection, reach $R_i(t) > 1$, and from that time onward begin to proliferate. This represents a *critical mutation*, which, using (2), translates to

$$F_c = \frac{1}{R_0}, \tag{7}$$

the *critical fitness* that, once crossed, may lead an initially non-pandemic pathogen to become pandemic. The smaller is R_0 the higher is F_c, as, indeed, weakly transmissible pathogens require a longer evolutionary path to reach pandemic spread.

Next, we analyze the spreading patterns of our evolving pathogens, seeking the conditions for the appearance of the critical mutation.

Phase-diagram of evolving pathogens

To examine the behavior of (1) - (6) we constructed an Erdős-Rényi (ER) network with $N = 5,000$ nodes and $\overline{k} = 15$, providing a *testing ground* upon which we incorporate a series of epidemic scenarios (Fig. 1). Each scenario is characterized by a different selection of our model's three epidemiological parameters: μ and β, which determine the pathogen's unmutated repro-

duction R_0, and σ, which controls the rate of mutation. We then follow the spread by measuring the prevalence $\rho(t)$, which monitors the fraction of infected individuals vs. time. We also track the pathogen's evolution via the population averaged fitness $\overline{F}(t) = (1/N)\sum_{i=1}^{N} F_i(t)$.

Pandemic (Fig. 1a, red). In our first scenario we set $\mu = 0.1$, $\beta = 8 \times 10^{-3}$ and $\sigma = 10^{-2}$. This captures a pandemic pathogen, which, using $\overline{k} = 15$, has $R_0 = 1.2 > 1$, namely it can spread even without mutation. Indeed $\rho(t)$ rapidly climbs to gain macroscopic coverage, congruent with the prediction of the classic SIS model, but this time constantly growing, due to the gradual, but continuous, increase in fitness $\overline{F}(t)$.

Mutated (Fig. 1c, green). Next we reduce the infection rate to $\beta = 1.67 \times 10^{-3}$, an initial reproduction of $R_0 = 0.25 < 1$. This describes a pathogen whose transmissibility is significantly below the epidemic threshold, and therefore, following the initial outbreak we observe a decline in $\rho(t)$, which by $t \sim 50$ almost approaches zero, as the disease seems to be tapering off. In this scenario, however, we set a faster mutation rate $\sigma = 1$. As a result, despite the initial remission, at around $t \sim 15$, the pathogen undergoes a critical mutation as $\overline{F}(t)$ crosses the critical $F_c = 1/R_0 = 4$ (grey dashed line) and transitions into the pandemic regime. Consequently, $\rho(t)$ changes course, the disease reemerges and the mutated pathogens successfully spreads.

Lethargic (Fig. 1b, blue). We now remain in the sub-pandemic regime, with $R_0 = 0.25$, but with a much slower mutation rate, set again to $\sigma = 10^{-2}$. As above, $\rho(t)$ declines, however the pathogen evolution is now too slow - it is *lethargic*, and cannot reach critical fitness on time. Therefore, the disease fails to penetrate the network, lacking the opportunity for the critical mutation to occur.

Taken together, the dynamics of the spread are driven by three parameters: the initial epidemiological characteristics of the pathogen, μ and β, which determine R_0, and the mutation rate σ, which governs the time-scale for the appearance of the critical mutation. Therefore, to determine the conditions for a mutation-driven contagion, as observed in Fig. 1c, we investigate the balance between the decay in $\rho(t)$ vs. the gradual increase in $\overline{F}(t)$.

The mutated phase

To understand the dynamics of the evolving pathogen model, we show in Supplementary Section 2 that at the initial stages of the spread, the prevalence $\rho(t)$ follows

$$\rho(t) = \rho(0)e^{\xi(t)}. \tag{8}$$

The time-dependent exponential rate $\xi(t)$ is determined by the epidemiological/mutation rates via

$$\xi(t) = -\mu(1 - R_0)t + \frac{1}{2}\sigma^2\mu^2 R_0^2 t^3, \tag{9}$$

whose two terms characterize the pre-mutated vs. post-mutated spread of the pathogen. The first term, linear in t, represents the initial patterns of spread, which are determined by the original pathogen parameters, μ, R_0. For $R_0 < 1$ this describes an exponential decay, *a là* SIS dynamics in the sub-pandemic regime. At later times, however, as t^3 becomes large, the second term begins to dominate, and the exponential decay is replaced by a rapid proliferation, now driven by the mutation rate σ. The transition between these two behaviors - decay vs. proliferation - occurs at $\tau_c = \sqrt{2(1 - R_0)/3\mu\sigma^2 R_0^2}$, which provides the anticipated time-scale for the appearance of the critical mutation F_c in (7).

This analysis portrays the mutated phase as a balance between two competing time-scales: on the one hand the exponential decay of the sub-pandemic pathogen, and on the other hand the evolutionary time-scale τ_c for the appearance of the critical mutation. For the evolution to *win* this race the pathogen must not vanish before $t = \tau_c$. This imposes the condition (Fig. 2a-c)

$$\rho(\tau_c) \geq \frac{1}{N}, \tag{10}$$

ensuring that at τ_c there are still one or more individuals hosting the pathogen. Indeed, $\rho(\tau_c) < 1/N$ indicates that *on average*, at $t = \tau_c$ less than a single individual is left in the infected pool. Under this condition, the critical mutation is too late, the spread has already tapered off, and the exponential growth driven by the positive term in (9) is averted.

Taking $\rho(\tau_c)$ from (8), we can now use (10) to express the boundary of the mutated phase, predicting the *critical mutation rate* as (Supplementary Section 2)

$$\sigma_c \sim \left(\frac{\sqrt{\mu(1-R_0)^3}}{2R_0} \right) \frac{1}{\ln(\mathcal{I}_0)}, \tag{11}$$

where $\mathcal{I}_0 = N\rho(t=0)$ is the number of individuals infected at $t = 0$. Equation (11) describes the minimal mutation rate required for the pathogen to evolve a pandemic strain. For $R_0 = 1$ it predicts $\sigma_c = 0$, as such pathogen can indeed spread even without mutation. However, as R_0 is decreased, for example under mitigation, the pathogen prevalence rapidly declines, and hence it must evolve at an accelerated rate to reach critical fitness. This is expressed in (11) by an increased σ_c, which approaches infinity as $R_0 \to 0$.

To test our predicted phase transition we simulate in Fig. 1d an array of $1,050$ realizations of Eqs. (1) - (6), representing different epidemiological scenarios. We varied R_0 from 0 to 1.5, *i.e.* from non-transmissible to highly contagious, and scanned a spectrum of mutation rates from $\sigma = 10^{-3}$ to $\sigma = 10$, spanning four orders of magnitude. Simulating each scenario 50 times we observe the probability P for the disease to spread. This is done by tracking the pathogen's long-term prevalence $\rho = \rho(t \to \infty)$ and counting the realizations in which $\rho \to 0$ vs. those where $\rho > 0$. As predicted, we find that the pandemic state, classically observed only at $R_0 \geq 1$, now extends to lower R_0 in the presence of sufficiently rapid mutations. This gives rise to the *mutated phase* (green), in which an initially decaying contagion suddenly turns pandemic. The transition between the lethargic and the mutated states (grey zone) is well-approximated by our theoretical prediction of Eq. (11), as depicted by the black solid line.

Equation (11) shows that σ_c depends not only on the epidemiological characteristics of the pathogen (μ, R_0), but also on the initial condition, here captured by the number of infected individuals $\mathcal{I}_0 = \rho(t=0)N$. If $\mathcal{I}_0$ is large the critical rate σ_c becomes lower, in effect expanding the bounds of the mutated phase. To understand this consider the evolutionary paths followed by the pathogens as they reproduce. These paths represent random trajectories in *fitness space*, each starting from $F_i(0) = 1$, and with a small probability crossing the critical fitness F_c. The more such attempts are made, the higher the chances that at least one of these paths will be successful. Therefore, a higher initial prevalence $\mathcal{I}_0$ of the pathogen increases the probability for the appearance of a critical mutation, enabling a mutated phase even with low σ. In simple words, even rare mutations may occur if the initial pathogen pool $(\mathcal{I}_0)$ is large enough. Indeed, in Fig. 2d we find that the phase boundary shifts towards lower σ_c as the initial prevalence is increased (grey shaded lines). Hence, a greater $\mathcal{I}_0$, indeed, expands the mutated phase.

Hysteresis. This dependence on $\mathcal{I}_0$ indicates that the transition of Eq. (11) behaves differently if we approach it from the pandemic state or from the healthy state. To observe this let us fix the mutation rate at $\sigma = 0.1$ and gradually increase R_0, seeking the critical point where the system shifts to the mutated phase. This is mapped to a vertical trajectory in the σ, R_0 plane (Fig. 2d, yellow dashed line). At each value of R_0 we instigate an outbreak with $\rho(0) = 0.2$, and observe its long-term prevalence ρ. For small R_0 this outbreak decays and the system reverts to the healthy state $\rho = 0$. However, as we transition into the mutated phase, here predicted at $R_0 = R_{\text{High}} \approx 0.6$, the pathogen turns pandemic and its prevalence abruptly changes to $\rho \approx 0.85$.

