Evolutionary Effects of Contagious Diseases

An analysis based on the mathematical Susceptible-Infected-Susceptible (SIS) model

Author: Henk-Jan van der Molen – 2025-04-12

Abstract

Background: according to Darwin's general principle of *the survival of the fittest*, the most adaptive species have the best chance off survival and should become dominant. As a result, the population becomes genetically more uniform. However, this principle has led to a paradox, because it seems incompatible with the biodiversity we see everywhere in the food chain. To overcome this paradox, the "Kill the Winner" (KtW) hypothesis was proposed. KtW should prevent a winner from emerging and thus maintains the coexistence of all species in the system.

The objective of this paper is to strengthen the understanding of the evolution *by* contagious diseases. More specifically, the focus is on the evolutionary link between contagious diseases and the diversity of host species, by evaluating the effect of host species diversity on the impact of epidemics. The question if species diversity will reach an equilibrium is addressed in the conclusion.

When pathogens multiply and mutate faster than their host organisms, on an evolutionary time scale host populations will endure epidemics, with possible high mortality.
The Herd Immunity Threshold (HIT) determines the impact of an epidemic on a host population. Contagious diseases can be an actor in the KtW hypothesis, if the HIT for epidemics significantly reduces when a host population genetically becomes more diverse. A mathematical model was analyzed to determine the effect of host diversity on the impact of epidemics, the results were verified with simulations.

Results: The analysis revealed that when the difference in susceptibility to a disease between two subpopulations increases, the HIT and therefore, the impact of an epidemic reduces. Populations with dominant species suffer more from epidemics.

Conclusions: based on this analysis, contagious diseases could serve as an actor in the KtW hypothesis.

Keywords: contagious diseases, SIS model, diversity, evolution, herd immunity, KtW.

Background

Darwin and biodiversity

The following quote on evolution is attributed to Darwin: "*It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change*". Within the context of the arms race between predator and prey species, this "*survival of the fittest*" principle suggests that the most adaptive species will become dominant, because the fittest species have the best chance to pass their genes to subsequent generations. This should increase the population size of the fittest species and decrease the population size of less fit species, leading to less biodiversity.

However, everywhere in the food chain there is a rich biodiversity of competing species with comparable population sizes, which seems incompatible with this explanation. Based on this view Hutchinson presented in 1961 the paradox of the plankton: why do so many coexisting species feed on the same nutrients, instead of one species out-competing all the others [1]? This latter expectation has been formulated as the so-called competitive exclusion principle [2]. The plankton paradox has been generalized as the Biodiversity Paradox, which has not yet been completely resolved [3].

The "Kill the Winner" hypothesis (KtW) was proposed to solve this paradox. KtW assumes that host specific predators reduce the population size of dominant species, so more species in the system can coexist [4].
The context of the KtW hypothesis is usually the Predator-Prey model relating to viruses that infect bacteria in marine ecosystems, but in its general form KtW is also applicable to terrestrial systems [5].
In addition to predation, the survival of a species also depends on environmental circumstances and evolutionary events that could lead to extinction. Recent research has made progress reconciling the biodiversity paradox with Darwin's principle, stating that populations with a large additive genetic variance are better able to adapt to environmental changes [6].

Prior research has simulated epidemics in host populations with 1, 2, 10 or 100 genotypes [7]. These researchers used the susceptible-infected-recovered (SIR) model to calculate the total number of infected individuals in each subpopulation as a measure of the impact of epidemics. The conclusion of the

researchers is that on average, epidemics in more heterogeneous populations are less likely to become catastrophic. Increased heterogeneity is associated with reduced mortality in epidemics and a shorter duration of epidemics. More recent research has used the related SVEIR (Vaccinated, Exposed) model to determine the effect of vaccination in heterogeneous populations. These researchers concluded that the Herd Immunity Threshold (HIT) values for heterogeneous populations are always significantly lower than those for corresponding homogeneous populations, which decreases the impact of epidemics [8].

Evolution due to ecology factors

On the survival of species, Robin Dunbar stated [9]: *"Species change through the gradual failure of some lineages to reproduce, resulting in a subtle but steady drift in the species' genetic make-up toward that of lineages that are more successful. Although in most cases these processes are quite slow, an entire species can go extinct catastrophically if none of its various lineages can reproduce fast enough to offset unusually high levels of mortality. There is always a steady trickle of such extinctions over time – there have been literally dozens within our own lineage during the course of our six-million-year evolutionary history. Sometimes, however, environmental conditions conspire to produce a rapid burst of extinctions".*

The MacArthur–Wilson theory of island biogeography also provides insights into the extinction of species [10]. According to this theory, the area of the island is one of the key factors that determines how many different species the island can support. One of the arguments that smaller islands have less biodiversity is that smaller populations are more prone to go extinct. This theory states that places with high diversity have a small influx of new species and vice versa. Research has shown that the more genetically diverse a microbial community becomes, the more likely antagonistic interactions are [11]. Another study concluded that there is an inverse relationship between soil microbial diversity and the survival of invading species [12]. The theory of island biogeography has inspired much criticism, but research has indeed provided evidence that links island size and the stable number of species on that island. Furthermore, the assumption relating population size to extinction likelihood is very well supported by data [13]. This indicates that large monocultures have a better chance to survive, because populations of species less successful in competition are (much) smaller and therefore, more likely to go extinct.

Generally, an organism can use available resources to increase growth, for reproduction or to counteract stress, e.g. predators, parasites, decay, unfavorable environmental conditions, or competitive ability [47]. The standard mathematical Lotka–Volterra predator–prey model consists of two differential equations to describe the change over time in the size of both the predator and prey populations connected in a food chain [14]. Each species will experience evolutionary pressure, because they must adapt to successful mutations of its predators, competitors and prey in the food chain network. In a study determining the cause of the high diversity in existing ecosystems, the mechanism of coevolution of predator-prey species was shown to maintain stable diversity [3]. Specifically, a significant inflow of genetically different members is necessary to avoid extinction of the population. This study introduces mutation to create genetic diversity, using the same rate of mutation for both predator and prey.

Evolution due to pathogen–host relations

Disease ecology and macro ecology share many similarities [15]. Within the context of the infection process, diseases act as parasitic predators in infected hosts preying on susceptible species. When adapting these variables in the predator-prey model accordingly, the predator-formula becomes the formula of the Susceptible-Infected-Susceptible (SIS) model, which describes the spread of contagious diseases. Despite these similarities, a pathogen as predator differs substantially from its host species. For example, the rate of mutation for each genome is roughly the same (less than 1 mutation per billion copies [16]), but a bacterium or virus reproduces much faster than multi-cell host organisms. Many viruses accumulate 10-10,000 more mutations per generation than cells do, because their reproduction process lack editing functions [17].

Pathogens can rapidly evolve when environmental conditions change which are related to their survival. Phages as bacterial predators exchange genetic material with each other and bacteria and thus stimulate diversity of bacterial species and contribute to accelerated evolution [18]. Other studies have demonstrated that vaccines cause pathogens to evolve [19, 20, 21, 22, 23]. Moreover, novel adaptations can significantly change the genome of pathogens and lead to rapid evolution [24]. This relationship works both ways: in areas where malaria is endemic, the frequency of mutations in red blood cells remains high [25]. Nature is very flexible in exploring new possibilities, even multi-year periodicity in epidemics occurs [26].

Research showed that tight (i.e. inflexible) interactions between microbial species could lead to ecological disasters [27, 28]. A study in 2012 revealed that chronic sinusitis patients have reduced bacterial diversity (dysbiosis) and that *Lactobacillus Sakei* cures some forms of sinusitis [29]. In 2013, a study suggested that

inflammatory diseases reduce the ecological diversity of the human gut micro-biome [30]. A study in 2014 on inflammatory bowel disease concluded that it is not yet clear whether dysbiosis is a cause or an effect of this disease [31]. An overview of metagenomic studies concluded in 2016 that science is still at the beginning of this research and that current methods are far from perfect. A better understanding of the effect of changes in microbiota depends on the development of better methods and techniques [32]. Another study on chronic inflammatory lung diseases in 2018 revealed that information on host-microbiota interactions is still very limited, so the causes and effects have not yet been determined [33]. A study in 2018 suggested that Crohn's disease causes dysbiosis [34]. Research in 2020 and 2021 has revealed that medical treatment sometimes causes dysbiosis, but that due to differences in human lifestyle and genetics, dysbiosis is not always associated with autoimmune, inflammatory or pernicious diseases [35, 36]. Thus, a steady balance between various pathogens appears to be an indicator of stability rather than a stabilizing factor.

