**The spread of Dutch Elm Disease in elm populations with resistance**

Kaylee Rosenberger

Calculating Biological Quantities, Fall 2022

14 December 2022

**Abstract:** Hundreds of millions of elm trees worldwide are threatened by Dutch Elm Disease (DED). However, genetically based resistance to DED exists in several species of elms. Since trees can hybridize with one another, it is possible for resistance to spread among populations through introgression. Here I model the spread of DED resistance (in the form of major gene resistance) across populations using a diploid selection model to determine 1) how long will it take for resistance to become fixed in a population? 2) what is the optimal initial proportion of resistance to introduce in a susceptible population of elms? 3) can resistance evolve/persist/fix without selective pressure? Additionally, I attempt to create a modified SIR model showing the dynamics of a population of elms experiencing an outbreak of DED.

1. **Introduction**

Plant pathogens and their host species have been co-evolving for hundreds of millions of years (Han, 2018). It is a perpetual arms race, with pathogens continually overcoming host defenses. Pathogens pose a threat to biodiversity, ecosystem stability, and society and have contributed to species extinctions (Potter et al., 2019). The number of diseases species face is increasing due to global connectedness, climate change, and other anthropogenic factors (Sniezko and Koch, 2017). Although plants lack the adaptive immune system of vertebrate species, plants species have developed several innate immune responses to pathogens (Han, 2018). Genetically based disease resistance provides species with an intrinsic mechanism to respond and adapt to pathogens. This has been seen in several plant species which have co-evolved with their pathogens for long periods of time. Several species of elm trees provide an example into this mechanism.

Elm trees in Europe and North America are threatened by Dutch Elm Disease (DED), which results in wilt, dieback, and death in elms (Brunet and Gurles, 2016). Elm tree cultivars were commonly planted across North America and Europe for their shade, aesthetics, and hardiness to diverse climatic conditions. Many species of elms are particularly susceptible to the disease since they were planted as cultivated varieties, resulting in low genetic diversity in many areas, which may correspond to a lack of ability to adapt (Martin et al., 2018). Depending on the age and susceptibility, the tree can die in as little as 2 months to a few years following DED infection (Grabowski, 2019). DED is caused by a group of fungal pathogens, *Ophiostoma ulmi,* and is vectored by bark beetles, which feed on the bark of elms (Brasier, 2001)*.* DED can also spread from tree to tree through root graphs (Grabowski, 2019). Taxa of *Ophiostoma* are rapidly evolving due to hybridization, contributing to more severe outbreaks of DED, threatening millions of elm trees (Brasier, 2001).

While some species of elm are extremely susceptible to the disease, particularly those native to Europe and North America, other species, mostly those found in Asia, are resistant and can overcome infection of DED (Gibbs, 1978). Resistance phenotypes range from susceptible to infection and death, to tolerant (survives with infection), to resistant (avoids infection). Since DED resistance is genetically based and in general, trees have the ability to hybridize with other species, resistance can spread between populations and species through introgression and be passed on to offspring. Introducing resistant genotypes in a population of susceptible trees should allow the resistance genotype to spread over time, theoretically allowing the population to survive DED infection. However, little is known about the gene or genes governing resistance to DED. Additionally, due to the long generation times of trees (at least 15 years to reach reproductive age), resistance spread would be slow, so introducing individuals after the population has been infected would likely not be an effective strategy.

As previously mentioned, several Asian elm species have genetic resistance to DED. This might be expected if these species co-evolved with this pathogen for long periods of time due to selection; however, *Ophiostoma taxa* are not found in Asia, nor is there evidence of *Ophiostoma* having a historical range in Asia (Gibbs, 1978). If *Ophiostoma* did not historically exist in Asia, this poses the questions: *could disease resistance have evolved by chance, from another selection pressure, or from a trait that is linked to resistance?* Additionally, *does resistance persist in species that are not facing the direct threat of disease?*

1. **Learning goals**

Here, I aim to explore the dynamics of DED spread in populations of elms with varying levels of resistance models of selection and disease dynamics models. I aim to answer the questions:

1. How many generations will it take for resistance to become fixed in the population?
2. What is the optimal initial proportion of resistance for the trait to become fixed in a population?
3. What are the allele frequency dynamics for a population that is not undergoing direct selective pressure?