To reverse this transition the naïve approach is to push R_0 slightly below this critical point, for instance, by practicing social distancing to reduce transmission. The challenge is that now, moving in the opposite direction - from large to small R_0 - our initial condition is pandemic, with prevalence of order unity ($\sim 85\%$), and hence $\mathcal{I}_0 \sim N$. Under these conditions, Eq. (11) predicts that, for our fixed σ, the critical R_0 is now lower, at $R_{\text{Low}} = 0.35$. This results in a hysteresis phenomenon, in which criticality occurs at different points depending on the state from which we approach the transition (Fig. 2e).

We find, therefore, that pathogen evolution fundamentally changes the phase space of epidemic spreading. First it predicts a broad range of conditions - the mutated phase - in which a sub-pandemic pathogen can gain prevalence. On top of that, it also predicts that this phase exhibits a discontinuous transition, characterized by hysteresis, a phenomenon unobserved in the classic SIS dynamics, yet congruent with other models[38,41,45–49] that incorporate feedback between a pathogen's prevalence ($\rho(t)$) and its potency ($R_i(t)$). These two observations have direct implications on mitigation:

- **Soft mitigation is risky**. Most mitigation strategies seek a *minimal* approach, aiming to drive R_0 just below unity. This is understandable as (i) major restrictions on social interactions are costly and difficult to sustain[50] for extended periods; (ii) having $R_0 < 1$, even by a small margin, is assumed to naturally suppress the spread, as it leads $\rho(t)$ to decay exponentially towards zero. Our analysis, however, shows that this is insufficient. For $R_0 \lesssim 1$ we have $\sigma_c \to 0$, indicating that even a relatively stable pathogen, with a low mutation rate, may eventually break through. Using Eq. (11) we can predict for a given σ, the level of *tolerable* R_0 that is sufficient to mitigate the mutated phase risk, providing guidelines for effective mitigation.
- **The sooner the better**. Another common assumption, driven by the classic epidemic phase-diagram, is that the projected state $\rho(t \to \infty)$ depends only on R_0, *i.e.* the epidemiological parameters. The current state of the spread $\rho(0)$ at the time we implement our mitigation, plays no role. The observed hysteresis, however, shows that successful mitigation strongly depends on the prevalence at the time of instigation. If the pathogen has already gained sufficient ground, we will need to suppress the reproduction number below R_{Low}, namely the lower phase-boundary in Fig. 2e. It is, therefore, crucial to respond early, and initiate our mitigation when $\rho(t)$ is still small, eradicating the pandemic before mutations may determine a risk for its reemergence.

Bounded fitness. Our mutation process in Eq. (6) allows the pathogen an unbounded random walk in fitness space. In reality, however, there are practical restrictions on fitness, as $R_i(t)$ cannot grow *ad infinitum*. Therefore, we now consider our evolving pathogen model, substituting the mutation in (6) with

$$F_i(t+1) = \min\Big(F_{\max}, \max\big(F_i(t) + \delta_i(t)\ ,\ 0\big)\Big), \tag{12}$$

in which the pathogen fitness is bounded from above by $F_{\max}$ and from below by zero. Setting $F_{\max} = 20$ we now revisit our phase-diagram (Fig. 3a). For small σ, mutations are slow, and the evolution path is unaffected by the upper bound on $F_i(t)$. Therefore, we continue to observe the same transition as in the unbounded model of Fig. 1d. As we increase σ, however, we witness a second transition, this time back to the healthy state, indicating that now, mutations are *too* rapid. This captures the final phase of our evolving pathogen model - the *volatile* phase:

Volatile (Fig. 3, blue). When the mutation rate is too high the pathogen fitness becomes unstable. On the one hand it can rapidly reach critical fitness, yet, on the other hand, due to the random nature of its frequent mutations, it fails to sustain this fitness - resulting in an irregular $\overline{F}(t)$, that fluctuates above and below the critical F_c (Fig. 3c).

To gain deeper insight into the volatile phase, consider the natural selection process, here driven by the reproduction benefit of the fitter strains. This process is not instantaneous, and requires several reproduction instances, *i.e.* generations, to gain a sufficient spreading advantage. With σ too high, natural selection is confounded, the pathogen shown no consistent gain in fitness and, as Fig. 3c indicates, $\rho(t)$ decays exponentially to zero. In Supplementary Section 3 we use a time-scale analysis, similar to the one leading to Eq. (11), to show that the volatile phase occurs when σ exceeds

$$\sigma_c \sim \sqrt{\frac{\mu}{3}}\frac{(F_{\max}R_0 - 1)^{\frac{3}{2}}}{R_0}. \tag{13}$$

This prediction is, indeed, confirmed by our simulated phase diagram in Fig. 3a (black solid line).

Our phase-diagram illustrates the different forces governing the spread of pathogens in the presence of mutations. While spread is prohibited classically for $R_0 < 1$, here we observe a new, previously undocumented pandemic phase, in which the disease can successfully permeate despite having an initially low reproduction rate. The conditions for this phase require a balance between three separate time-scales: (i) The time for the initial outbreak $\rho(0)$ to reach near zero prevalence τ_r; (ii) The time for the pathogen to evolve beyond critical fitness τ_c; (iii) The time for the natural selection to *lock-in* the fitter mutations τ_s. Pathogens with small R_0, we find, can still spread provided that

$$\tau_r > \tau_c > \tau_s. \tag{14}$$

The l.h.s. of (14) ensures that the pathogen can reach critical fitness *before* reaching zero prevalence. This gives rise to the first transition of Eq. (11), between the *lethargic* and the *mutated* phases. The r.h.s. of (14) is responsible for the second transition, from *mutated* to *volatile*. It ensures that fitter pathogens do not undergo additional mutation before they have time to proliferate via natural selection. Therefore, we observe a *Goldilocks zone*, in which the mutation rate σ is just right: on the one hand, enabling unfit pathogens to cross the Rubicon towards pandemicity, but on the other hand, avoiding aimless capricious mutations.

Mutation risk in vaccine distribution - the case of COVID-19

Vaccination during an ongoing pandemic is, by nature, a competition between the rate of the vaccine roll-out and the spread of the pathogen.[51–53] Therefore, naïvely, to win this race all one has to do is disseminate the vaccine as efficiently as possible, aiming to reach the majority of the population before the pathogen does. This, however, ignores the role of mutations, which may gravely impact even the most efficient vaccination campaign. Such mutations may, generally, be less fit epidemiologically, *i.e.* have a lower F and consequently a lower $R_i(t)$. Therefore, absent a vaccine, they will be rapidly overcome by the faster spreading pathogen strains. However, once the vaccine becomes widespread, resistance, even if less contagious, becomes a highly desirable trait, and a resistant mutation, if occurs, will take over the population.

To examine this in a realistic setting we consider the spread of SARS-CoV-2, currently battled by a global vaccination effort. To model the disease dynamics we collected data on the COVID-19 infection cycle (Fig. 4a), which includes a well-documented and elaborate set of transitions.[54–64] Upon infection, individuals enter a pre-symptomatic state, which lasts, on average 5 days. During this period, typically within $2-4$ days they begin to shed the virus and infect their network contacts (PS, purple). This continues until the onset of mild (I_M), severe (I_S) or critical (I_C) symptoms, at which point they enter isolation and cease to spread the virus. A fraction ($\sim 30\%$) of infected individuals never go on to develop noticeable symptoms (AS, top arrow), and hence they continue to spread the virus until their full recovery (R), typically within ~ 7 days.

To evaluate the infection rate β we used empirical data on the observed spread in 12 different countries.[68] Focusing on the early stages of the contagion, prior to the instigation of mitigation strategies, we find that $\beta = 5\times10^{-2}$ best fits the observed spreading dynamics. This corresponds to a reproduction rate of $R_0 \approx 2.6$, congruent with existing valuations of R_0 under COVID-19.[55,69] For details on the data analysis see Supplementary Section 4.

Here we complement this disease cycle by two additional processes

- **Vaccination**. The population is vaccinated at a rate ν, quantifying the percentage of the (susceptible) population that receives the vaccine per unit time (day).
- **Resistance**. At each time-step, the pathogen may undergo a vaccine-resistant mutation with probability p. This mutation has no bearing on its epidemiological parameters μ, β, thus providing no additional spreading advantage, other than being resistant to the vaccine. The larger is the infected population ($\rho(t)N$) the greater is the risk for such mutation, hence we quantify the *mutation risk* via $\mathcal{P} = pN$, and examine two scenarios: high risk with $\mathcal{P} = 2.5$ and low risk, setting $\mathcal{P} = 0.25$. For a population of $N \sim 10^9$ both cases capture a very rare mutation with $p \sim 10^{-9}$ and 10^{-10}, respectively.

Two factors drive the level of risk in this process. The prevalence $\rho(t)$ determines the size of the pathogen pool, which must be large for the rare mutation to be realized. The vaccine coverage $V(t)$ determines the selective advantage of the resistant strain, which becomes marginal if only a small fraction of the population is inoculated. Therefore the highest risk occurs under the coexistence of both infected and vaccinated individuals. This enables the interaction between these two populations paving the way for both mutation (large $\rho(t)$) and selection (large $V(t)$), and hence potentially driving the system towards vaccine resistance (Fig. 4b).