A study divided pathogens into three classes based on their antigenic variation, which enables them to circumvent host immunity systems [37]. In Table 1 Class I is typical for STDs with a low basic reproduction number (R_0) and a corresponding low HIT for epidemics. The characteristics of class II are typical of flu-like diseases with high antigenic variation and a low R_0, which translates to a low HIT for epidemics. For instance, the influenza A virus (IAV) is very successful to evade the host immune system and is worldwide able to infect 350 – 1750 million people each year [38]. Against a world population of 7000 million the HIT of IAV is between 5 – 25%. Pathogens in class III are characterized by a high R_0 and low antigenic variation, so vaccines provide lifelong protection.

Infection	Class	Antigenic variation	R_0	HIT(%)
STD: HPV [39]	I	High	0.5	N/A
Influenza H1N1 [40]	II	High	1.3 - 2	23 - 50
SARS [41]	II	High	2 - 4	50 - 75
Mumps [42]	III	Low	4 - 7	75 - 86
Polio [10]	III	Low	5 - 7	80 - 86
Rubella [10]	III	Low	6 - 7	83 - 86
Pertussis [10]	III	Low	12 - 17	92 - 94
Measles [10]	III	Low	12 - 18	83 - 94

Table 1: Classes of contagious diseases with antigenic variation, basic reproduction number (R_0) and corresponding Herd Immunity Threshold (HIT)

Using their genetic flexibility, contagious diseases in class I and II are able to exploit new genetic susceptibilities to infect species. Some contagious diseases have a combined effect, because the immune system can be impacted by previous infections. Research supported the hypothesis that infection with measles decreases resistance to other infectious diseases for up to 5 years [43].
Because pathogens tend to specialize in specific host species, they are more likely to reduce the populations of common species and therefore maintain diversity in host species [44]. Because virulent diseases are by far the most effective eradicators of species [45], contagious diseases have an effect on the evolution of species, independent of the regular predator-prey interaction, shown in Figure 1 [46].

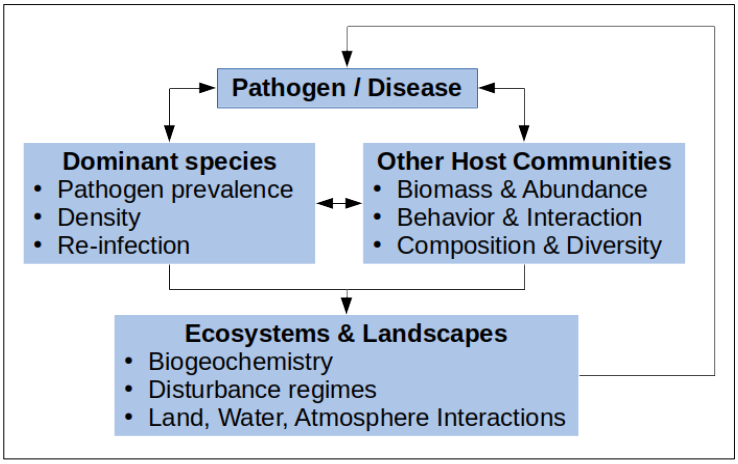

Figure 1. Framework Relating Pathogens to Host Communities and Ecosystems

Resistance to contagious diseases has evolutionary benefits for a species, but on the long term it seems that higher lifeforms cannot avoid the risk of epidemics, no matter how fit they are compared to their competitors and predators in the food chain [47].

Disease resilience in species

Many of the species concepts and their associated definitions used in research are incompatible and can lead to different conclusions concerning the boundaries and numbers of species [4]. A species in this context is defined as a collection of individuals with a set of uniform characteristics, such as the susceptibility for diseases.
Disease resilience, defined as an animal's ability to maintain productive performance in the face of infection [48], is essential for coping with the continuously changing polymicrobial challenges. Most pathogens show plasticity and are able to adapt to every intervention, as the increasing antibiotic resistance of bacteria has demonstrated. The arms race between pathogens and susceptible species is asymmetrical: a pathogen survives if it can exploit one susceptibility in any infected species; a susceptible individual can only survive if it can recover from all infections with the various pathogens in its environment. Given enough time, any population will come in contact with a pathogen that will trigger an epidemic.

Research indicates that resilience against diseases has a low heritability, even at intermediate prevalence [49]. For a species, it is costly to change its genome, because it requires considerable investment throughout development (of disease resilience) in response to feedback from the environment [50]. Other research has associated the development of resistance with a significant fitness penalty for resistant populations [47]. This suggests that the evolutionary costs of implementing resilience in the genome for diseases are usually greater than the benefits, especially if pathogens can evolve faster then susceptible species can become resilient. Evolution seems to value genetic diversity more as a measure to counter diseases in host species. It is also known that viral infections can be a source of genetic diversity, because the genetic material of viruses and their hosts can be mixed [51]. For both hosts and viruses this can lead to acquisition of new functional genes that can be beneficial for the host and for viral fitness.

Research shows that genetic variation decreases the HIT for a disease, which provides positive feedback that can lead to eradication of the disease [52]. Thus, genetic variation offers much potential to reduce the prevalence of infectious diseases. However, several studies have shown that genetic variation in disease resilience is difficult to define, understand and locate, because it encompasses a variety of traits linked to hosts, pathogens and the environment [48, 49, 53, 54].

Defining host variation in infectivity by two distinct homogeneous subgroups, research has shown that genetic variation can greatly reduce the probability of disease emergence [55]. Another study predicts similar results when traits like host susceptibility, recovery and infectivity are controlled by multiple genes [56].

Methods

The analysis in this article is based on the Susceptible-Infected-Susceptible (SIS) model, which predicts how contagious diseases spread in a population. The SIS model assumes that each individual (node) in the network is in one of two binary states: healthy (and susceptible to infection), or infected. After infection, the node immediately spreads the disease to other, susceptible nodes. Infected nodes can 'heal', but then immediately become susceptible to the same disease. Therefore, the SIS model is only fit for diseases such as tuberculosis or flu-like diseases that do not create permanent immunity after recovery. The transfer of infections between nodes follows a Poisson process, so infections are independent of each other.

The underlying differential equation of the SIS-model requires three parameters: the chance that an infected individual will transfer its infection to a connected susceptible individual (β), the fixed number of contacts a node has (k), and the chance that an infected individual will "cure" (δ).

The basic reproduction number ($R_0 = \beta.k/\delta$) predicts, on average, how many susceptible individuals an infected individual will infect. If the R_0 of a population for a contagious disease is more than unity, the SIS model predicts an epidemic. Calculations can be simplified using R_0, because the steady state of the mathematical model only depends on the value of R_0. The infected fraction of the SIS model in steady state is used as a measure for the impact of epidemics, which is equal to the Herd Immunity Threshold (HIT).

The topology structure considered is the regular graph in which each node has the same amount (k) of connections with other nodes. The network must also be completely symmetrical: each node must link to equally many nodes in each subpopulation. When a graph complies to this requirements, the steady state of the SIS model can be calculated for populations with more than one genotype, even when the system of differential equations cannot be solved [57].

A population is defined as a collection of species (or subpopulations) with individuals which are a member of one species. All species in the population are susceptible as hosts to contagious diseases, which will lead to epidemics on an evolutionary timescale. The basic hypothesis is that the genetic diversity within the host population helps reduce the impact of epidemics.

The mathematical SIS-model was analyzed to determine the effect of host diversity on the impact of epidemics. In this context, diversity is defined as the weighted difference between the R_0 factors of two subpopulations (R_1, R_2). Genetic drift can redistribute genetic factors related to susceptibility of contagious diseases, which influences the diversity of the population. The impact of epidemics on (more or less diversified) populations can be determined, if the R_0 of the whole population remains constant. If varying diversity changes the R_0 of the population, the steady state of the model would follow, which corrupts the analysis.