The optimal initial proportion of resistant individuals to introduce in a population would be useful for land managers and conservationists to plan to protect and supplement elm populations to survive DED outbreaks. We would want to introduce the smallest number of individuals in the population due to logistical constraints. The number of generations for resistance to become fixed is also important for planning purposes. Finally, I wish to know the allele frequency dynamics for a population that is not under selection to explore if resistance could evolve in a population without a DED outbreak (as in the case of Asian elm trees).

I hypothesize that it will take many generations for resistance to become fixed in a population, even with strong selective pressure, due to the long generation time of elm trees. However, resistance will become fixed quicker with a higher selective pressure. Additionally, the greater the initial proportion of resistant individuals, the faster resistance will become fixed. Lastly, I hypothesize that resistance will not become fixed in a population that is not under selection; instead, the alleles will exist in an equilibrium that is in between 0 and 1 (related to their initial values).

First, I will model resistance in the form of major gene resistance (resistance controlled by a single locus). Full resistance will be represented by a homozygous dominant genotype (AA), tolerance will be represented by a heterozygous genotype (Aa), and susceptibility will be represented by a homozygous recessive genotype (aa). Most elms are diploid, so I will use the diploid selection model to analyze the allele frequency dynamics of this population based on different initial proportions of alleles. Then, I will model the population dynamics of a population of elm trees with DED infections using a disease dynamics model that is similar to the standard SIR model. Again, over time, I hypothesize that resistance will become fixed when a population is experiencing DED infections.

1. **Methods**

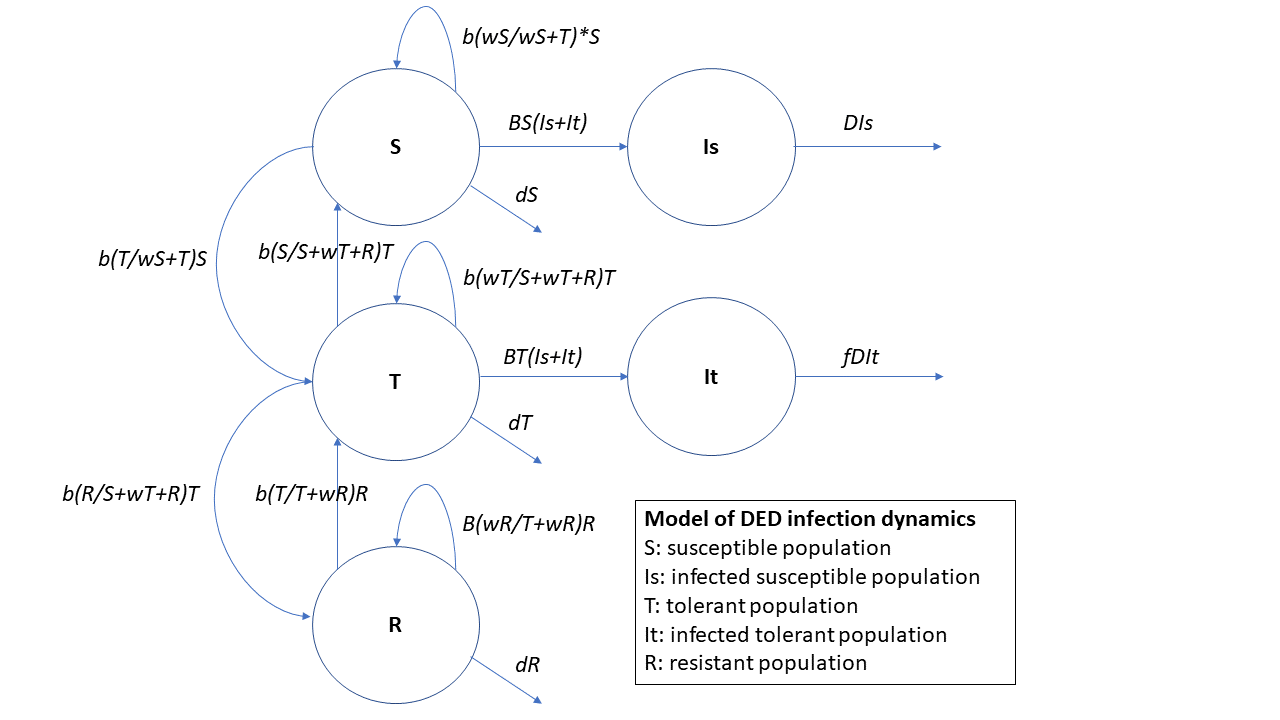
Diploid allele frequency dynamics of a population under selection are represented in Equation 1 in discrete time (Otto and Day, 2007):