To observe this we simulated three vaccination strategies, under the high risk $\mathcal{P} = 2.5$ scenario • **Slow** (Fig. 4c,d). First we assume a slow vaccination rate of $\eta = 10^-3$, a 0.1% daily coverage. Such slow vaccination is insufficient to suppress $\rho(t)$, allowing us, after some time to enter the

risk zone in which $\rho(t)$ coexists with $V(t)$ (shaded). Mutations occurring within this window (orange) are likely to proliferate. Indeed, we find that in the long term, vaccination fails, and the resistant strain gains coverage. • **Rapid** (Fig. 4e,f). To overcome this we simulate a rapid vaccine roll-out with $\eta = 10^{-2}$, capturing an optimistic scenario, in which 1% of the population is inoculated per day. Despite these favorable conditions we continue to enter the risk zone, as the pathogen is allowed to spread freely in parallel to our vaccination efforts. The result is, as before, an increased likelihood of a resistant mutation, which, once again, regardless of our efficient dissemination, renders our vaccination void. • **Combined effort** (Fig. 4g,h). The only way to avoid the risk zone is to minimize the potential interaction between infected and vaccinated individuals. Since $V(t)$ will inevitably grow - indeed, this is the goal of vaccine distribution - we must contain $\rho(t)$, namely aim for the right-most branch of the risk curve in Fig. 4b. This requires a combined effort of both rapid vaccination ($\eta = 10^{-2}$) and fierce mitigation to suppress R_0. The result is a successful elimination of the pathogen with $V(t) \to 1$ and $\rho(t) \to 0$.

In Fig. 4i,j we systematically plot the spreading probability P in function of η, R_0 under our high/low risk scenarios. We find that for COVID-19, having $R_0 \approx 2.6$ (black solid line) the risk of vaccine resistance is significant, even under large η. Reducing R_0 via social distancing helps alleviate this risk. For example, for $\mathcal{P} = 2.5$, even is we assume a rapid roll-out (large η), we must reach $R_0 \lesssim 2$ to remain within a low mutation risk (blue). Under $\mathcal{P} = 0.25$, it is sufficient to aim for $R_0 \lesssim 3$, roughly the natural state of SARS-CoV-2.

Discussion and outlook

The phase diagram of epidemic spreading is a crucial tool for forecasting and mitigating pandemic risks. First, it identifies the relevant control parameters, such as μ, β and $\overline{k}$ in our SIS framework, or additional parameters in more complex contagion processes, whose value determines R_0 and hence the expected patterns of spread. The phase boundaries, then, help us assess the state of the system - healthy or pandemic - and provide guidelines for our response. For example, social distancing to reduce $\overline{k}$, therapeutic treatment to increase μ or mask wearing to suppress β, all aimed to navigate the system's location along the pandemic phase-diagram towards the desired healthy state.

The common thread binding all of these strategies is the assumption that the epidemiological control parameters themselves are constant in time, and hence our intervention must *just* push them beyond the static phase-boundary, from which point on the epidemic will decay towards $\rho \to 0$ spontaneously. This is, indeed, relevant if the temporal evolution of μ, β is slow compared to the epidemic spreading dynamics - as observed in the case of our lethargic phase. However, once the epidemiological parameters can change at a sufficiently high rate, it fundamentally changes the *rules of the game*. This is because now, not only are the parameters *dynamic*, but, thanks to natural selection, they also become *responsive*. If, for instance, we develop drug-based treatment to increase the recovery rate μ, we inevitably also generate selection pressure towards drug persistence. Similarly, if we vaccinate or practice distancing to reduce $\overline{k}, \beta$, we initiate an evolutionary race towards higher transmissibility or vaccine resistance. This was clearly observed in our analysis of the COVID-19 vaccine dissemination.

The result is a complex interplay between the spreading dynamics (R_0), the instantaneous prevalence of the pathogen ($\rho(t)$), and the dynamic evolution of its parameters (σ), which reshapes the pandemic phase diagram. It not only expands the pandemic risk to a range of $R_0 < 1$, but also predicts an explosive transition pattern, *i.e.* the hysteresis of Fig. 2e, that is

not observed in standard epidemiological transitions. This altered phase diagram, and its abrupt first-order like transition, we have shown, has crucial implications pertaining to our mitigation strategies. Yet, more broadly, as a physical phenomenon, it offers an interesting mechanism for explosive transitions. Most often, such abrupt phase-shifts are caused by internal suppression rules, that hold back the transition until it breaks through in an explosive fashion.[37,70–72] In contrast, here what holds back the transition is the waiting time for the critical mutation. Until its appearance the system behaves in one way ($R_0 < 1$), but once it occurs, the system suddenly enters the pandemic regime ($R_0 > 1$). The explosiveness is therefore traced to a local event, whose probability depends on the system's initial parameters ($R_0, \mathcal{I}_0, \sigma$). This local event then changes fundamentally the state of the system - capturing a feedback between the system's phase and its intrinsic control parameters. We believe this describes a unique mechanism, inherent to the basic ingredients of our biological system, reproduction, mutation and selection.

Acknowledgements

X.Z. thanks Dr. Xiaobo Chen, Dr. Tingting Shi, Dr. Xing Lu and Prof. Weirong Zhong for useful discussions and supports in numerical calculations. This work was partially supported by the National Natural Science Foundation of China under Grants No. 12075008 and No. 1200050749. This research was also supported by the Israel Science Foundation (grant No. 499/19) and the Bar-Ilan University Data Science Institute grant for COVID-19 related research.

Author contribution

X.Z. developed the concept. X.Z., B.B. and S.B. designed the framework. X.Z. and Z.R. performed the numerical simulations. All authors jointly analyzed the results and developed the analytical framework. X.Z., B.B. and S.B. wrote the paper.

Code availability

All code to study and reproduce the results shown here will be made freely available online upon publication.

References

[1] Hufnagel, L., Brockmann, D. and Geisel, T., Forecast and control of epidemics in a globalized world, Proc. Natl. Acad. Sci. U. S. A. 101 15124-15129 (2004).

[2] Colizza, V., Barrat, A., Barthélemy, M. and Vespignani, A., The role of the airline transportation network in the prediction and predictability of global epidemics, Proc. Natl. Acad. Sci. U. S. A. 103, 2015-2020 (2006).

[3] Balcan, D. et al. Multiscale mobility networks and the spatial spreading of infectious diseases, Proc. Natl. Acad. Sci. U. S. A, 106, 21484-21489 (2009).

[4] Brockmann, D. and Helbing, D., The Hidden Geometry of Complex, Network-Driven Contagion Phenomena, Science, 342, 1337-1342 (2013).

[5] Liu, Q. et al. Measurability of the epidemic reproduction number in data-driven contact networks, Proc. Natl. Acad. Sci. U. S. A, 115, 12680-12685 (2018).

[6] Pastor-Satorras, R., Castellano, C., Van Mieghem, P. and Vespignani, A. Epidemic processes in complex networks, Rev. Mod. Phys. 87, 925–979 (2015).

[7] Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China, Science, 368, 1481-1486 (2020).

[8] Maier, B. F. and Brockmann, D., Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China, Science, 368, 742-746 (2020).

[9] Schlosser, F. et al. COVID-19 lockdown induces disease-mitigating structural changes in mobility networks, Proc. Natl. Acad. Sci. U. S. A, 117, 32883-32890 (2020).

[10] Aleta, A. et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, Nature Human Behaviour, 4, 964971 (2020).

[11] Vespignani, A. et al. Modelling COVID-19, Nature Reviews Physics, 2, 279281 (2020).

[12] te Vrugt, M., Bickmann, J. and Wittkowski, R., Effects of social distancing and isolation on epidemic spreading modeled via dynamical density functional theory, Nature Communications 11, 5576 (2020).

[13] Kogan, N. E. et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time, Science Advances 7, eabd6989 (2021).

[14] Yao, H. et al. Patient-derived mutations impact pathogenicity of SARS-CoV-2, medRxiv 2020.04.14.20060160 (2020).

[15] Konno, Y. et al. SARS-CoV-2 ORF3b is a potent interferon antagonist whose activity is further increased by a naturally occurring elongation variant, bioRxiv 2020.05.11.088179 (2020).

[16] Hu, J. et al. The D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity and decreases neutralization sensitivity to individual convalescent sera, bioRxiv 2020.06.20.161323 (2020).

[17] Korber, B. er al. on behalf of the Sheffield COVID-19 Genomics Group, Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus, Cell (2020).

[18] Grubaugh, N.D., Hanage, W.P., and Rasmussen, A.L., Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear, Cell (2020).

[19] Ojosnegros, S. and Beerenwinkel, N., Models of RNA virus evolution and their roles in vaccine design, Immunome Research 6(Suppl 2):S5 (2010).

[20] Nougairède, A., Charrel, R. N. and Raoult, D., Models cannot predict future outbreaks: A/H1N1 virus,the paradigm, Eur J Epidemiol 26:183, 186 (2011).

[21] Bedford, T., et al. Integrating influenza antigenic dynamics with molecular evolution, eLife 3:e01914 (2014).

[22] Nelson, M. I. and Holmes, E. C., The evolution of epidemic influenza, Nature Reviews Genetics 8, pages196,205 (2007).

[23] Ghedin, E. et al. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution, Nature 437, pages 1162, 1166 (2005).

[24] Earn, D. J.D., Dushoff, J. and Levin, S. A., Ecology and evolution of the flu, TRENDS in Ecology & Evolution 17, Pages 334-340 (2002).

[25] Lopman, B., Zambon, M. and Brown, D. W., The Evolution of Norovirus, the Gastric Flu, PLoS Medicine 5(2): e42. (2008).