In Annex A the effect of diversity on the impact of epidemics was calculated for homogeneous populations (i.e. one R_0 value for the whole population) and heterogeneous populations (each subpopulation (i) can have a different value for R_i, while the R_0 value for the whole population remains constant).
The Sagemath software was used to calculate the derivative of the steady state function of a population with two equally sized and static subpopulations.

The calculated effect of diversity on the impact of epidemics was verified with simulations. Compared to the mathematical model, the simulation model requires additional parameters.
For instance, all the simulations were based on a population with two equally sized and static subpopulations of 2500 nodes each. Each node has an equal number of connections (k), with half of these connections directed to nodes in each subpopulation.
The simulated networks are *undirected*, because edges must allow transmission of infections between any linked pair of individuals. An example of an undirected symmetrical network: node (0) is connected to nodes (2498, 2499, 1, 2) in both subpopulations, node (1) links to nodes (2499, 0, 2, 3), and so on. Thus all links in an undirected network are bidirectional, but the infection rate (β_1, β_2) in each direction may be different.
Note that *directed* symmetrical networks can also be simulated, but it is unrealistic to assume that the spread of pathogens is constrained by edges that allow transmission of infections in one direction only.
To initialize the simulation, 1% randomly chosen nodes are infected (node value => 1), and the remaining 99% of the nodes are susceptible (value = 0). It is easy to verify that the initially infected fraction (1% = 50 nodes) is sufficient to trigger an epidemic. Simulations that start with 100% of the population infected, lead to a comparable mean and standard deviation of the steady-state value.

In a single time step the condition of each node is evaluated within its context. In every time step an infected node has a chance δ to “cure” and become susceptible again. The differential equation of the SIS model states that a susceptible node becomes infected with chance $\beta*(k*v)$. In simulations the component ($k*v$) is defined as: how many of the k contacts of that particular node are infected.
These rules are straightforward to implement using a random sample ($\underline{u}$) drawn from the uniform [0..1] distribution. An infected node is cured when $u \leq \delta$, a susceptible node is infected if $u \leq \beta*(k*v)$.

A simulation regularly includes 2000 *linear* time steps, because the analysis only focuses on the steady state of the epidemic. To reach the steady state as fast as possible, β and δ are maximized under the constraint that the chances $\beta.k$ and δ must be ≤ 1. The steady state value was calculated as the mean value of the function (v) from time steps [1000 – 2000], averaged over 50 simulation runs to reduce stochastic noise. The Python simulation script is listed in Annex B.

Results & Discussion

Diversity in the homogeneous SIS Model

The simplest form of the SIS model has one constant and uniform value for R_0 for the whole population. Assuming $R_0 > 1$, solving the differential equation of the SIS model leads to the characteristic S-shaped logistic function, showing how the disease spreads over time. The homogeneous SIS model only permits situations where subpopulation #1 is susceptible and the other subpopulation #2 is not. The size of subpopulation #1 is a fraction n_1 of the population, with a basic reproduction number $R_1 > 0$. Subpopulation #2 is the remaining fraction n_2, with $R_2 = 0$. Diversity in this context starts off with a mutation that introduces immunity for the disease in the host population with $n_2 > 0$. When immunity increases the chance of survival of the species, subpopulation #2 will increase over time.

When the size of one of the two subpopulations approaches zero, the population becomes a monoculture by definition, with the diversity index $D \rightarrow 0$. Therefore, the full diversity range ($0 \leq D \leq 1$) is available when the

minimum value of n_1 = 0.5. When the R_0 factor remains constant for the population as a whole, this leads to $D = 2n_2$ and $R_1 = R_0 / n_1$. Preventing n_1 to approach zero also prevents the corresponding R_1 to reach infinity.

The analysis in annex A shows that when D increases, the infected fraction of the host population decreases steadily. This result is verified in the next section with the heterogeneous SIS model.

Diversity in the heterogeneous SIS Model

To analyze the effects of diversity in a population the least complex heterogeneous SIS model is used, set up as neutrally as possible: two equally sized and static subpopulations that can have different R_0 values. If the graph is completely symmetrical, the basic reproduction number R_0 of a population is the sum of the weighted R_0 factors of its subpopulations [57]. Similar to the homogeneous model, an epidemic will occur when for the population as a whole $R_0 > 1$. The analysis assumes a genetic drift that changes the gap between the R_0 of both subpopulations, while the R_0 for the population as a whole remains constant.

As an example, assume $R_0 = 2$ for the whole population. For a monoculture, the diversity index (D) is zero when the basic reproduction numbers of both subpopulations are equal ($R_1 = R_2 = R_0$). The effect of diversity can be analyzed by increasing the gap between the R_0 values of both subpopulations ($D = (R_1 - R_2) / 2R_0$), while maintaining $R_0 = 2$ for the population as a whole. Therefore, when $R_1 = 3$ for the first sub-population, then $R_2 = 1$ for the second sub-population. At maximum diversity $D = 1$, $R_1 = 2R_0 = 4$ for subpopulation #1 and $R_2 = 0$ for subpopulation #2.

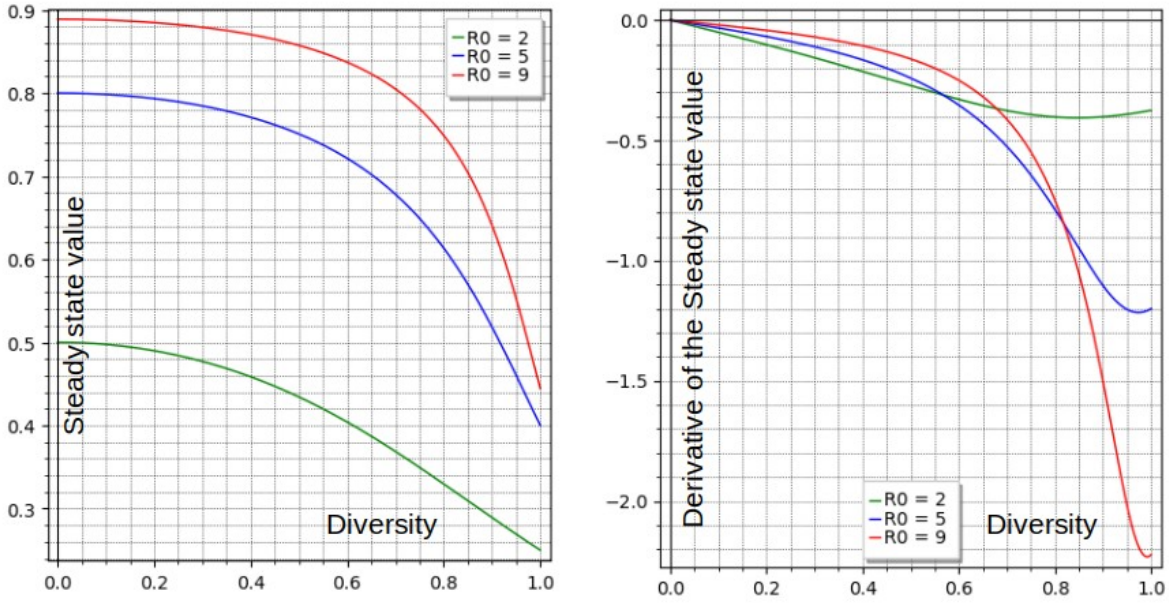

Figure 2. As functions of diversity, the steady state value (left) and its derivative (right) were calculated for the heterogeneous SIS model with $n_1 = n_2 = 0.5$, $R_0 = 2$, 5 and 9

As a function of diversity, the infected fraction (v) in the steady state is shown in Figure 2 by the graphs on the left. The derivative for diversity of (v) shows that increasing diversity steadily decreases the impact of epidemics. Similar to the homogeneous SIS model, the infected fraction of the host population gradually decreases when diversity increases. For almost the entire diversity interval it also holds that the more diverse a host population becomes, the *faster* the impact of epidemics reduces. For populations with nondominant species the HIT, and therefore the impact of epidemics is lower, improving their chance of survival.