Where *p* is the proportion of the dominant allele, A, *q* is the proportion of the recessive allele, a*,* and represent the relative fitness of each genotype. The equation for *q* can be calculated using the formula: 1=p+q. *W* represents the mean fitness of the population, shown in Equation 2:

I explored the allele frequency dynamics of the population for different initial proportions of the resistance allele, *A*, in the population at several increments ranging from 0 to 0.5. Then I test a variety of scenarios of fitness for each genotype: represents a scenario we would expect in the case of major gene resistance, with resistant individuals having the highest fitness; represents a case we might not expect in reality—when tolerant and susceptible individuals have the same fitness; and , where the fitness of each genotype is the same—that is, the population is not under selection (no DED infections). The scenarios of relative fitness strengths are shown in Table 1. I simulated all combinations of these scenarios in R for 50 generations and examined the allele frequencies over time.

Table 1: Simulation values chosen for the relative fitness values:

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 0.1 | 0.1 | 0.1 |
| 0.1 | 0.05 | 0.05 |
| 0.1 | 0.05 | 0.025 |
| 0.1 | 0.05 | 0.01 |
| 0.2 | 0.05 | 0.01 |

The modified SIR model in Figure 1 represents the dynamics for a population undergoing DED infection. There are five categories of the population—susceptible to infection, infected susceptible, tolerant, infected tolerant, and resistant to infection. Susceptible and tolerant individuals have a chance to become infected, modeled as a per capita infection rate. Infected individuals die at a per capita death rate which is higher for infected susceptible individuals compared to infected tolerant individuals (scaled by *f*). In this model, infected individuals do not have the opportunity to recover. Resistant individuals completely avoid infections. Natural births and deaths are modeled as per capita rates for all categories except infected, which I’ve assumed to be expending all their energy towards survival and not reproduction. A proportion of births are selfed (births into the same category), and a proportion of births swap categories in a stepping-stone manner from susceptible to tolerant, tolerant to resistant, and vice versa. Here, selection arises from the susceptible and tolerant categories becoming infected and eventually dying. Eventually, this should result in the entire population to become resistant. No outside intervention is assumed (DED mitigation strategies like vaccination, pruning, fungicides, etc.). The dynamics of this flow diagram is shown in equation form in Equation 3.

*Figure 1: Dynamics of a population undergoing DED infection, modeling susceptible, tolerant, resistant, and infected populations. Infections occur in susceptible and tolerant populations at a rate B. The death rate due to DED is given by D. Natural birth and death rates are given by b and d respectively. Offspring can change compartments from susceptible -> tolerant -> resistant (and vice versa) in steps.*

I modeled this system in R with the parameter values: b=0.15, d=0.1, B=0.5, D=0.5, w=0.5, f=0.5 and initial values of S=0.7, Is=0.01, T=0.25, It=0.01, R=0.03. I calculated solutions for the system using the function ode() in the package *deSolve* (Soetaert et al., 2010), which solves initial value problems for a system of ordinary differential equations. I plotted the simulation results using the function *matplot* to explore the dynamics of the system.

The code for the diploid selection model and the disease dynamics model can be found at: https://github.com/kayleejorose/CSCI2897\_project.