[26] Uwimana, A. et al., Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda, Nature Medicine (2020).

[27] Birnbaum, J. et al., A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites, Science 367, 51-59 (2020).

[28] Alexander, H. and Day, T., Risk factors for the evolutionary emergence of pathogens. J. R. Soc. Interface 7, 1455-1474 (2010).

[29] Uekermann, F. and Sneppen, K., Spreading of multiple epidemics with cross immunization. Phys. Rev. E, 86, 036108 (2012).

[30] Hébert-Dufresne, L., Patterson-Lomba, O., Goerg, G. M. and Althouse, B. M., Pathogen Mutation Modeled by Competition Between Site and Bond Percolation, Phys. Rev. Lett. 110, 108103 (2013).

[31] Alexandrou, C., Harmandaris, V., Irakleous, A., Koutsou, G. and Savva, N., Modeling the evolution of COVID-19, arXiv:2008.03165 (2020).

[32] Eletreby, R. et al., The effects of evolutionary adaptations on spreading processes in complex networks, Proc. Natl. Acad. Sci. U. S. A. 117,5664 (2020).

[33] Anderson, R. M., Heesterbeek, H., Klinkenberg, S. and Hollingsworth, T.D., How will country-based mitigation measures influence the course of the COVID-19 epidemic? The Lancet, 395,10228 (2020).

[34] Shen, C. and Bar-Yam, Y, COVID-19: How to win (2020).

[35] Meidan, D. et al. Alternating quarantine for sustainable epidemic mitigation, Nature Communications 12,220 (2021).

[36] Böttcher, L., Woolley-Meza, O., Araújo, N. A. M., Herrmann, H. J. and Helbing, D., Disease-induced resource constraints can trigger explosive epidemics, Scientific Reports 5, 16571 (2015).

[37] Boccaletti, S. et al. Explosive transitions in complex networks? structure and dynamics: Percolation and synchronization, Physics Reports 660, 1-94 (2016).

[38] Liu, Q., Wang, W., Tang, M., Zhou, T. and Lai, Y., Explosive spreading on complex networks: The role of synergy, Phys. Rev. E 95, 042320 (2017).

[39] Pires, M. A., Oestereich, A. L. and Crokidakis, N., Sudden transitions in coupled opinion and epidemic dynamics with vaccination, J. Stat. Mech. 053407 (2018).

[40] Mieghem, P. V., Explosive phase transition in susceptible-infected-susceptible epidemics with arbitrary small but nonzero self-infection rate, Phys. Rev. E, 101, 032303 (2020).

[41] Zhang, X., Ruan, Z., Zheng, M., Barzel, B. and Boccaletti, S., Epidemic spreading under infection-reduced-recovery, Chaos, Solitons and Fractals 140, 110130 (2020).

[42] Pastor-Satorras, R., Castellano, C., Mieghem, P. V. and Vespignani, A., Epidemic processes in complex networks, Rev. Mod. Phys. 87, 925 (2015).

[43] Yang, H., Gu, C., Tang, M., Cai, S. and Lai, Y., Suppression of epidemic spreading in time-varying multiplex networks, Applied Mathematical Modelling, 75, 806-818 (2019).

[44] Zhai, Z., Long, Y., Tang, M., Liu, Z. and Lai, Y., Optimal inference of the start of COVID-19, Phys. Rev. Research 3, 013155 (2021).

[45] Gómez-Gardeñes, J., Lotero, L., Taraskin, S. N. and Pérez-Reche, F. J., Explosive contagion in networks, Sci. Rep. 6, 19767 (2016).

[46] Wu, J., Zheng, M., Xu, K. and Gu, C., Effects of two channels on explosive information spreading, Nonlinear Dynamics, 99, 2387-2397, (2020).

[47] Hébert-Dufresne, L., Scarpino, S. V. and Young, J., Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement, Nature Physics 16, 426 (2020).

[48] Gross, T., Lima, C. J. D. and Blasius, B., Epidemic dynamics on an adaptive network, Phys. Rev. Lett. 96, 208701 (2006).

[49] Böttcher, L., Woolley-Meza, O., Goles, E., Helbing, D. and Herrmann, H. J., Connectivity disruption sparks explosive epidemic spreading, Phys. Rev. E 93, 042315 (2016).

[50] Epstein, J.M., Modelling to contain pandemics. Nature 460,687 (2009).

[51] Harush, U. and Barzel, B., Dynamic patterns of information flow in complex net- works, Nat. Comm. 8, 2181 (2017).

[52] Hens, C., Harush, U., Cohen, R. and Barzel, B., Spatiotemporal signal propagation in complex networks, Nat. Phys. 15, 403 (2019).

[53] Hacohen, A., Cohen, R., Efroni, S., Barzel, B. and Bachelet, I., Digitizable therapeutics for decentralized mitigation of global pandemics, Scientific Reports 9, 14345 (2019).

[54] Ferguson, N. M. et al. Impact of Non-Pharmaceutical Interventions (npis) to Reduce COVID-19 Mortality and Health-Care Demand. (Imperial College COVID-19 Response Team, 2020).

[55] Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel Coronavirus-infected pneumonia. N. Engl. J. Med. 382, 1199 (2020).

[56] WHO. Coronavirus Disease (COVID-2019) Situation Report 30. (2020).

[57] Ferretti, L. et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 368, 6936 (2020).

[58] Backer, J. A., Klinkenberg, D. and Wallinga, J. Incubation period of 2019 novel Coronavirus (2019-ncov) infections among travelers from Wuhan, China, 2028 January 2020. Eurosurveillance 25, 2000062 (2020).

[59] Linton, N. M., et al. Incubation period and other epidemiological characteristics of 2019 novel Coronavirus infections with right truncation: a statistical analysis of publicly available case data. J. Clin. Med. 9, 538 (2020)

[60] Tao, Y. et al. High incidence of asymptomatic SARS-CoV-2 infection, Chongqing, China. medRxiv https://doi.org/10.1101/2020.03.16.20037259 (2020).

[61] Mizumoto, K., Kagaya, K., Zarebski, A. and Chowell, G. Estimating the asymptomatic proportion of Coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance 25, 2000180 (2020).

[62] Pan, X. et al. Asymptomatic cases in a family cluster with SARS-CoV-2 infection. Lancet Infect. Dis. 20, 410 (2020).

[63] Lu, X. et al. SARS-CoV-2 infection in children. N. Engl. J. Med. 382, 16631665 (2020).

[64] Al-Tawfiq, J. A. Asymptomatic coronavirus infection: MERS-CoV and SARS-CoV-2 (COVID-19). Travel Med. Infect. Dis. 35, 101608 (2020).

[65] Colson, P. et al. Children account for a small proportion of diagnoses of SARS-CoV-2 infection and do not exhibit greater viral loads than adults. Eur. J. Clin. Microbiol. Infect. Dis. 26, 15 (2020).

[66] Song, H. et al. A considerable proportion of individuals with asymptomatic SARS-CoV-2 infection in Tibetan population. medRxiv https://doi.org/10.1101/2020.03.27.20043836 (2020).

[67] Dong, Y. et al. Epidemiology of COVID-19 among children in China. Pediatrics 145, e20200702 (2020).

[68] Dong, E., Du, H. and Gardner, L., An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 20, 533534 (2020).

[69] Bar-On, Y. M., Flamholz, A., Phillips, R. and Milo, R., SARS-CoV-2 (COVID-19) by the numbers. eLife 9, e57309 (2020).

[70] Leyva, I. et al., Explosive first-order transition to synchrony in networked chaotic oscillators, Phys. Rev. Lett. 108 (16), 168702 (2012).

[71] Zhang, X., Hu, X., Kurths, J. and Liu, Z., Explosive synchronization in a general complex network, Phys. Rev. E 88, 010802(R) (2013).

[72] Zhang, X., Boccaletti, S., Guan, S. and Liu, Z., Explosive synchronization in adaptive and multilayer networks, Phys. Rev. Lett. 114, 038701 (2015).