Simulating the heterogeneous SIS model in undirected networks

Based on the formulas of the heterogeneous SIS model in annex A, various populations are simulated in a range of input values (R_0 = [2, 2.25, 2.5, 3, 4, .. , 9] and diversity index D = [0.0, 0.25, 0.5, 0.75, 1.0]). For each host population in the steady state, the size of the infected fraction is calculated. When $D = 0$ the R_0 values of both subpopulations are equal to the R_0 of the population; at maximum diversity $D = 1$ and the R_0 values of the subpopulations are $R_1 = 2R_0$ and $R_2 = 0$.

The number of nodes influences the outcome of the simulation. Increasing the network from 500 to 5000 nodes reduced stochastic noise with 45%.
Simulating populations with low steady state values requires higher values of (k) to reach the predicted

steady state value. For instance, at maximum diversity with R_0 = 2 and k = 8, approximately 80% of simulated epidemics die out. To reach 99% of the steady state value predicted by the mathematical model (0.25), the value of (k) must be at least 128. To stay above 99% accuracy, the value of (k) for various values of R_0 is set as follows: (R_0 < 3, k = 128); (R_0 = 3, k = 64); (R_0 = 4, k = 32); (R_0 > 4, k = 16). Figure 3 shows an example of a simulation.

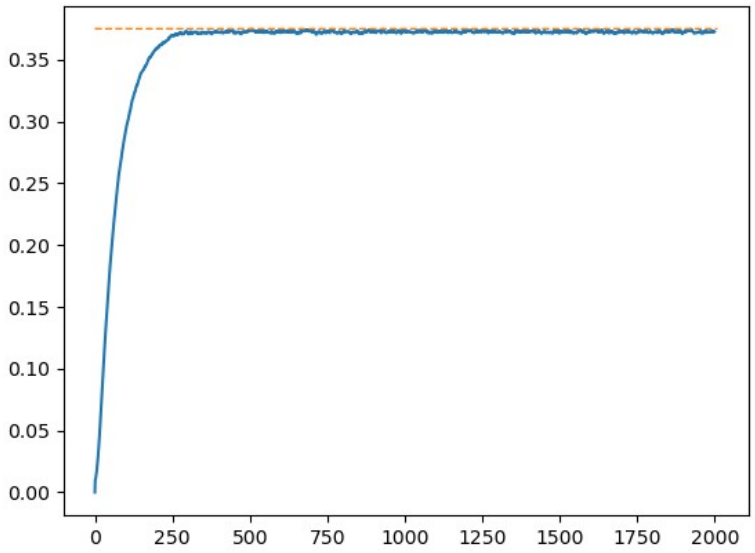

Figure 3. Simulation of a undirected network with 2 subpopulations of 2500 nodes each, with R_0= 4.0, k = 32 and D = 1 (R_1, R_2 = 8.0, 0.0) in 2000 time steps (x-axis). The calculated steady state (the horizontal dotted line) = 0.375; the simulation average over 50 runs = 0.373 and its standard deviation = 0.0005

Table 2 summarizes the results of simulations of undirected networks for R_0 = [2, 5, 9]. The Columns R_0, R_1 and R_2 contain the basic reproduction numbers of the population as a whole and the first and second subpopulation respectively. Column v_∞ shows the values calculated with the deterministic model in Annex A for the discrete values of D (diversity index). The columns *Sim μ* and *Sim σ* list the value and standard deviation of the steady state, averaged over 50 simulation runs.

R_0	D	R_1	R_2	v_∞	Sim μ	Error % of v_∞	Sim σ
2.00	**0.00**	2.00	2.00	0.50000	0.49805	0.39%	0.00057
2.00	**0.25**	2.50	1.50	0.48414	0.48211	0.42%	0.00047
2.00	**0.50**	3.00	1.00	0.43426	0.43140	0.66%	0.00053
2.00	**0.75**	3.50	0.50	0.34946	0.34676	0.77%	0.00071
2.00	**1.00**	4.00	0.00	0.25000	0.24767	0.93%	0.00067
5.00	**0.00**	5.00	5.00	0.80000	0.79735	0.33%	0.00036
5.00	**0.25**	6.25	3.75	0.78950	0.78732	0.28%	0.00020
5.00	**0.50**	7.50	2.50	0.75081	0.74771	0.41%	0.00036
5.00	**0.75**	8.75	1.25	0.64912	0.64486	0.66%	0.00038
5.00	**1.00**	10.00	0.00	0.40000	0.39690	0.78%	0.00042
9.00	**0.00**	9.00	9.00	0.88889	0.88814	0.08%	0.00025
9.00	**0.25**	11.25	6.75	0.88236	0.88168	0.08%	0.00018
9.00	**0.50**	13.50	4.50	0.85731	0.85629	0.12%	0.00022
9.00	**0.75**	15.75	2.25	0.78108	0.77930	0.23%	0.00030
9.00	**1.00**	18.00	0.00	0.44444	0.44357	0.20%	0.00030

Table 2: steady-state values of the whole population with two subpopulations from the deterministic model v_∞ and simulations (Sim μ), with the difference as percentage of v_∞

As Table 2 shows, v_∞ and *Sim μ* are very similar. The correlation is 0.99998 and the average of *Sim σ* is 0.0004. The steady-state values of the simulations of an undirected network (*Sim μ*) are slightly lower than the values of the deterministic model v_∞, especially at lower R_0 values and higher D values. Therefore, simulations confirm the predictions of the deterministic model that diversity reduces the impact of epidemics significantly. When the nodes in the simulated networks have less then the required number of contacts (k), diversity reduces the impact of epidemics at low R_0 values even more than the deterministic model indicates.

The deterministic model predicts that for m equally sized subpopulations the maximum reduction factor going from ($D = 0 \rightarrow 1$) is *(m)*. This means that creating a new subpopulation always reduces the impact of epidemics. But the added value of creating a new subpopulation at (D=1) going from (*m-1*) to (m) subpopulations will vanish for large values of (m).

Conclusion and open questions

If Darwinian evolution is summarized as *the survival of the fittest,* the arms race between predator and prey is usually viewed as the most important factor. The MacArthur–Wilson theory of Island Bio Geography states that dominant species are less likely to become extinct because they out compete other species in terms of population size. However, populations with dominant species will be genetically more uniform.

As stated in the Pathogen-Host section, pathogens can be seen as fast mutating predators. Susceptible host species require time to become resilient to new epidemics. When certain host species are highly susceptible for a deadly pathogen, epidemics could even lead to its extinction, if that species cannot reproduce fast enough to compensate for the mortality of these epidemics [46].

In this paper populations with two equally sized and static subpopulations were analyzed, that differ in their susceptibility to a contagious disease. The SIS model can be used to determine the effect of diversity in this case. Both the analysis of the SIS model and the simulations show that the more genetically different the two subpopulations in a population become, the lower the HIT and the impact of epidemics becomes. This increases the chances for survival for such a population. Conversely, the impact of epidemics is maximal for monocultures, because there is no diversity within the host population to obstruct the infection process. On an evolutionary time scale, virulent contagious diseases therefore oppose the mechanism to produce a “winner takes all” genome that scores the best in the food chain network. Therefore the conclusion of this analysis is that contagious diseases could serve as a KtW actor.

The calculations in paragraph 4 of Annex A show that when a population with m subpopulations reaches maximum diversity, the impact of epidemics is reduced by a factor (m), as compared to a monoculture. That means in this context that when nature increases the number of subpopulations, this continues to reduce the impact of epidemics. But the added value of creating a new subpopulation at maximum diversity will vanish when already many subpopulations exist.

This analysis revealed several areas for further research. First, it should be noted that this analysis relies on the SIS model, which greatly simplifies the real world. It should be applied with care because of its limitations. The SIS model requires a static population that is completely symmetrical, i.e. the infected and susceptible nodes are distributed evenly in the network. Because the SIS model is based on the exponential behavior of contagious diseases, the results of this analysis are significant on the same order of magnitude.

To overcome these restrictions, studies could use more complex models that better resemble reality to determine the effect of diversity on the impact of epidemics. The analysis should focus on genetic vulnerabilities that result in varying susceptibility for contagious diseases, but are unrelated to the fitness of host species in the food chain. For instance, by simulating host populations with mutating pathogens and multiple predator-prey relations that are not perfectly mixed.
Although the simulations confirm the predictions of the deterministic model, at low steady state values the simulation require higher values of the number of contacts (k) for each node in the network to reach the calculated steady state value. Why this is necessary is unclear. Future research could clarify the reasons behind this by using more sophisticated simulation techniques.