1. **Findings and Results**

***How many generations will it take for resistance to become fixed in the population?***

With strong selection favoring the resistant genotype, the allele can become fixed in 10-20 generations (Figure 2). The stronger selection is, the faster the allele can become fixed (compare rows 3 and 4 of Figure 2). Increasing the initial proportion of resistant genotypes in the population decreases the amount of time to reach fixation by a very small amount (compare columns 1 and 2).

***What is the optimal initial proportion of resistance for the trait to become fixed in a population?***

It is possible for resistance to become fixed in a population with an initial allele frequency of just 1% in the population if there is strong selection (panels C and D). If there is a genotype that is favored through selection (i.e., resistant genotypes favored in a population with DED infections), that allele increases in frequency over time until it is fixed (frequency = 1).

Increasing the initial proportion of the resistant allele in the population makes a small difference in the amount of time it takes for the allele to reach fixation (comparing columns 1 and 2). However, the difference may not be enough to recommend a larger introduction size. Even in a scenario with an initial proportion of 20% resistant allele in the population, it still takes roughly 10 generations to reach fixation (panels G and H).

Graphical user interface, application

Description automatically generated

*Figure 2: A selection of results from the diploid selection model, showing two initial values of resistance in the columns (0.01, 0.2), and rows showing a range of selection strengths. The plots for the rest of the scenarios can be found here: https://github.com/kayleejorose/CSCI2897\_project/tree/main/figures*

***What are the allele frequency dynamics for a population that is not undergoing direct selective pressure?***

A population that is not under selective pressure has constant allele frequency dynamics (panels A and E). That is, the alleles will stay constant at their initial frequencies—no increase or decrease in proportion.

***Population disease dynamics model***

The solutions of the population disease dynamics model in Figure 3 show that over time, the resistant population (blue) increases rapidly. This is because there is selection on the other populations—susceptible and tolerant individuals can become infected and die, whereas resistant individuals cannot. However, it's interesting that tolerant individuals persist over time and continually become infected (yellow and green).

Chart, line chart, histogram

Description automatically generated

*Figure 3: Population disease dynamics for a population of elms with DED infections (solutions from Eq. 3). Here, S=red, Is=orange, T=yellow, It=green, R=blue.*

1. **Discussion and Conclusions**

***Diploid selection model***

These results indicate that genetically based resistance in the form of major gene resistance can become fixed in a population under selection from a disease in roughly 10-20 generations. The amount of time this will take depends on the generation time of the species of interest and on the relative fitness of each genotype. Due to the amount of time it takes for juvenile elm trees to become reproductive (around 15 years), we can expect a slightly longer time lag for resistance to become fixed, perhaps adding on a generation. Resistance can become fixed faster if selection is strong, which should be expected in the case of DED since it is so deadly and fast-acting. If the population has even 1% resistant allele, the trait can become fixed.

Increasing the initial proportion of resistance allows the trait to become fixed slightly faster, but potentially not enough to be recommended in practice, due to the effort, time, and cost it would take to introduce that many individuals in large populations. However, if DED kills off the entire susceptible population before resistance can spread, 1% of the population size remaining may not allow the population to recover due to the lack of genetic diversity. Therefore, potentially a slightly larger proportion of the population (considering logistical constraints) should be introduced resistant individuals (5-10%).

Throughout this project, I realized that major gene resistance may not adequately model the system of DED resistance in elms. Since resistance phenotypes range on a scale of susceptible to resistant with a variety of tolerant phenotypes, it’s likely that resistance is a quantitative trait, resulting from the additive value of multiple genes. If DED resistance is a quantitative trait, then the amount of time for the trait to become fixed would increase, since an individual would need a larger specific set of alleles to have the full resistance phenotype.

Additionally, these results indicate that resistance cannot become fixed in a population that is not under direct selective pressure, for a single-locus unlinked trait. However, resistance is fixed in Asian elm species, and there is no evidence of those species co-evolving with *Ophiostoma*. Therefore, for resistance to become fixed in those populations, it must have either been linked to another trait that was under selection at the time, or major gene resistance is not the best model of genetic DED resistance in elms.