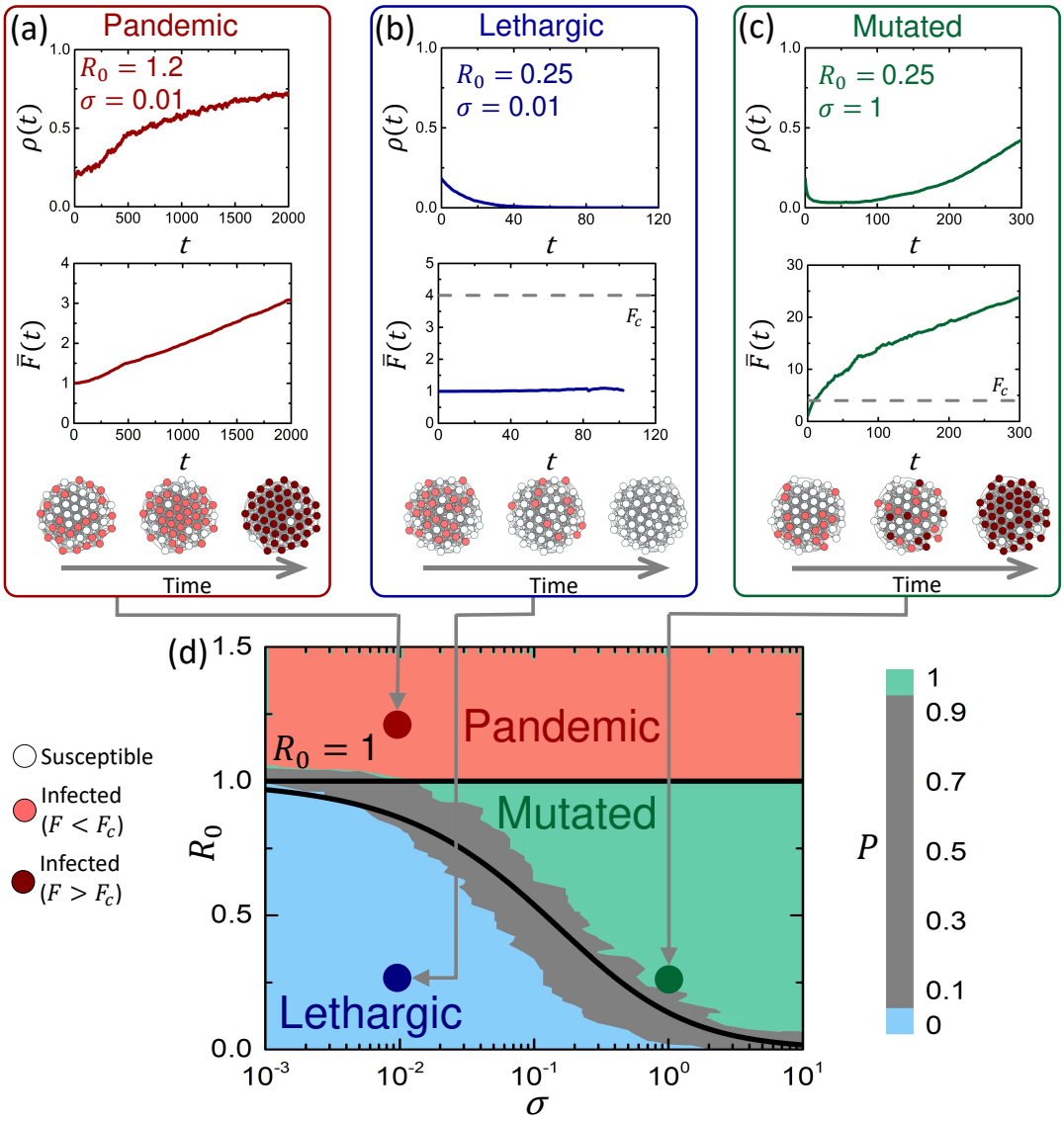

Figure 1: **The phases of a pandemic under pathogen mutation**. (a) *Pandemic phase*. For $R_0 > 1$ we observe the classic pandemic phase. The prevalence $\rho(t)$ vs. t (top) grows continuously as the fitness $\overline{F}(t)$ (bottom) increases due to mutation and natural selection. (b) *Lethargic phase*. For $R_0 < 1$ we have $\rho(t)$ exponentially decaying to zero. The mutation rate $\sigma = 0.01$ is too slow, $\overline{F}(t)$ remains almost constant (bottom), and the pathogen fails to reach critical fitness F_c (grey dashed line) on time. (c) *Mutated phase*. We now remain in the sub-pandemic regime $R_0 < 1$, but increased the mutation rate to $\sigma = 1$. For small t we observe $\rho(t)$ rapidly decaying (top). However, thanks to the rapid mutations $\overline{F}(t)$ reaches critical fitness (grey dashed line) within a short time. Following this point the disease reemerges and $\rho(t)$ changes course, turning pandemic. This is observed in the snapshots at bottom through the appearance of sporadic instances of high fitness pathogens (middle, dark red nodes), which then spread to infect the majority of the population. (d) σ, R_0 *phase diagram*. To systematically observe the different phases we varied $R_0 \in (0, 1.5)$ and $\sigma \in (10^{-3}, 10)$, capturing a total of $1,050$ epidemiological scenarios, with different μ, β and σ. For each scenario we ran 50 stochastic realizations and measured the probability P to have $\rho(t \to \infty) > 0$, *i.e.* pandemic. We observe three phases with sharp boundaries between them. First, the pandemic phase (red) for $R_0 > 1$, independent of σ, as predicted by the classic SIS model. In addition to that the sub-pandemic regime $R_0 < 1$ is split into two phases: Under small σ, P tends to zero (blue) and the pathogen fails to spread, giving rise to the lethargic phase. For large σ, the spreading probability becomes almost certain, as $p \sim 1$ (green), and we observe a mutation driven contagion. The gap between these phases (grey) indicates an abrupt transition from $P \to 0$ to $P \to 1$, a dramatic shift occurring within a narrow range of R_0, σ values. This grey range is well-approximated by our theoretical prediction (solid black line) as appears in Eq. (11). All simulations, here and throughout, were done on a random network of $N = 5,000$ nodes and $\overline{k} = 15$. The disease parameters were set to $\mu = 0.1$ and the infection rate was set variably to $\beta = \mu R_0/\overline{k}$, to obtain the different values of R_0. The mutation rate σ is specified in each figure. In each scenario we set the initial condition to $\rho(t) = 0.2$.