The MacArthur–Wilson theory and the diversity induced decrease of the impact of epidemics seem two opposing forces, since both are based on the size of genetically identical groups in a population. A study could investigate whether this leads to equilibrium in species diversity, i.e. an optimal population size and number of species that vary in susceptibility [58].

Henk-Jan van der Molen CISSP CISM CISA has a background in electrical engineering, audits, information systems design, software development, project management, cyber security and privacy. He is a freelance teacher at Security Academy (henk-jan.van.der.molen@securityacademy.org)

Annex A – mathematical calculations

1. Diversity in the homogeneous SIS model

The homogeneous SIS model is described by the following differential equation:

$$\frac{dv}{dt}=\beta k\, v(1-v)-\delta v \;\rightarrow\; \frac{dv}{dt}=\delta v\left(R_0(1-v)-1\right) \tag{1}$$

The curing rate is denoted as δ, and the infection rate of all incoming links is denoted as β. Each node has an equal number of connections (k), with half of these connections directed to nodes in each subpopulation. The fractions of infected and susceptible nodes are represented by v and $(1 – v)$, respectively. The basic reproduction number R_0 is used to replace $\beta.k / \delta$. When $R_0 > 1$, at $t = \infty$ the steady state v_∞ can be calculated by setting the differential equation to zero and ignoring the solution $v_\infty = 0$:

$$\frac{dv}{dt}=0=v_\infty\left(R_0(1-v_\infty)-1\right) \;\Rightarrow\; v_\infty=1-\frac{1}{R_0} \tag{2}$$

For brevity v is used instead of v_∞. If the fraction n_2 of the population becomes immune to a disease ($R_2 = 0$) and the rest of the population n_1 ($= 1 – n_2$) remains susceptible ($R_1 > 0$), the susceptible part of the population in (2) changes from $(1 – v)$ to $(1 – n_2 – v)$. Thus:

$$v\left(R_0 v(1-n_2-v)-1\right)=0 \;\Leftrightarrow\; v=1-n_2-\frac{1}{R_0};\quad \text{when}\quad n_2\geqslant\left(1-\frac{1}{R_0}\right)\quad \text{the HIT is reached.} \tag{3}$$

To perform a fair analysis of the effect of diversity on the steady state of v, the value of R_0 for the population as a whole must be kept constant, above the epidemic threshold. If R_0 is not kept constant, v will fluctuate with the value of R_0, which would invalidate the analysis.

When n_1 decreases, the value of R_0 for the population as a whole remains constant if $R_1 = R_0 / n_1$ (4)

When $\{n_1 = 0\}$ or $\{n_1 = 1\}$, this by definition constitutes a monoculture, so the diversity index (D) is defined as: $D = 2n_1n_2R_1 / R_0$, with $\{½ \leq n_1 < 1\}$, so $\{0 < D \leq 1\}$. (5)

Inserting (4) in eq. (3) and D in eq. (5): $v=1-n_2-\frac{(1-n_2)}{R_0};\quad D=\frac{2\,n_1\,n_2\,R_1}{R_0}=2\,n_2$ (6)

To investigate how v behaves as a function of diversity, its derivative v' is computed:

$$v' \;=\; \frac{dv}{dD} \;=\; \frac{dv}{2\,dn_2} \;=\; -\frac{1}{2}\left(1-\frac{1}{R_0}\right) \tag{7}$$

When diversity increases in the interval [D = 0..1], v decreases steadily between its extremes:

$$v_{D=0}=1-\frac{1}{R_0};\quad v_{D=1}=\;\frac{1}{2}\left(1-\frac{1}{R_0}\right) \tag{8}$$

This means that $v_{D=0}$ at maximum diversity (D=1) is reduced by a factor of 2.

2. Diversity in the heterogeneous SIS model with (m) subpopulations

For clarity, some formulas for the heterogeneous SIS model are restated of a previous publication [57]. The general heterogeneous SIS model for a regular graph is described by the following equations:

$$\frac{dv_i}{dt}=\delta_i\left(R_i v(1-v_i)-v_i\right);\quad \sum_{i=1}^{m}(n_i)=1;\quad v=\sum_{i=1}^{m}(n_i v_i);\quad R_0=\sum_{i=1}^{m}(n_i R_i) \tag{9}$$

In a population with m subpopulations, the curing rate of subpopulation i is denoted as δ_i, and the basic reproduction number is denoted as R_i, for $i = \{1, 2, \ldots, m\}$. The fractions of infected nodes of subpopulation i and the population as a whole are represented by v_i and v, respectively. As in the homogeneous SIS model, a disease becomes epidemic when its $R_0 > 1$. The steady state v_∞ is calculated, ignoring $v_\infty = 0$ as a solution. Again, for brevity v is used instead of v_∞.

$$\frac{dv_i}{dt}=0 \;\rightarrow\; v_i=\left(\frac{R_i v}{R_i v+1}\right);\;\; v=\sum_{i=1}^{m} n_i v_i \;=\; v\sum_{i=1}^{m}\left(\frac{n_i R_i}{R_i v+1}\right) \;\rightarrow\; (\forall\, v\neq 0)\;\; \sum_{i=1}^{m}\left(\frac{n_i R_i}{R_i v+1}\right)=1 \qquad (10)$$

3. Diversity in the heterogeneous SIS model with 2 static, equally sized subpopulations

For two equally sized subpopulations {m = 2, $n_1 = n_2 = ½$ } the steady state of (v) only depends on the parameters R_1 and R_2.

Derived from (10): $n_1R_1(R_2 v + 1) + R_2n_2(R_1 v + 1) = (R_1 v + 1)(R_2 v + 1)$, thus:

$$av^2+bv+c=0 \;\Rightarrow\; v=\frac{-b+\sqrt{b^2-4ac}}{2a} \quad (a=R_1R_2;\;\; b=R_1+R_2-R_1R_2;\;\; c=1-R_0) \qquad (11)$$

When {$R_0 > 1$} then {$c < 0$}, so the heterogeneous model always yields a solution with $v > 0$.

Diversity (D) is defined similarly to (5): $D = 2n_1n_2|R_1 - R_2| / R_0$, with {$0 \le n_1, n_2 \le ½$}, so {$0 \le D \le 1$}.

Keeping the value of R_0 for the population as a whole constant, filling in $n_1 = n_2$ set to a constant value guarantees maximum diversity. So R_1 and R_2 are the only remaining variables:

$$n_1=n_2=\frac{1}{2} \;\rightarrow\; R_1+R_2=2R_0 \;\rightarrow\; R_2=2R_0-R_1 \qquad (12)$$

Assuming {$R_1 \ge R_2$} then {$R_0 \le R_1 \le 2R_0$} and $D = (R_1 – R_0) / R_0 \equiv R_1 / R_0$.

To investigate how the steady state of v behaves as a function of diversity, its derivative v' is computed:

$$v'=\frac{dv}{dD}=\frac{-2ab'+2a'b}{4a^2}+\frac{ad'}{4a^2\sqrt{d}}-\frac{2a'\sqrt{d}}{4a^2}=\frac{(2a'b-2ab')\sqrt{d}+ad'-2a'd}{4a^2\sqrt{d}} \qquad (13)$$

In the following R is used instead of R_1 and ($2R_0 – R$) is used for R_2.

$$\begin{aligned} &a=-R^2+2R_0R;\;\; b=R^2-2R_0R+2R_0;\;\; c=1-R_0;\\ &d=b^2-4ac=R^4-4R_0R^3+4R_0^2R^2+4R^2-8R_0R+4R_0^2 \end{aligned} \qquad (14)$$

$$a'=\frac{da}{dD}=R_0\frac{da}{dR}=2R_0(R_0-R);\quad b'=-a';\;\; d'=R_0\frac{dd}{dR}=4R_0(R^2-2RR_0+2)(R-R_0) \qquad (15)$$

Filling in (14) and (15) in (13) yields:

$$v'=2R_0(R-R_0)\frac{(R^2-2R_0R+2R_0^2-R_0\sqrt{d})}{R^2(R-2R_0)^2\sqrt{d}}$$

Because {$R_0 > 1$} and {$R_0 \le R < 2R_0$} and therefore {$R^2 - 2R_0R + 2R_0^2 > 0$}, the sign of v' depends only on:

$$(R^2-2R_0R+2R_0^2)^2-R_0^2d \;=\; -(R_0^2-1)R^2(R-2R_0)^2 \qquad (16)$$

When D increases, v decreases steadily between the same extremes as stated in (8). In this case the same reduction factor of 2 applies to $v_{D=0}$ at maximum diversity (D=1).