***Population disease dynamics model***

Figure 3 does not create the population dynamics that I expected to see. Resistant individuals increase rapidly on an exponential scale likely because I used the parameters *b > d.* However, when I modeled *b = d*, the resistant category stayed constant (equal to its initial proportion) for the entire timescale. This makes me believe the way I modeled mutations/introgression (offspring switching categories) was not the best way to do it. Changing other parameters like B, D, and w also drastically changed the system dynamics. I’ve only shown the one result in Fig. 3 as it was the closest to what I expected to see. Perhaps more work in determining realistic parameter values is required.

Additionally in Figure 3, tolerant individuals persist in the population which is interesting since they also can become infected with DED and die. However, this is likely because infected tolerant individuals have a lower chance of dying from infection than infected susceptible individuals. In fact, only 25% of the infected tolerant individuals die. In the case of my diploid selection model, tolerant individuals carry one copy of the resistance allele (genotype Aa), so in this system, resistance has almost reached fixation.

***Conclusions***

Taken with the caveats listed above, these results can be applied for other systems with genetically determined disease resistance that is controlled by a single, unlinked locus. These results show a range of optimal proportions of resistant individuals to introduce in a population, and the timeframe for which you can expect resistance to become fixed if the system follows major gene resistance. Even though the model may not be perfect, it still offers a rough time frame for a trait to become fixed, allowing conservation practitioners time to prepare for DED infections.

1. **Outlook and Future Work**

To build on this project, I would try modeling selection on a number of loci instead of just one to model quantitative resistance and compare the results to major gene resistance. That way, I would be able to determine if DED resistance is better modeled as a quantitative trait than a single-locus trait. For example, I could determine if resistance could become fixed as a quantitative trait without selective pressure. If so, resistance is likely quantitative, and if not, resistance must be linked with another trait that is (or was at some point) under selection.

Additionally, I would continue refining the population disease dynamic model shown in Figure 1. The model currently does not represent what I would expect to see in the system, with resistant individuals increasing in frequency over time due to selection, until they reach fixation. I would like to modify the model such that we see resistance fixed at a proportion of 1 in the population instead of exponentially growing to infinity. It is good to see the population growing in spite of DED infections, but this doesn’t seem like it would represent reality. This could be fixed by tweaking the parameters of *b* and *d* or changing the way I’ve modeled individuals switching categories. I noticed that when I changed several of the parameters in the model, the dynamics would change drastically. I would work on creating more realistic parameter values to use for this system.

1. **References**

Brasier, C. (2001). Rapid evolution of introduced plant pathogens via interspecific hybridization: Hybridization is leading to rapid evolution of Dutch elm disease and other fungal plant pathogens. *Bioscience,* 51(2).

Brunet, J. and Gurles, R. (2016). Elm genetic diversity and hybridization in the presence of Dutch Elm Disease. *Proceedings of the American elm restoration workshop*.

Gibbs, J. (1978). Intercontinental epidemiology of Dutch Elm Disease. *Annual Review of Phytopathology*, 16.

Grabowski, M. (2019). Dutch elm disease. *University of Minnesota extension.*

Han, G. (2018). Origin and evolution of the plant immune system. *New Phytologist*, 222(1).

Martin, J., Sobrino-Plata, J., Rodríguez-Calcerrada, J., Collada, C., and Gil, L. (2018). Breeding and scientific advances in the fight against Dutch elm disease: Will they allow the use of elms in forest restoration? *New Forests, 50.*

Otto, S. and Day, T. (2007). Deriving classic models in ecology and evolutionary biology. *A biologist’s guide to mathematical modeling in ecology and evolution.* Princeton University Press.

Potter, K., Escanferla, M., Jetton, R., and Man, G. (2019). Important insect and disease threats to United States tree species and geographic patterns of their potential impacts. *Forests, 10(4).*

Sniezko, R. and Koch, J. (2017). Breeding trees resistant to insects and diseases: putting theory into application. *Forest Invasions,* 19(11).

Soetaert K, Petzoldt T, and Setzer RW. (2010). Solving differential equations in r: Package deSolve. *Journal of Statistical Software*, 33.