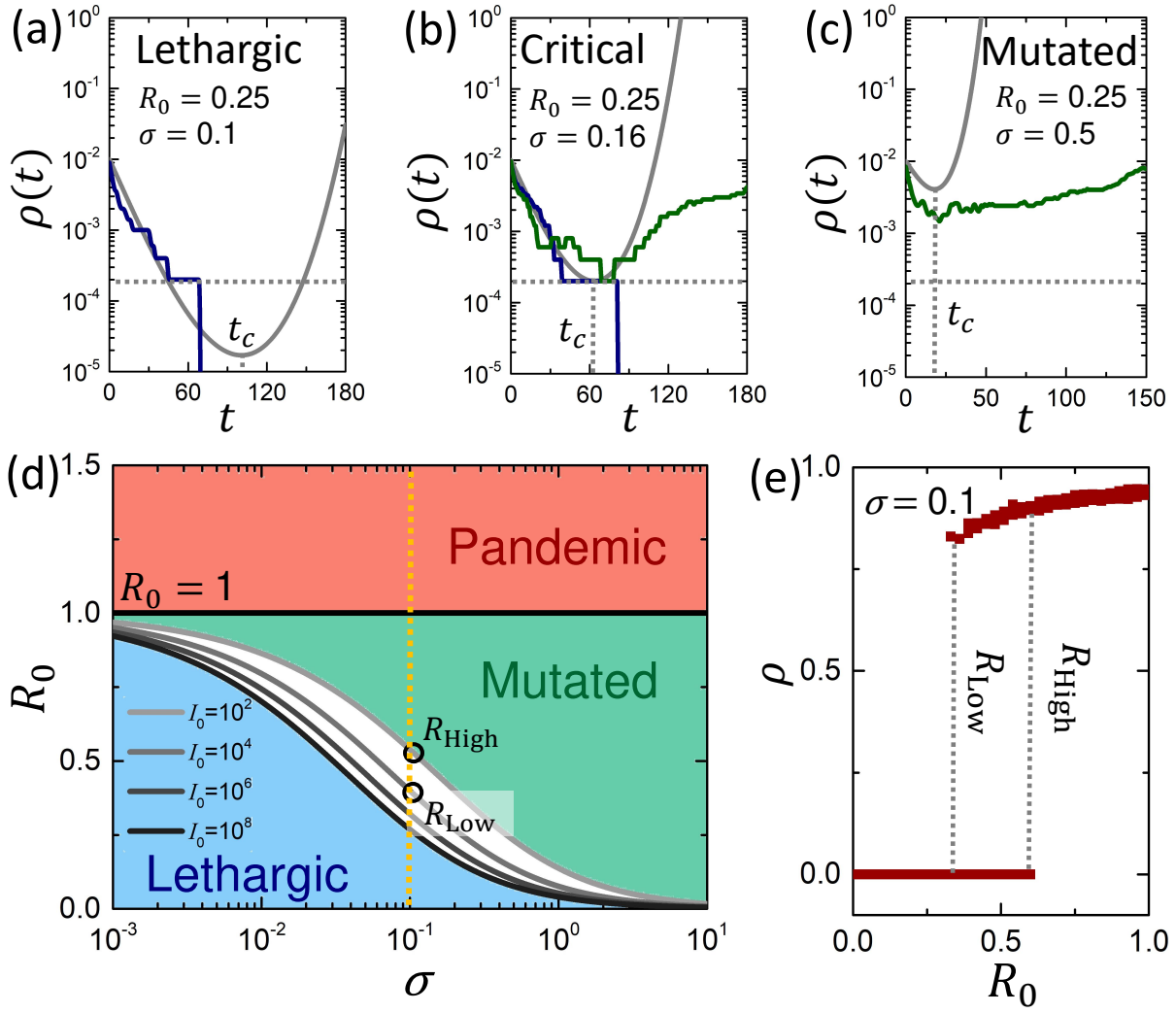

Figure 2: **The transition to the mutated phase**. To observe a mutated phase a critical mutation must arise before the pathogen is eliminated, namely before $\rho(t)$ crosses $1/N$ (grey dashed lines), capturing the *unit line* in which there is a single infected individual among the N node population. (a) $\rho(t)$ vs. t (grey solid line) as obtained from Eq. (8) in the lethargic phase ($R_0 = 0.25, \sigma = 0.01$). The critical mutation occurs at the minimum point (t_c), which is below the unit line. Therefore the epidemic decays prior to the appearance of the critical mutation. Indeed, the stochastic simulation (blue solid line) approaches zero prevalence, never reaching the positive branch of $\rho(t)$. (b) Setting $\sigma = 0.16$ the system is at criticality. $\rho(t_c)$ is adjacent to the unit line, and hence we observe critical behavior: some realizations decay (blue), whereas others successfully mutate (green). (c) Under $\sigma = 0.5$, the system is in the mutated phase, $\rho(t_c)$ is sufficiently above the unit line and the critical mutation is reached with probability $P \to 1$. (d) The lethargic-mutated phase boundary in Eq. (11) depends on the initial size of the infected population $\mathcal{I}_0$. Here we show this boundary for $\mathcal{I}_0 = 10^2, \ldots, 10^8$ (grey solid lines). (e) The long term prevalence $\rho = \rho(t \to \infty)$ vs. R_0 under $\sigma = 0.1$ (yellow dashed path in panel (d)). Approaching from small R_0 (left to right) we begin with an initial infection of $\mathcal{I}_0 = 10^2$ and observe a transition to the mutated phase at $R_0 = R_{\rm High}$. In the opposite direction, however, as we begin with large R_0 we approach the transition from an already pandemic state with $\mathcal{I}_0 \sim 10^4$. Now the phase boundary traverses through $R_0 = R_{\rm Low}$. Both transitions are also marked by circles in panel (d). We, therefore arrive at a hysteresis phenomenon, in which the critical transition point depends on the current state of the spread. Consequently, preemptive mitigation, done when the spread is still at its embryonic stage ($\mathcal{I}_0$ small), is more effective than reactive mitigation, applied when $\mathcal{I}_0$ is already large.

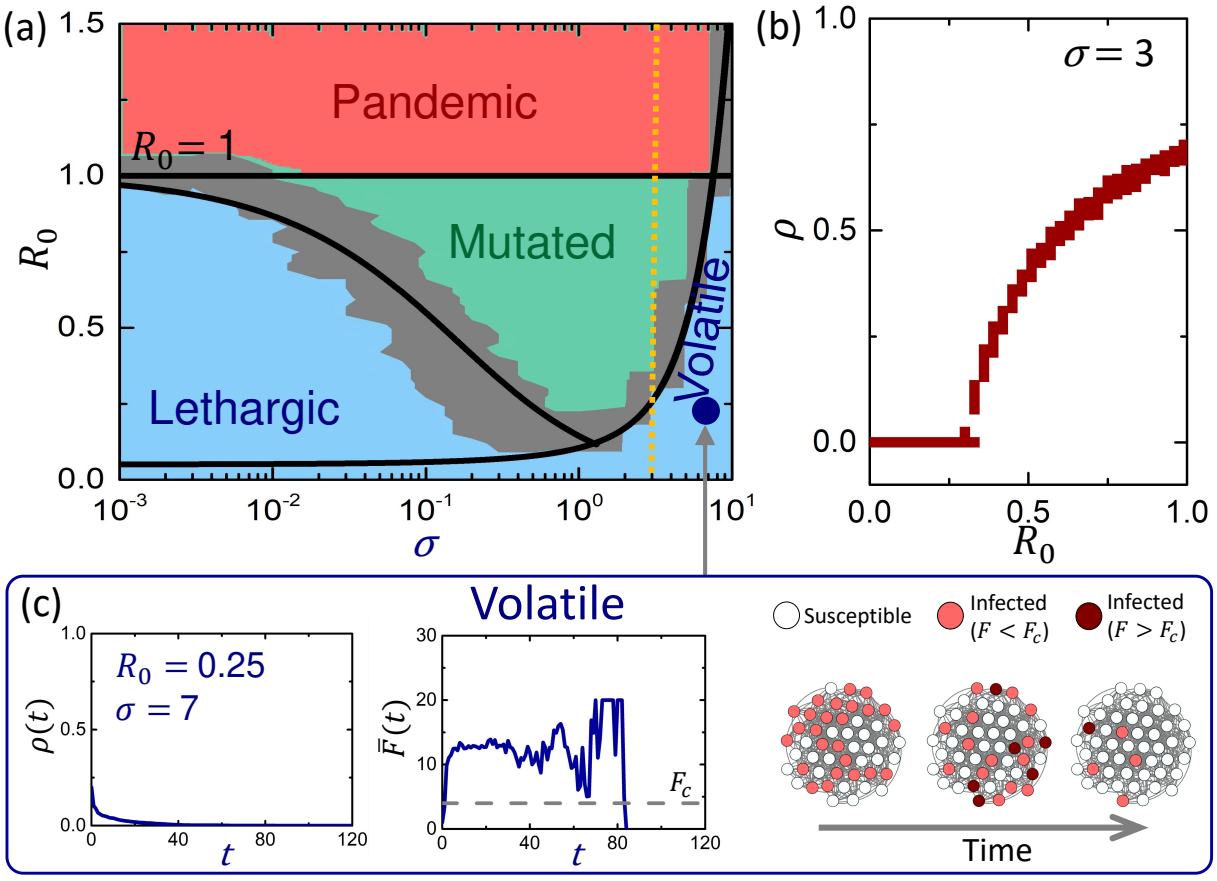

Figure 3: **The volatile phase**. (a) The σ, R_0 phase diagram under the bounded fitness of Eq. (12). We now observe a volatile phase, in which $\rho \to 0$ (blue), when σ is too large. Hence, the mutated phase (green) now only appears in the *Goldilocks zone* in which the mutation rate in not too high nor too low. The theoretical prediction of (13) is also shown (black solid line on right). (b) ρ vs. R_0 under $\sigma = 3$ (yellow path in panel (a)). As opposed to the lethargic-mutated phase transition, the shift from volatile to mutated follows a continuous second order transition. (c) $\rho(t)$ vs. t in the volatile phase decaying, as predicted, to the healthy state $\rho = 0$. (d) $\overline{F}(t)$ changes rapidly thanks to the large σ, and crosses the critical F_c (grey dashed line) early on. However the rapid mutations prevent the slower natural selection from securing a steady increase in $\overline{F}(t)$. Hence, the achieved fitness cannot be stably sustained for the pathogen to continually spread. (e) Indeed, we observe multiple instances of critical fitness (dark red) that fail to reproduce and dominate the pathogen population.