Note: is $R_2 = 0$, then this (m=2) model can be reduced to an (m=1) model, see equation (3).

4. Diversity in the heterogeneous SIS model with 3 static, equally sized subpopulations

Although not verified by simulations, for completeness the mathematical analysis of this model is included. From (10) the formula can be derived for the steady state with 3 static subpopulations (m=3).

$$\begin{aligned} &av^3+bv^2+cv+d=0;\;\; with\;\; a=R_1R_2R_3;\;\; b=R_1R_2+R_1R_3+R_2R_3-R_1R_2R_3;\\ &c=R_1+R_2+R_3-R_1R_2(n_1+n_2)-R_1R_3(n_1+n_3)-R_2R_3(n_2+n_3);\;\; d=1-R_0 \end{aligned} \qquad (17)$$

Assuming that $\{R_1 \geq R_2 \geq R_3 \geq 0\}$, the diversity factor D is derived from its general form as follows.

$$D = 3\left(\frac{n_1 n_2 |R_1 - R_2| + n_1 n_3 |R_1 - R_3| + n_2 n_3 |R_2 - R_3|}{2 R_0}\right) = \frac{R_1 - R_3}{3 R_0}$$

In table 3 different scenario's are listed with increasing diversity for equally sized and static subpopulations $\{n_1 = n_2 = n_3 = ⅓;\ R_1 + R_2 + R_3 = 3R_0\}$ The base case is a monoculture $\{D = 0\}$, thus $R_1 = R_2 = R_3 = R_0$, and $v = 1 - 1/R_0$. In the scenarios with $D > 0$ the impact of epidemics is significantly decreased, compared to the base case. The values in the Column *Reduction factor* v_∞ were calculated for $(2 \leq R_0 \leq 9)$.

D	R_1	R_2	R_3	**Reduction factor for v_∞**
0.000	R_0	R_0	R_0	1.000
0.167	1.25 * R_0	R_0	0.75 * R_0	1.005 – 1.022
0.333	1.50 * R_0	R_0	0.50 * R_0	1.024 – 1.097
0.500	1.75 * R_0	R_0	0.25 * R_0	1.086 – 1.262
0.667	2.00 * R_0	R_0	0	1.520 – 1.591
0.750	2.25 * R_0	0.75 * R_0	0	1.555 – 1.727
0.833	2.50 * R_0	0.50 * R_0	0	1.631 – 1.968
0.917	2.75 * R_0	0.25 * R_0	0	1.849 – 2.373
1.000	3.00 * R_0	0	0	3.000 (=m)

Table 3: steady-state reduction factors calculated for varying diversity factors (D) using the deterministic model with three static and equally sized subpopulations

Populations with m equally sized subpopulations reach maximum diversity (D=1), when $R_1 = mR_0$, and $R_i = 0\ \forall\ i \in [2, \ldots, m]$. It is easy to verify that for m subpopulations the maximum reduction factor is (m), as compared tot the monoculture situation (D=0). This means that each new subpopulation reduces the impact of epidemics. But the added value of a new subpopulation decreases with the number of subpopulations that already exists, because going from (m-1) to (m) subpopulations the reduction factor increases with $m/(m\text{-}1)$, which reaches unity when $\{\text{Lim } m \rightarrow \infty\}$.