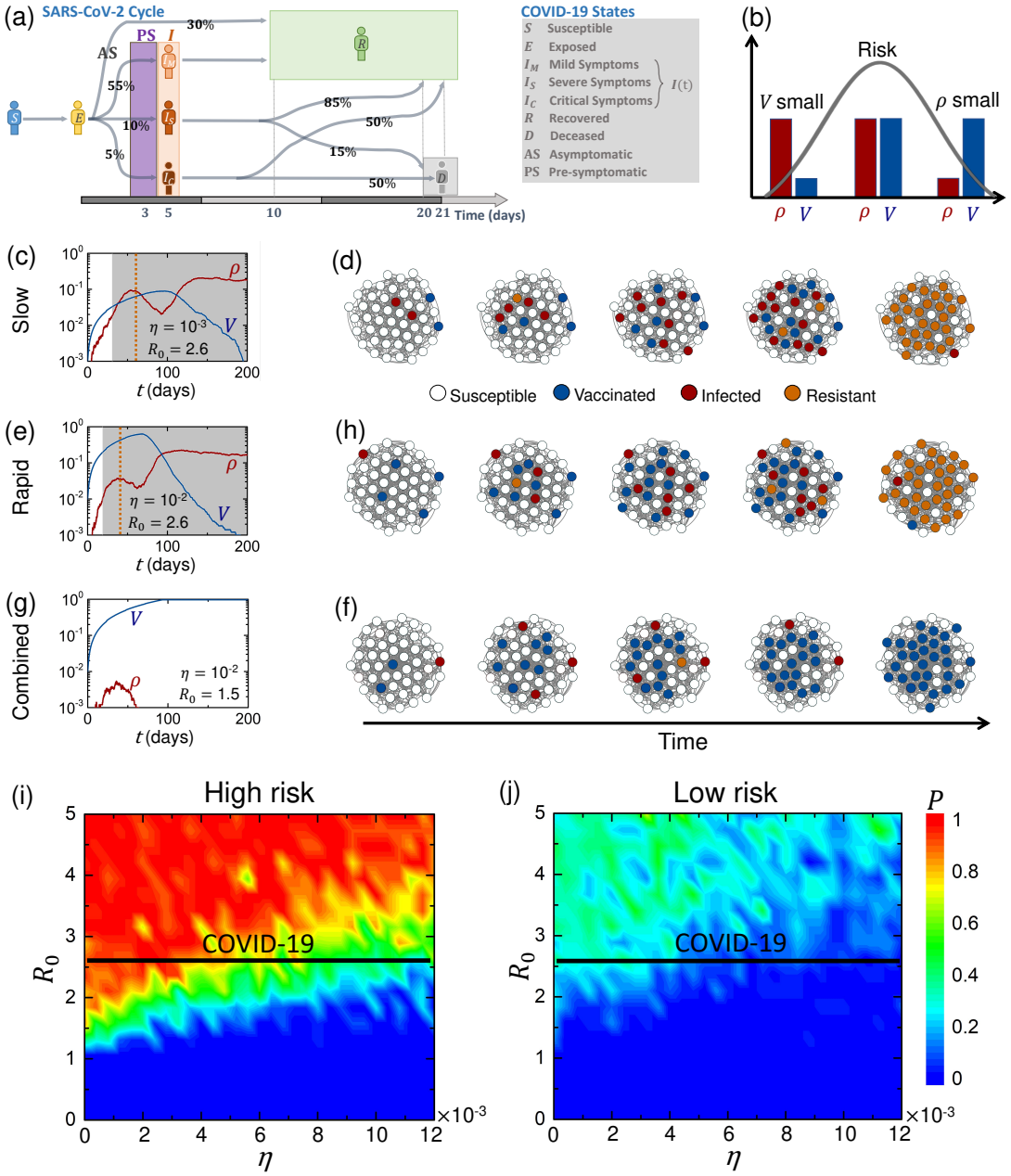

Figure 4: **Vaccination under the threat of mutation**. (a) The SARS-CoV-2 disease cycle. Upon exposure (yellow) individuals enter a pre-symptomatic phase (purple), from which they later develop mild (I_M), severe (I_S) or critical symptoms (I_C), determining the duration of their infected phase and their probability to recover (green) or decease (grey). (b) Vaccine resistance is risky under a coexistence of both infected (ρ) and vaccinated (V) individuals (center). When ρ is small, the probability of mutation is marginal (right); when (V) is small the selection pressure for resistence is weak (left). (c) Under slow vaccination $\rho(t)$ increases (red). As a result, when vaccines gain coverage we enter the risky zone (shaded), and become potentially vulnerable to resistance mutation. Indeed, when such mutation occurs (orange line), the trend is reversed, $\rho(t)$ increases and the vaccine coverage $V(t)$ plummets (blue). (d) We present several snapshots to track the state of the spread. In snapshot 2 we observe a *premature* mutation (orange node) that fails to spread, since $V(t)$ at that point is still small (blue nodes). Later (snapshot 4), with the system in the risky zone of high $\rho(t)$ and $V(t)$, such mutations rapidly take over, as seen by the coverage of the orange nodes in snapshot 5. (e) - (f) Rapid vaccination in and of itself may be insufficient. The system quickly enters the risky zone (shaded) and with $R_0 > 1$, a single resistance mutation eventually outruns our vaccination efforts. (g) - (h) Successful eradication of the disease is achieved under a combination of rapid vaccination (blue) and suppression of R_0, *e.g.*, through social distancing. Pushing R_0 down suppresses $\rho(t)$, and hence avoids the risky zone by locating the system in the right hand side of panel (b). (i) The probability P to observe a pandemic state as a function of the vaccination rate η for different values of R_0. To alleviate the risk of vaccine resistant spread we must remain in the blue zone, in which we not only invest in the vaccine roll-out (η), but also in suppressing the spread (reducing R_0). (j) Similar, albeit less dramatic results are also onserved under our low risk scenario $\mathcal{P} = 0.25$.

Figures

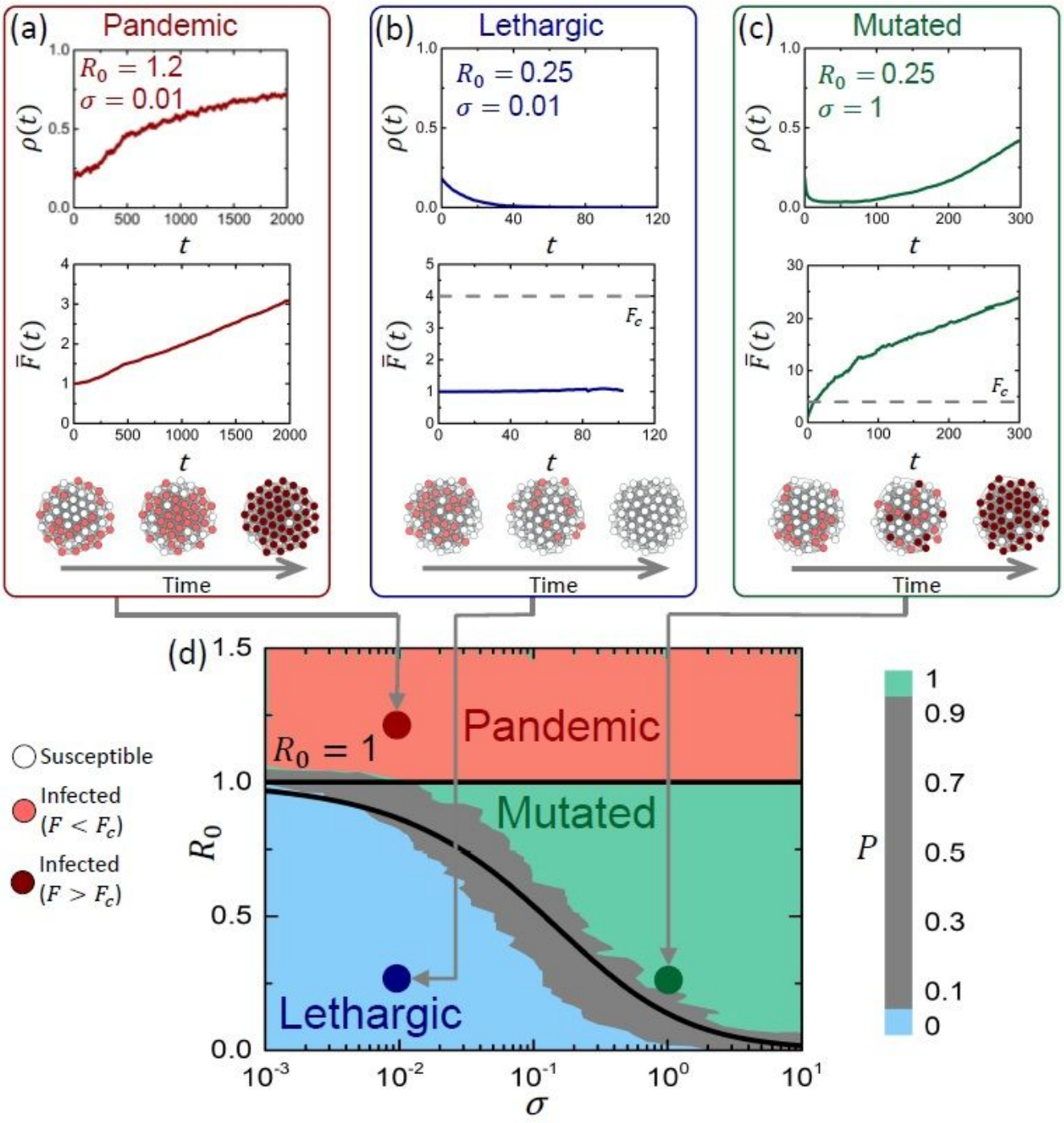

Figure 1

The phases of a pandemic under pathogen mutation. (see Manuscript file for full figure caption)