Annex B – Python 3 source code of the simulation

```
''' Run a simulation of the heterogeneous SIS model with 2 subpopulations.
Graph implementation based on: https://stackabuse.com/courses/
graphs-in-python-theory-and-implementation/lessons/representing-graphs-in-code/

*** Formulas - rnd = random sample from U [0..1] ***
chance node is infected from k infected contacts  = beta * k  = (rnd < beta * k)
chance infected node recovers from infection      =     delta = (rnd < delta)
chance infected node does NOT recover             = 1 - delta = (rnd > delta)

If  node = 0    # Susceptible, chance of (re)infection =
    node = 1 * (rnd < beta * count_infected_contacts)
Else            # Infected, chance to remain infected =
    node = 1 * (rnd > delta)

Author: Henk-Jan van der Molen
Last source code versions on: https://github.com/HJvdMolen/SIS-model.git
'''

def calc_stats(floatlist):  # calculate μ, σ  from a list of values
    from math import sqrt
    total = total2 = 0
    for value in floatlist:
        total  += value
        total2 += value ** 2

    n = len(floatlist)
    mu = total / n
    sd = sqrt((total2 - total ** 2 / n) / n)
    return mu, sd

def plot_function(function, title, xlabel, ylabel, v_8):
    import matplotlib.pyplot as plt
    plt.plot(function)
    plt.plot( [v_8 for _ in range(10 + len(function))], linestyle='dashed', linewidth=1)
    plt.title(title)
    plt.xlabel(xlabel)
    plt.ylabel(ylabel)
    plt.show()
    return

def calc_beta_delta(r0, k): # calculates β, δ as big as possible
    assert k % 4 == 0,      f"Variable k must be a multiple of 4, got {k}"
    assert r0 >= 0,         f"r0 must be >= 0, got {r0}"

    if r0 > 1:
        delta = 1 / r0
    else:
        delta = 1
    assert 0 < delta <= 1,  f"0 < δ <= 1 expected, got: {delta}"

    beta = r0 * delta / k
    assert beta * k <= 1,   f"β.k <= 1 expected, got: {beta * k}"

    print(f" r0 = {r0}; k = {k}; beta = {beta:.3f}; delta = {delta:.3f}")
    return beta, delta

class Graph:
    # Constructor
    def __init__(self, network_size, n1, no_of_steps, r0, D):
        self.network_size = network_size
        self.n1 = n1
        self.n2 = 1 - n1
        self.no_of_steps = no_of_steps
        self.r0 = r0
        self.D  = D

        # Calculated later...
```

```
        self.v_8 = 0.0
        self.mu = self.sd = 0.0

        # Create simulation + average simulation list
        self.sim_tot = [0 for j in range(self.no_of_steps)]
        self.sim_avg = [0 for j in range(self.no_of_steps)]
        return

    def calc_steady_state(self, p1, p2):
        from math import sqrt

        a           = p1.r0 * p2.r0       # calculate steady state of v
        if a != 0:
            b       = p1.r0 + p2.r0 - a
            c       = 1 - self.r0
            d       = b * b - 4 * a * c
            self.v_8 = (sqrt(d) - b) / (2 * a)
        elif p1.r0 == 0:
            self.v_8 = self.n2 - 1 / p2.r0
        elif p2.r0 == 0:
            self.v_8 = self.n1 - 1 / p1.r0
        return

    # Execute complete simulation
    def run_simulation(self, p1, p2):
        p1.init_simulation()
        p2.init_simulation()
        self.sim_tot = [0 for j in range(self.no_of_steps)]

        for j in range(1, self.no_of_steps):
            p1.simulation_step(p2, j)
            p2.simulation_step(p1, j)

            p1.node = p1.temp [:]
            p2.node = p2.temp [:]

            avg_v = self.n1 * p1.sim [j] + self.n2 * p2.sim [j]
            self.sim_tot [j]  = avg_v
            self.sim_avg [j] += avg_v

        # Calculate (μ, σ) from steady state part = right half of sim []
        self.mu, self.sd = calc_stats(self.sim_tot [self.no_of_steps//2:])
        return

class SubPopulation:
    def __init__(self, num_of_nodes, beta, k, delta, no_of_steps, directed):
        self.num_of_nodes = num_of_nodes
        self.no_of_steps = no_of_steps
        self.directed = directed

        self.beta = beta
        self.k = k
        self.delta = delta
        self.r0 = beta * k / delta

        # Create the nodes list
        self.node = [0 for j in range(self.num_of_nodes)]

        # Create Temp copy of self.node []
        self.temp = [0 for j in range(self.num_of_nodes)]

        # Create simulation list for SubPopulation
        self.sim = [0 for j in range(self.no_of_steps)]

        # Initialize & fill the adjacency list
        self.adj_list = {node: set() for node in range(num_of_nodes)}

        # adj_list is also used for contacts between SubPopulations = /2
        if self.directed:   # directed graph
            contacts_out = self.k // 2
```

```
        else:               # undirected = include "mirror" edges
            contacts_out = self.k // 4

        for key in self.adj_list.keys():
            for j in range(contacts_out):
                # Link node(n) to node(n+1, n+2, .. , n+contacts_out)
                self.add_edge(key, (key + 1 + j) % self.num_of_nodes, directed)
        return

    # Add the edge from node1 to node2 and v.v. in undirected graphs
    def add_edge(self, node1, node2, directed):
        assert node1 != node2, f"node {node1} cannot connect to itself"

        self.adj_list [node1].add(node2)
        if not (self.directed or node1 == node2):   # also add "mirror" edge
            self.adj_list [node2].add(node1)
        return

    def init_simulation(self):
        import random

        # Set all nodes to be Susceptible (== 0)
        for j in range(self.num_of_nodes):
            self.node [j] = 0
            self.temp [j] = 0

        if False:   # switch to start with 100% infected
            self.node = [1 for j in range(self.num_of_nodes)]
            self.nodes_infected = self.num_of_nodes
        else:       # Randomly infect 2% of susceptible nodes (0 => 1)
            self.nodes_infected = self.num_of_nodes // 100
            for j in range(self.nodes_infected):
                while True:
                    p = random.randrange(0, self.num_of_nodes)
                    if self.node [p] == 0 : break
                self.node [p] = 1    # node is infected

        # Fill in first value in simulation list
        self.sim = [0 for j in range(self.no_of_steps)]
        self.sim [0] = self.nodes_infected / self.num_of_nodes
        return

    # #infected contacts (node []==1); p1 + p2 have identical adjacency list
    def count_infected_contacts(self, other, key):
        cic = 0
        for nd in self.adj_list [key]:
            cic +=  self.node [nd]   # If node [nd] = 1 -> Infected
            cic += other.node [nd]
        return cic

    # Execute single time step in the simulation for all nodes
    def simulation_step(self, other, step_no):
        import random
        self.nodes_infected = 0
        for j in range(self.num_of_nodes):
            rnd = random.uniform(0,1)
            if self.node [j] == 0:    # Susceptible node, can be infected
                contacts = self.count_infected_contacts(other, j)
                self.temp [j] = 1 * (rnd < self.beta * contacts)
            else:                     # Infected node, can recover
                self.temp [j] = 1 * (rnd > self.delta)
            self.nodes_infected += self.temp [j]

        self.sim [step_no] = self.nodes_infected / self.num_of_nodes
        return

# Main program
if __name__ == '__main__':
    network_size  = 5000            # no. of nodes in the network
    k             = 8               # no. of contacts for each node
```

```
    n1              = 0.5               # fraction of SubPopulation 1
    n2              = 1 - n1            # fraction of SubPopulation 2
    directed        = True              # False = bidirectional network
    assert n1 == n2 == 0.5, f"n₁, n₂ must be 0.5, got n1 = {n1}, n2 = {n2}"

    r0              = 3                 # basic reproduction number R₀
    assert r0 > 1, f"R₀ > 1 expected, got: {r0}"

    D               = 0.00              # Diversity index, [0 = min, 1 = max]
    assert 0 <= D <= 1, f"0 <= Diversity <= 1, got {D}"

    no_of_steps     = 2000              # no. of time steps in 1 simulation
    g               = Graph(network_size, n1, no_of_steps, r0, D)

    # SubPopulation #1 & #2
    num_of_nodes1 = int(network_size * n1)
    num_of_nodes2 = network_size - num_of_nodes1
    r0_1          = r0 * (1 + D / (4 * n1 * n2))
    r0_2          = (r0 - n1 * r0_1) / n2
    # β: infection chance from 1 infected contact | δ: recovery from infection
    beta1, delta1 = calc_beta_delta(r0_1, k)
    beta2, delta2 = calc_beta_delta(r0_2, k)
    p1            = SubPopulation(num_of_nodes1, beta1, k, delta1, no_of_steps, directed)
    p2            = SubPopulation(num_of_nodes2, beta2, k, delta2, no_of_steps, directed)

    g.calc_steady_state(p1, p2)    # calculates D and v(∞)
    g.run_simulation(p1, p2)

    title  = f"SIS simulation: R₀, R₁, R₂ = {g.r0}, {p1.r0}, {p2.r0}; v(∞) = {g.v_8}"
    xlabel = f"#time steps: {no_of_steps} with {network_size} nodes, directed network:
{directed}"
    ylabel = f"μ(v) = {g.mu:.5f}, σ(v) ={g.sd:.5f}"
    plot_function(g.sim_avg, title, xlabel, ylabel, g.v_8)
```

Funding: this research was not funded by any organization. It was supported by individuals investing time in reviewing draft versions of the article. The reviewers had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1 Hutchinson GE. The paradox of the plankton. Am Nat. 1961;95-882:137
2 Hardin G. The competitive exclusion principle. Science 1960;131:1292-7
3 Xue C., Goldenfeld N. Co-evolution maintains diversity in the stochastic “kill the winner” model. Phys Rev Lett. 2017;119:268101
4 Kevin De Queiroz K, Species Concepts and Species Delimitation, Systematic Biology, Volume 56, Issue 6, December 2007, p. 879–886, https://doi.org/10.1080/10635150701701083
5 Winter C, Bouvier T, Weinbauer MG, Frede Thingstad T. Trade-Offs between competition and defense specialists among unicellular planktonic organisms: the “killing the winner” hypothesis revisited. Microbiol Mol Biol Rev. 2010: doi:10.1128/MMBR.00034-09
6 Bonnet T. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. Science 2022; DOI:10.1126/science.abk0853
7 Springbett AJ, MacKenzie K, Woolliams JA, Bishop SC. The contribution of genetic diversity to the spread of infectious diseases in livestock populations. Genetics 2003;165:1465–1474
8 Elbasha EH, Gumel AB, Vaccination and herd immunity threshold in heterogeneous populations. J. Mathematical Biology 2021; https://doi.org/10.1007/s00285-021-01686-z
9 Dunbar R. How many friends does one person need? – Dunbar's Number and Other Evolutionary Quirks. Harvard University Press 2010; p.143
10 MacArthur RH, Wilson EO. The Theory of Island Biogeography. Princeton University Press, 1967
11 Pérez-Gutiérrez R, López-Ramírez V, Islas A, Alcaraz LD, Hernández-González I, Olivera BCL et al. Antagonism influences assembly of a Bacillus guild in a local community and is depicted as a food-chain network. The International Society for Microbial Ecology Journal. 2013;doi:10.1038/ismej.2012.119
12 van Elsas JD, Chiurazzi M, Mallona CA, Elhottovāb D, Krištůfekb V, Falcão Salles J. Microbial diversity determines the invasion of soil by a bacterial pathogen. PNAS. 2012; www.pnas.org/cgi/doi/10.1073/pnas.1109326109
13 The Theory of Island Biogeography Revisited. Editors: Losos JB, Ricklefs RE. Princeton University Press. 2010; p. 65
14 López-Gómez J, Ortega R, Tineo A. The periodic predator-prey lotka-volterra model. Adv Differ Equ Volume 1. 1996;3:403-23
15 Hall SR, Lafferty KD, Brown JH, Cáceres CE, Chase JM, Dobson AR et al. Is infectious disease just another type of predator-prey interaction? In: Ostfeld RS, Keesing F, Eviner VT, editors. Infectious disease ecology. Princeton: Princeton University Press; 2010. p. 223-41
16 Dennet DC. From bacteria to bach and back: The evolution of minds. 1st ed. New York: WW Norton & Company. 2017
17 Lostroh P. Molecular and Cellular Biology of Viruses. CRC Press. 2019. p. 429
18 Xia G. and Wolz C. Phages of staphylococcus aureus and their impact on host evolution. Infect Genet Evol. 2014;21:593–601
19 Read AF, Baigent SJ, Powers C, Kgosana LB, Blackwell L, Smith LP et al. Imperfect vaccination can enhance the transmission of highly virulent pathogens. PloS Biol 13. 2015;13:e1002198
20 Mooi FR, van Loo HM, King AJ. Adaptation of Bordetella pertussis to Vaccination: A cause for its reemergence? Emerg Infect Dis. 2001;7-3
21 Lam C, Octavia S, Ricafort L, Sintchenko V, Gilbert GL, Wood N et al. Rapid increase in pertactin-deficient bordetella pertussis isolates. Emerg Infect Dis Vol. 20, 2014; DOI: http://dx.doi.org/10.3201/eid2004.131478
22 Leong J, Lin D, Nguyen MH. Hepatitis B surface antigen escape mutations: Indications for initiation of antiviral therapy revisited. World J Clin Cases. 2016; DOI:10.12998/wjcc.v4.i3.71
23 Snitkin ES, Zelazny AM, Thomas PJ, Stock F, NISC Comparative Sequencing Program, Henderson DK. Tracking a hospital outbreak of carbapenem-resistant klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med. 2012;4-148
24 Maddamsetti R, Hatcher PJ, Green AG, Williams BL, Marks DS, Lenski RE. Core genes evolve rapidly in the long-term evolution experiment with escherichia coli. Genome Biol. Evol. 2017; doi:10.1093/gbe/evx064
25 Weatherall DJ. Genetic variation and susceptibility to infection: the red cell and malaria. Br J Haematol. 2008; doi:10.1111/j.1365-2141.2008.07085.x
26 Kamo M, Sasaki A. Evolution toward multi-year periodicity in epidemics. Ecol Lett. 2005;doi: 10.1111/j.1461-0248.2005.00734.x
27 Butler S, O’Dwyer J. Stability criteria for complex microbial communities. Nat Commun. 2018;9
28 Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: Networks, competition, and stability. Science. 2015;350-6261
29 Abreu NA, Nagalingam NA, Song Y, Roediger FC, Pletcher SD, Goldberg AN et al. Sinus microbiome diversity depletion and corynebacterium tuberculostearicum enrichment mediates rhinosinusitis. Sci

References

Transl Med. 2012; DOI:10.1126/scitranslmed.3003783

30 dos Santos, MBQ, Nielsen HB, Sicheritz-Pontén T, Metagenomic analysis of the human gut microbiome. technical university of Denmark (DTU). 2013

31 Comito D, Cascio A, Romano C. Microbiota biodiversity in inflammatory bowel disease. Ital J Pediatr. 2014; doi:10.1186/1824-7288-40-32

32 Cui H, Li Y, Zhang X. An overview of major metagenomic studies on human microbiomes in health and disease. Quant Biol. 2016;4(3)

33 van der Does AM, Amatngalim GD, Keijser B, Hiemstra PS, Villenave R. Contribution of host defence proteins and peptides to host-microbiota interactions in chronic inflammatory lung diseases. Vaccines 2018; doi:10.3390/vaccines6030049

34 O'Brien CL, Kiely CJ, Pavli P. The Microbiome of Crohn's disease aphthous ulcers. Gut Pathog. 2018;10-44

35 Colquhoun C, Duncan M, Grant G. Inflammatory bowel diseases: host-microbial-environmental interactions in dysbiosis. Diseases. 2020; doi:10.3390/diseases8020013

36 Khor B, Snow M, Herrman E, Ray N, Mansukhani K, Pate KA. Interconnections between the oral and gut microbiomes: reversal of microbial dysbiosis and the balance between systemic health and disease; Micro organisms. 2021;9:496.

37 Lange A, Ferguson N. Antigenic Diversity, Transmission Mechanisms and the Evolution of Pathogens. PLoS Computational Biology 5. 2009;10:e1000536

38 Lostroh P. Molecular and Cellular Biology of Viruses. CRC Press. 2019. p. 383

39 Lee SL, Tamerụ AM. A mathematical model of human papilloma virus (HPV) in the United States and its impact on cervical cancer. Science Xpress. 2012; doi:10.7150/jca.4161

40 Fraser C, Donnelly CA,Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD et al. Pandemic potential of a strain of influenza a (H1N1): early findings. Science Xpress. 2009

41 Consensus document on the epidemiology of severe acute respiratory syndrome (SARS), World Health Organization, 2003

42 Smith PG. Concepts of herd protection and immunity. Procedia Vaccinol. 2010; doi:10.1016/j.provac.2010.07.005

43 Gadroen K, Dodd CN, Masclee GMC, de Ridder MAJ, Weibel D, Mina MJ et al. Impact and longevity of measles-associated immune suppression: a matched cohort study using data from the THIN general practice database in the UK. BMJ Open. 2018; doi:10.1136/bmjopen-2017-021465

44 Clay K, Reinhart R, Rudgers J, Tintjer, T, Koslow J, Luke Flory S. Red Queen Communities. In: Ostfeld RS, Keesing F, Eviner VT, editors. Infectious disease ecology. Princeton: Princeton University Press; 2010. p. 156

45 Rifkin RF, Potgieter M, Ramond J, Cowan DA. Ancient oncogenesis, infection and human evolution. Evolutionary Applications. 2017;DOI:10.1111/eva.12497

46 Collinge SK, Ray C, Cully JF jr. Effects of disease on keystone species, dominant species and host communities. In: Ostfeld RS, Keesing F, Eviner VT, editors. Infectious disease ecology. Princeton: Princeton University Press; 2010. p. 126, 127, 140

47 Shanks GD, Brundage JF. Pathogenic Responses among Young Adults during the 1918 Influenza Pandemic. Emerg Infect Dis. 2012; DOI: http://dx.doi.org/10.3201/eid1802.102042

48 Bai X, Plastow GS. Breeding for disease resilience. CABI Agric Biosci. 2022;3:6

49 Dekkers JCM, Hospital F. The use of molecular genetics in the improvement of agricultural populations. Nature Reviews | Genetics. 2002; DOI:10.1038/nrg701

50 Diamond SE, Martin RA. Buying Time: Plasticity and Population Persistence. In: Pfennig DW, editor. Phenotypic Plasticity & Evolution – Causes, consequence, Controversies. Evolutionary Cell Biology, Taylor & Francis group. 2021: DOI:10.1201/9780429343001

51 Wirth J, Young M. Viruses in Subsurface Environments. Annu. Rev. Virol. 2022. 9:99–119; https://doi.org/10.1146/annurev-virology-093020-015957

52 Hulst AD, de Jong MCM, Bijma P. Why genetic selection to reduce the prevalence of infectious diseases is way more promising than currently believed. Oxford genetics, 2021; DOI:10.1093/genetics/iyab024

53 Bai X, Cheng J, Fortin F, Harding JCS, Dyck MK, Dekkers JCM et al. Indicators of disease resilience from complete blood count and in vitro immunoassays data from young-healthy pigs. WCGALP. 2022: DOI:10.3920/978-90-8686-940-4_161

54 Raymond RR, Doerksen T, Lu A, Sheahan M, Lunney J, Dekkers J, Palinski RM et al. Effect of the host genotype at a Porcine Reproductive and Respiratory Syndrome (PRRS) resistance marker on evolution of the modified-live PRRS vaccine virus in pigs. Virus Res 316. 2022; https://doi.org/10.1016/j.virusres.2022.198809

55 Yates A, Antia R, Regoes RR. How do pathogen evolution and host heterogeneity interact in disease emergence? Proc Biol Sci. 2006; doi:10.1098/rspb.2006.368

References

56 Doeschl-Wilson AB, Davidson R, Conington J, Roughsedge T, Hutchings MR, Villanueva B. Implications of host genetic variation on the risk and prevalence of infectious diseases transmitted through the environment. Genetics. 2011; doi:10.1534/genetics.110.125625

57 Kooij RE, van der Molen HJ. On the malware front. Int J Comput Netw (IJCN). 2012;4;5:72-81

58 Futuyma D, Kirkpatrick M. Evolution. 4rd ed. Oxford University Press. 2018, p 507