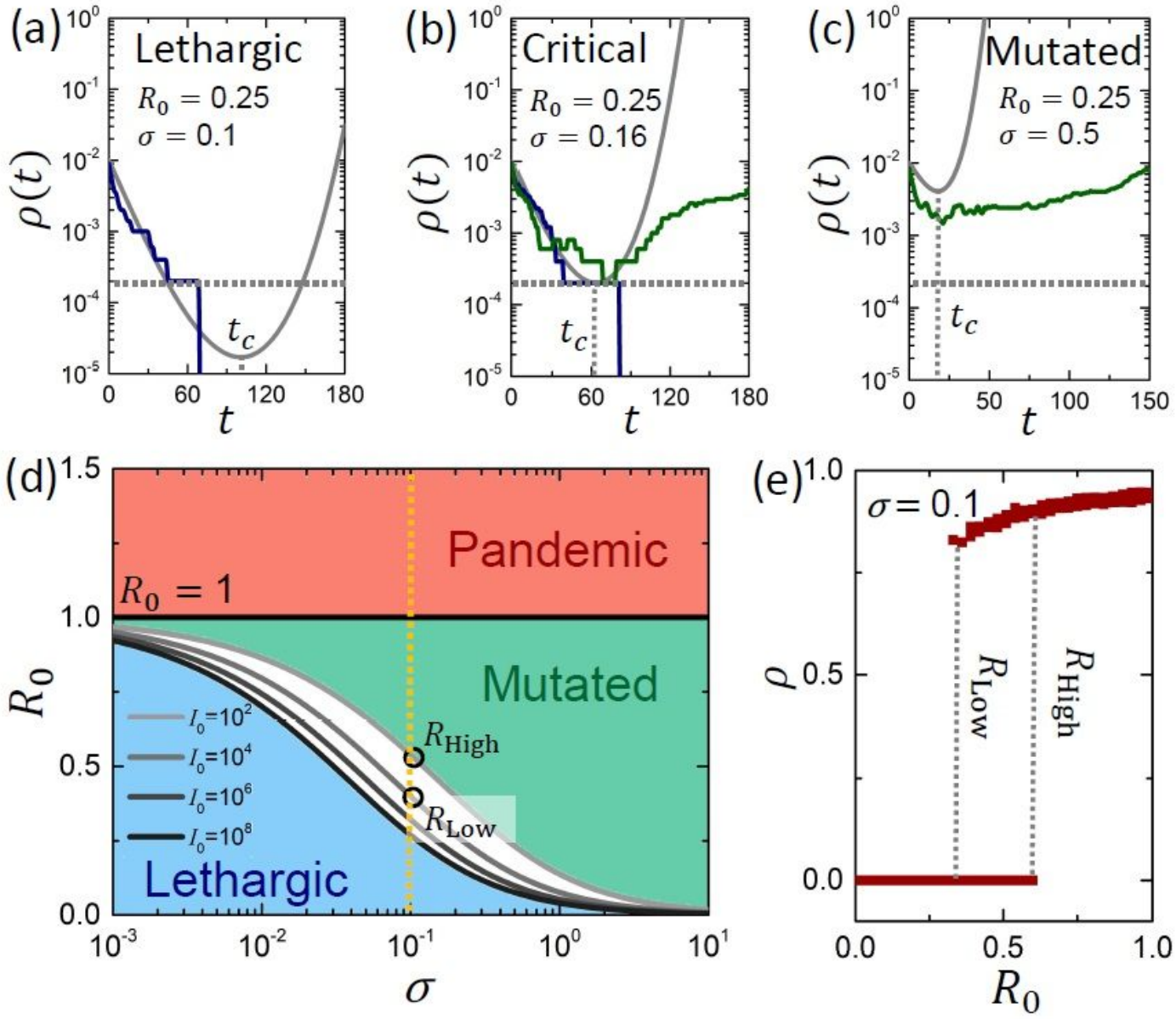

Figure 2

The transition to the mutated phase. (see Manuscript file for full figure caption)

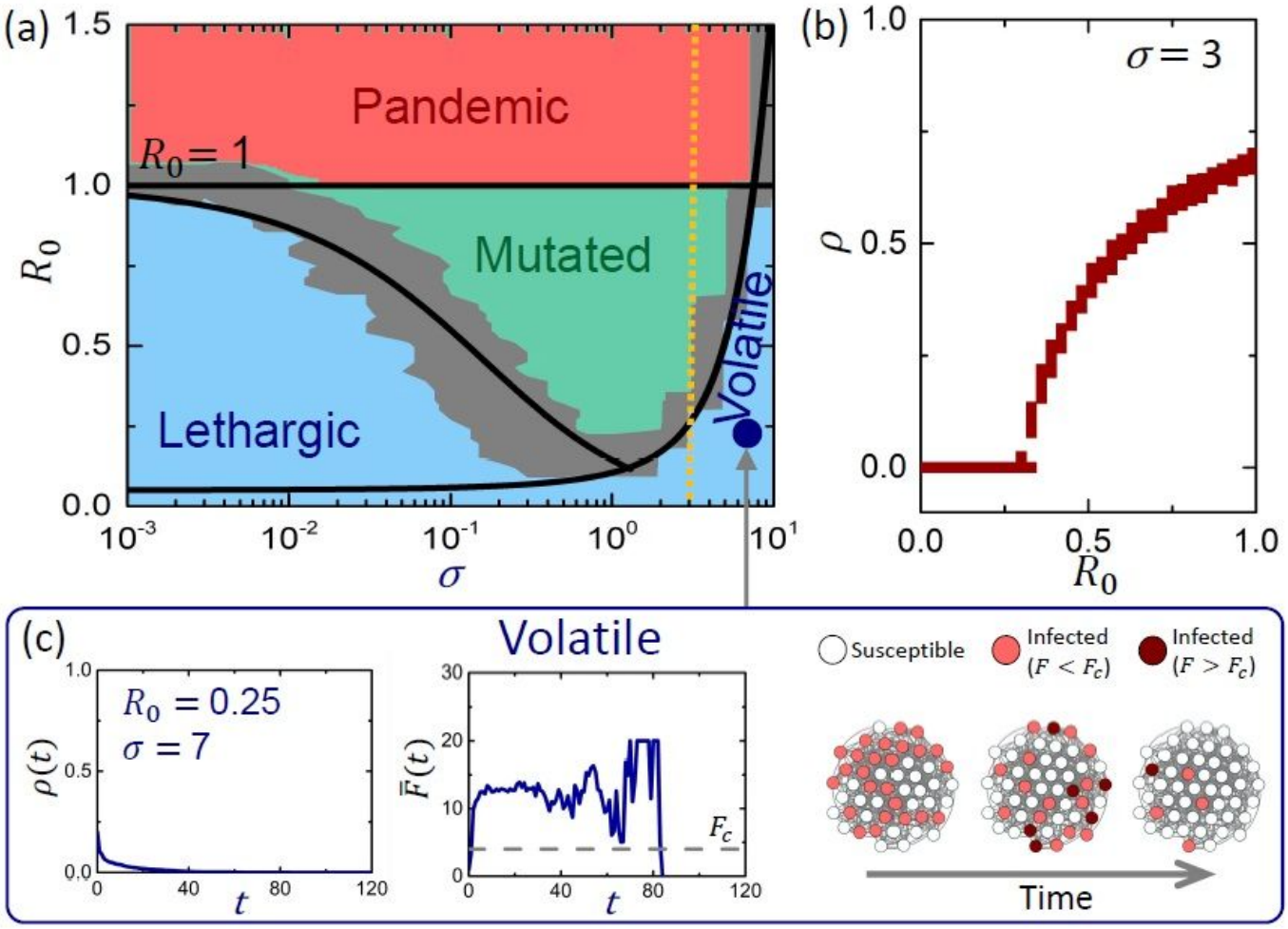

Figure 3

The volatile phase. (see Manuscript file for full figure caption)

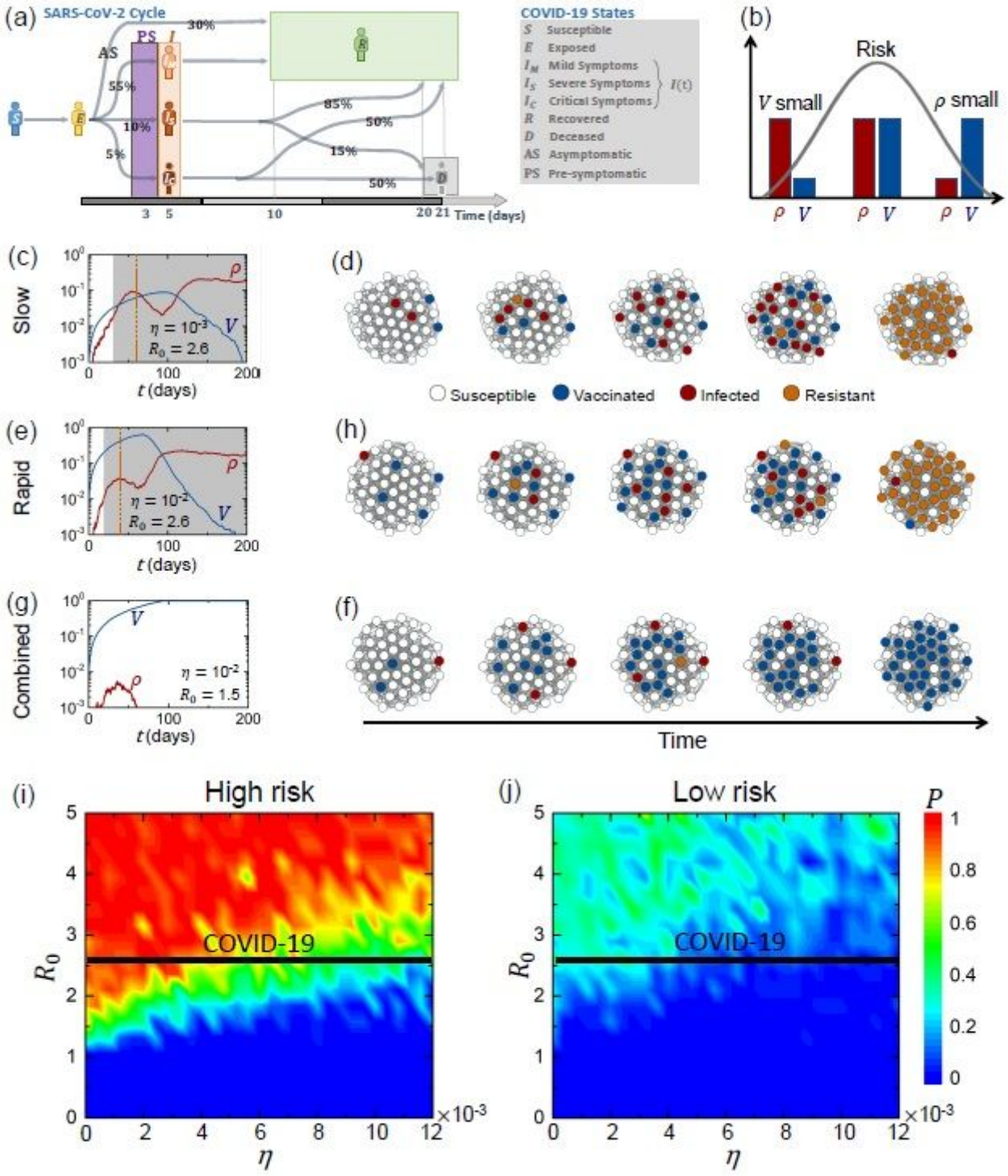

Figure 4

Vaccination under the threat of mutation. (see Manuscript file for full figure caption)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf