Host diet differentially affects virulence in two baculovirus morphotypes

Ari Freedman

Advised by Dr. Greg Dwyer

May, 2020

Abstract

Host diet has been found to have an effect on pathogen performance in many systems, including a wide array of lepidopteran species and their associated baculoviruses. However, the extent to which this host-diet effect differs by viral morphotype in systems infected by multiple baculoviruses is largely unknown. Understanding how different host plants influence viral transmission can be important in predicting the dynamics of local epizootics, and tradeoffs in transmission over these host plants can elucidate the processes regulating morphotype distribution. I tested for a differential effect of host plant on the mortality rate of the Douglas-fir tussock moth (*Orgyia pseudotsugata*) by isolates from two baculovirus morphotypes (*Op*SNPV and *Op*MNPV) applied on foliage from one of two host-tree species (grand fir or Douglas fir). Mortality was significantly higher on Douglas fir and among *Op*MNPV isolates, with a more pronounced difference in mortality between tree species for *Op*MNPV than *Op*SNPV. Generalized linear models and Bayesian hierarchical models fit to the data confirm that morphotype, host-tree species, and morphotype-tree interaction are all significant predictors of mortality. The virus's speed of kill was also affected by host tree and morphotype, indicating that host-tree distribution may have played a role in the morphotypes evolving distinct infection strategies. While this study is concerned only with virulence in individual larvae, discrepancies between these results and those of preliminary field experiments measuring between-larvae transmission in the same system highlight the challenges with inferring large-scale epidemiological processes from isolated lab-measured mechanisms.

Contents

1 Introduction

1.1 General background

Fundamental to the study of wildlife infectious diseases is the problem of predicting the scales and epicenters of epizootics (i.e. disease outbreaks in animals). Insect outbreaks, which are a leading concern for the forestry and agriculture industries (Singh and Satyanarayana 2009), can be closely linked to the outbreaks of infectious diseases within their populations (Dwyer et al. 2000). Thus, understanding the factors affecting when and where outbreaks of both pathogen and host will occur is of utmost importance for creating mathematical models to determine how control measures can be implemented (Dwyer et al. 2005).

One such factor that has received some attention is the role of host diet in disease dynamics, in insects and other animal groups. Traditionally, host-pathogen systems have been modeled without regard to host behavior, including diet (Parker et al. 2010). However, numerous studies have found connections between variability in host diet and unexplained variability in disease dynamics: from birds' nutritional status and the transmission of avian malaria (Cornet et al. 2013); to diet quality of *Daphnia* and the interspecific interactions between their parasites (Lange et al. 2014); to the feeding behaviors of gypsy moth larvae and their risk of baculovirus consumption (Parker et al. 2010).

Baculoviruses, a class of arthropod-specific viruses known mostly to infect lepidopterans (Federici 1997), are environmentally transmitted by host consumption of contaminated foliage. Because insect hosts often feed on multiple plant species, they are a particularly appealing candidate for the study of how host diet can alter viral performance (Dwyer et al. 2005). Differences in mortality of up to 100 times by the same virus have been observed in many lepidopteran-baculovirus systems on different plant types even within a species, often with secondary foliage compounds being the proposed driver of such differences (Ali et al. 2002; Duffey et al. 1995; Hodgson et al. 2002; Hoover et al. 1998; Keating et al. 1988; Shikano et al. 2017).

These secondary compounds are often highly variable between host-plant species and can greatly influence larval gut chemistry, a motivating reason for the study of host-plant effect on virulence (Keating et al. 1988). Indeed, plant chemistry has been shown in various lepidopteran systems to alter the efficacy of host immune system at every step of the infection process, from the sloughing of virus particles out of the gut (Washburn et al. 1998) to the immune response once the

virus has established in the host's bloodstream (Lee et al. 2006).

Most baculovirus research focuses on the control of agricultural pests over non-agricultural defoliators such as the Douglas-fir tussock moth (*Orgyia pseudotsugata*), the focus of this study. The Douglas-fir tussock moth (DFTM) is one of the most destructive North American forest defoliators, capable of killing large swaths of its host trees in outbreak years (Mason and Wickman 1991). Furthermore, DFTM feeds on several different tree species throughout its range with varying foliage chemistry, suggesting that larval host-tree species may play a part in viral transmission (Moore et al. 2000; Schaupp et al. 2008). However, no studies have tested the impact of host-tree species on the efficacy of its associated baculovirus, *Op*NPV.

1.2 Study system

*Op*NPV is a nucleopolyhedrovirus (NPV), a class of obligately-killing baculoviruses defined by a polyhedral protein matrix called the occlusion body, each of which contains thousands of virions and dissolves upon entering its host's alkaline midgut (Federici 1997). Transmission occurs when uninfected larvae consume occlusion bodies on foliage from virus-killed cadavers (Federici 1997). DFTM outbreaks occur cyclically approximately once every ten years, beginning in localized epicenters and leading to widespread defoliation of their host trees (Maclauchlan et al. 2009). The USDA Forest Service regularly uses an *Op*NPV solution called TMB-1 as a means of biocontrol to projected outbreak epicenters (Martignoni 1999; Otvos et al. 1987; Scott and Spiegel 2002).

Two morphotypes, analogous to species in eukaryotes, exist within *Op*NPV: the single-capsid form *Op*SNPV and the multi-capsid *Op*MNPV (written as SNPV and MNPV from now on). Morphotype is categorized by how virions are bundled within an occlusion body and by extensive genotypic differentiation (Hughes and Addison 1970; Rohrmann et al. 1978; Sadler et al. 1998). There is significant genetic variation within a morphotype (Cory et al. 1997), leading to a multitude of strains, or isolates as they are called when sequestered from a host without being fully categorized. Recently, a genotypic cline has been observed in DFTM's range, in which strains in the southwestern US are mostly SNPV, strains in British Columbia are mostly MNPV, and various levels of coexistence arise in between (Williams et al. 2011). Even though MNPV strains seem to dominate over SNPV within a single larva (Hughes 1979), the two often coexist in DFTM populations within a small geographical area (Hughes 1976). This cline largely coincides with the

distribution pattern of DFTM's major host-tree species, with the Douglas fir (*Pseudotsuga menziesii*) overlapping the moth's entire range, grand fir (*Abies grandis*) inhabiting the northern part, and white fir (*Abies concolor*) replacing it in the south (Schaupp et al. 2008). These facts suggest that (1) the morphotypes may have evolved distinct life strategies, allowing for coexistence (de Roode et al. 2008b; Fleming-Davies et al. 2015), and (2) these life strategies may be specific to tree species, with each morphotype best suited to infect on a different tree (de Roode et al. 2008a).

Little work has been done to test for such tradeoffs in performance over a range of diet in any baculovirus system, although Hodgson et al. (2002) showed that two host-plant species differentially affected the fitness of two baculovirus genotypes within a morphotype. And no plant-virus interaction of any kind has been documented for DFTM. Dwyer et al. (2005) found evidence of a host-plant effect on NPV transmission for the closely-related gypsy moth (*Lymantria dispar*, same subfamily as DFTM), which has been linked to its host plants' tannin contents (Elderd et al. 2013). Preliminary results from Dwyer's group have also observed a host-plant effect on transmission for the two NPV morphotypes infecting DFTM. Simulating small-scale epizootics *in situ*, MNPV was found to have higher transmission rates on grand fir than on Douglas fir, while SNPV showed intermediate levels of transmission on both trees.

In this study, I sought to document a host-plant effect on *Op*NPV's mortality rate with two hypotheses: (1) A tree-virus interaction exists between DFTM's host trees and *Op*NPV, differentially affecting the mortality rate of the morphotypes depending on host-tree species; and (2) the two morphotypes have distinct speed-of-kill distributions which also differ by tree species, signifying different life strategies adapted to the geographical distributions of its host-tree species.

2 Methods

2.1 Experimental overview

The preliminary experiments measured transmission rate in the field, which is in fact a complex parameter encompassing many processes (Dwyer et al. 2005). A myriad of factors control the amount of virus in the environment and its ability to infect its host, including larval feeding patterns (Parker et al. 2010), virus avoidance (Capinera et al. 1976; Eakin et al. 2015), branch and needle geometry (Dwyer et al. 2005), outbreak timing and spatial spread (Dwyer and Elkinton

1995; Hunter and Elkinton 2000), and virus breakdown from UV rays or general exposure (Griego et al. 1985).

Once the virus has successfully entered a larva, the next factor underlying transmission is mortality rate, the probability of death given virus consumption (Dwyer et al. 2005), which was the focus of the present study. By confining this study's scope to the narrow process of mortality given consumption, it can hopefully contribute to the overarching basic ecological question of how well small-scale mechanisms can be scaled up to explain large-scale processes such as transmission and long-term epizootic dynamics (Levin 1992; Mihaljevic et al. 2020).

To accurately measure mortality rate for different treatments of morphotype and host-tree species, experiments were conducted in a controlled laboratory setting, with each larva being fed a known viral dose on a single fir-tree needle and then monitored until death. For a more holistic notion of virus-induced mortality, the number of days from infection to death, or speed of kill, was also recorded for each larva. It is generally expected that mortality rate increases and speed of kill (as measured in days) decreases with higher doses, as it becomes more likely that a lethal amount of virions will successfully infect the host (Kennedy et al. 2014; Zwart et al. 2009). To measure variation within morphotype, four viral isolates were tested from each morphotype (Cory et al. 1997).

2.2 Rearing methods

Egg masses were collected in May, 2019, from trees in Sage Hen Reservoir, Idaho, then sterilized in 10% formalin before incubating (Dwyer and Elkinton 1995). Larvae were then raised on artificial diet at room temperature with ambient light and humidity (Chauthani and Claussen 1968). Larvae were reared until the fourth instar, which was chosen to emulate the transmission from neonates infected upon hatching to fourth-instar larvae, a key part of the epizootic (Dwyer and Elkinton 1993). All larval rearing and experiments took place at the Wenatchee Forestry Sciences Laboratory in Wenatchee, Washington.

2.3 Dose-response and speed-of-kill experiments

Along with the commercially available TMB-1, seven viral isolates were used, each one sequestered from an individual infected larva collected in one of seven sites across DFTM's range. Of the eight

isolates, four are of the SNPV morphotype and the other four MNPV, identified by transmission electron microscopy (Table A3). Using a hemocytometer, each isolate was diluted to a viral concentration of 1,800 ± 150 occlusion bodies/μL, which had been previously established as a relatively low dose for fourth instars.

The dose-response experiment consisted of three doses, the two tree species grand fir and Douglas fir, and the eight isolates, as well as control treatments for all doses and both tree species. Control treatments were used primarily to assess the baseline levels of viral contamination in the pre-treatment larvae. Each larva was starved for 24 hours to remove any effects from the artificial diet (Dwyer et al. 2005), then fed a single fir needle with 1, 2, or 3 μL of virus solution pipetted onto it corresponding to low, medium, or high doses, respectively, while control larvae were fed a fir needle with 1 μL of deionized water. Isolating larvae took place over three days, one for each dose, with fir needles collected fresh from live trees every day. For each treatment, needles were pooled together from several trees in an area with very low presence of DFTM larvae and *Op*NPV, to average out any differences between individual trees (Dwyer et al. 2005). Each treatment started with 60 larvae, but larvae which did not eat their entire needle within 36 hours were discarded to ensure consumption of the entire virus dose. The remaining larvae were then reared on artificial diet for a total of 28 days post infection. Dead checks were performed daily starting six days after infection (the minimum time needed for the virus to kill its host from first consumption; Morris 1963), and the speed of kill was recorded as the number of days after infection until death. After 28 days, the remaining larvae or pupae were assumed to have overcome the infection and labeled "survived". Dead larvae were autopsied with a compound microscope, and larvae with visible NPV occlusion bodies were marked "virus-killed".

2.4 Statistical methods

2.4.1 Generalized linear models

I estimated the dose response for each treatment of dose, isolate, and tree species using nested generalized linear models (GLMs) in the R statistical computing language (R Core Team 2019), with each of the 48 experimental treatments representing one data point. To test the relative importance of dose (D), morphotype (M), and tree species (T), the GLMs test every combination of these three predictors and their interaction terms (Table 2). Because dose is a numeric variable while

morphotype and tree species are categorical, any term involving D controls the slopes of the linear fits while all other terms control the intercepts for different treatments. GLMs use the binomial method, which transforms all data to a logistic scale before performing linear regressions, since the response variable, proportion virus-killed, is restricted between 0 and 1. For example, the best model $D \cdot M + M \cdot T$ represents the regression

$$\mathbb{P}\{\text{virus-killed}\} = \text{logit}^{-1}\left[\alpha + \beta_1 \cdot D + \beta_2 \cdot M + \beta_3 \cdot T + \beta_4 \cdot (D \cdot M) + \beta_5 \cdot (M \cdot T)\right], \qquad (1)$$

where D = dose in occlusion bodies; $M = 0$ for SNPV or 1 for MNPV; $T = 0$ for grand fir or 1 for Douglas fir; $\text{logit}(x) = \log(\frac{x}{1-x})$ is the logistic-link function; and α and β_i are fitted parameters.

To compare different GLMs, AIC_c scores were calculated for each model. The standard Akaike information criterion (AIC) compares models with the same likelihood function by penalizing their overall log-likelihoods by the number of parameters being fit, while the AIC_c adds a second-order correction term for models with a small sample size relative to number of parameters (Burnham and Anderson 2002). As with standard AIC, the model with the smallest AIC_c score is considered the most informative model. In Table 2, the AIC_c scores are presented along with Δ_i (the difference in AIC_c score between model i and the best model) and the AIC_c weights w_i (the conditional probabilities for each model being the correct model given the model set; Burnham and Anderson 2002). Models with $\Delta_i < 2$ are considered just as informative as the best model.

2.4.2 Bayesian hierarchical models

The GLMs only consider viral morphotype as a variable, but not the individual isolate within the morphotype. This is because isolates genetically fall within a hierarchy grouped by morphotype (Rohrmann et al. 1978), and so fitting each isolate independently, as GLMs are restricted to do, would be inappropriate. Thus, a reasonable hierarchical model may pull the parameters for each isolate (α, β in Eqs. 2-8) from a distribution whose mean and standard deviation are shared at the morphotype level, meaning that the parameters for every isolate in a morphotype share the same means and standard deviations. However, because there is so much variation in mean mortality between isolates even within a morphotype (Cory et al. 1997), it is more biologically realistic to fit the means of each isolate's linear parameters independently, while having their standard deviations shared at the morphotype level (Mihaljevic et al. 2020). In this way, isolate is a random effect of morphotype, while tree species is assumed to be a fixed effect. Hierarchical models

were implemented in a Bayesian framework through the statistical inference language Stan, which uses a variant of the Hamiltonian Monte Carlo algorithm to sample from parameters' posterior distributions (Stan Development Team 2018).

The Bayesian hierarchical models (BHMs) test the importance of morphotype and tree species with four models, for which parameter means are fit independently for each isolate: one where parameter standard deviations are shared at the morphotype level, and parameter means also vary by tree species ("Morphotype and tree species"); one where parameter standard deviations are shared at the morphotype level, but parameter means do not vary by tree species ("Morphotype only"); one where all parameters share the same standard deviations, and parameter means vary by tree species ("Tree species only"); and one where all parameters share the same standard deviations, but parameter means do not vary by tree species ("Neither morphotype nor tree species").

To further justify the use of a hierarchy, the data were also fit to a null model with both parameter means and standard deviations fit independently for each isolate ("No hierarchy"). And to justify not grouping parameter means by morphotype, another model was tested with both means and standard deviations grouped by morphotype ("Complete hierarchy"). Both of these last two models were based off of the best BHM from the first set, "Morphotype and tree species".

All BHMs are based on the basic underlying logistic-linear equation

$$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha + \beta x_n)\right), \tag{2}$$

where $\text{Bin}(k, p)$ is the binomial distribution with k trials and probability p of success; a treatment n has mortality rate y_n, sample size s_n, and dose x_n; and α and β are fitted parameters whose values, means, and standard deviations vary in different ways for each model:

Morphotype and tree species. For this model, α and β vary by both isolate (i) and tree species (j), and so are indexed as $\alpha_{i,j}$ and $\beta_{i,j}$. Their means also vary by isolate and tree species, written as $\mu_{\alpha,i,j}$ and $\mu_{\beta,i,j}$; but their standard deviations vary only by morphotype (h), and so are written as $\sigma_{\alpha,h}$ and $\sigma_{\beta,h}$. Thus, the model's specifications are

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right) \\ \alpha_{i,j} &\sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_{\alpha,h}) \\ \beta_{i,j} &\sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_{\beta,h}) \end{aligned} \tag{3}$$

(where $\mathcal{N}(\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ).

Morphotype only. This model tests the importance of varying parameters by tree species, by removing tree species as a factor. Here, α, β, and their means vary by isolate only, and their standard deviations vary by morphotype:

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_i + \beta_i x_n)\right) \\ \alpha_i &\sim \mathcal{N}(\mu_{\alpha,i}, \sigma_{\alpha,h}) \\ \beta_i &\sim \mathcal{N}(\mu_{\beta,i}, \sigma_{\beta,h}). \end{aligned} \tag{4}$$

Tree species only. This model tests the importance of varying parameter standard deviations by morphotype, by assuming no distinction between morphotypes. Here, α, β, and their means vary by isolate and tree species, but they all share the same standard deviations:

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right) \\ \alpha_{i,j} &\sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_\alpha) \\ \beta_{i,j} &\sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_\beta). \end{aligned} \tag{5}$$

Neither morphotype nor tree species. This model removes effects of both morphotype and tree species. α, β, and their means vary by isolate only, and they all share the same standard deviations:

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_i + \beta_i x_n)\right) \\ \alpha_i &\sim \mathcal{N}(\mu_{\alpha,i}, \sigma_\alpha) \\ \beta_i &\sim \mathcal{N}(\mu_{\beta,i}, \sigma_\beta). \end{aligned} \tag{6}$$

No hierarchy. This acts as a null model by removing any morphotype hierarchy. It is the same as "Morphotype and tree species", except standard deviations vary by isolate, not morphotype:

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right) \\ \alpha_{i,j} &\sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_{\alpha,i}) \\ \beta_{i,j} &\sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_{\beta,i}). \end{aligned} \tag{7}$$

Complete hierarchy. This model assumes a morphotype hierarchy not only for parameter standard deviations, but for parameter means as well. It modifies "Morphotype and tree species" by varying means by morphotype and trees species, rather than by isolate and tree species as before:

$$\begin{aligned} y_n &\sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right) \\ \alpha_{i,j} &\sim \mathcal{N}(\mu_{\alpha,h,j}, \sigma_{\alpha,h}) \\ \beta_{i,j} &\sim \mathcal{N}(\mu_{\beta,h,j}, \sigma_{\beta,h}). \end{aligned} \tag{8}$$

See Table A1 for a summary of the BHMs. All BHMs were run with 4 chains for 5,000 iterations each. Priors for a BHM's parameter means and standard deviations were chosen to be normally distributed about the estimates and standard errors, respectively, taken from a corresponding GLM's parameter estimates (see Appendix). The predictive accuracies of BHMs were compared using a variant of the Watanabe-Akaike information criterion (WAIC, a Bayesian analog to AIC) called leave-one-out cross validation, or the LOO information criterion (LOOIC). The LOOIC is calculated similarly to WAIC by averaging likelihoods from samples taken from the parameters' posterior distributions, but is more suitable for smaller data sets with highly influential observations (Vehtari et al. 2016). LOOIC scores for all BHMs are given in Table 3.

	SNPV	MNPV	Total
Grand fir	31 / 294 = .105	92 / 342 = .269	123 / 636 = .193
Douglas fir	84 / 359 = .234	214 / 324 = .660	298 / 683 = .436
Total	115 / 653 = .176	306 / 666 = .459	421 / 1319 = .319

Table 1: Overall results from the infection experiments, disregarding dose. For a treatment of certain morphotype (column) and tree species (row), mortality rate is presented as (Number virus-killed) / (Sample size) = (Proportion virus-killed). Totals show the mortality data for all larvae within a morphotype (bottom row), within a tree species (right column), or overall (bottom-right cell).

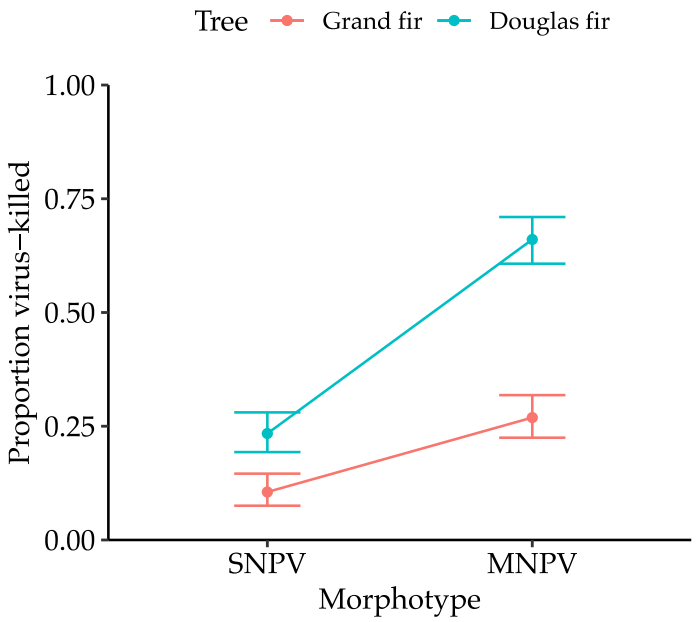

Figure 1: Mortality rates by morphotype and tree species. Error bars represent 95% binomial confidence intervals (Wilson method).

2.4.3 Speed-of-kill distribution models

To formally determine how morphotype and tree species affect the speed at which NPV kills its host, speed-of-kill distributions for each treatment were fit to maximum-likelihood gamma distributions using the R `optim` function. Gamma distributions were chosen for their flexibility in shape and scale, and their ability to mimic the approximate bell shape of the speed-of-kill distributions for a wide range of parameters. Speed-of-kill distributions were fit in four models: (1) one with data grouped all together, (2) one grouped by morphotype only, (3) one grouped by tree species only, and (4) one grouped by morphotype and tree species. The likelihood of each model was calculated as the product of the gamma likelihoods for each data point, from which AIC_c scores were calculated with sample size $N = 421$ (the number of virus-killed larvae; Table 4).

Rank	Model	AIC_c	Δ_i	K	log-likelihood	w_i
1	$D \cdot M + M \cdot T$	334.3	0	6	−160.1	.73
2	$D \cdot M + D \cdot T + M \cdot T$	336.8	2.5	7	−160	.20
3	$D \cdot M \cdot T$	339.3	5.0	8	−159.8	.06
4	$D \cdot M + D \cdot T$	342.9	8.6	6	−164.4	.01
5	$D \cdot T + M \cdot T$	362.2	27.9	6	−174.1	$< 10^{-4}$
6	$D + M + T$	362.9	28.6	4	−177	$< 10^{-4}$
7	$M \cdot T$	372.5	38.2	4	−181.8	$< 10^{-4}$
8	$M + T$	376.3	42.0	3	−184.9	$< 10^{-4}$
9	$D \cdot M$	462.1	127.8	4	−226.6	$< 10^{-4}$
10	$D + M$	483.8	149.5	3	−238.6	$< 10^{-4}$
11	M	492.4	158.1	2	−244.1	$< 10^{-4}$
12	$D + T$	518.2	183.9	3	−255.8	$< 10^{-4}$
13	$D \cdot T$	520.4	186.1	4	−255.8	$< 10^{-4}$
14	T	526.1	191.8	2	−260.9	$< 10^{-4}$
15	D	610.9	276.6	2	−303.3	$< 10^{-4}$
16	1	615.6	281.3	1	−306.7	$< 10^{-4}$

Table 2: AIC_c table for the dose-response GLMs. D, M, and T stand for dose, morphotype, and tree species, respectively. A difference $\Delta_i > 2$ signifies a model is insufficient for explaining the data. $\text{AIC}_\text{c} = -2\log\mathcal{L}(\hat{\theta}) + 2K + \frac{2K(K+1)}{N-K-1}$, where $\hat{\theta}$ is the parameter vector that maximizes the model's likelihood function $\mathcal{L}$, K is the number of parameters, and $N = 48$ is the sample size. Model weight is calculated as $w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_j \exp(-\frac{1}{2}\Delta_j)}$ where the sum is over all models being considered.

3 Results

Virus contamination was low enough in the control treatments, with only 1 virus death out of 74 larvae, to effectively ignore viral contamination from the lab during rearing as a significant factor. Thus, the control treatments and the possibility of contamination were not considered in any subsequent analyses.

Although the dose-response experiments were expected to yield positive relationships between dose and mortality, the data produced mixed results at the isolate level. Only 5 out of 16 treatments displayed strictly increasing dose-response curves, out of which 4 were MNPV treatments (Fig. A1). In general, larvae on Douglas fir tended to have higher mortality than those on grand fir, with 43.6% virus-killed overall on Douglas fir and 19.3% on grand fir; similarly, MNPV isolates tended to be more lethal than SNPV isolates, with 45.9% virus-killed on MNPV and 17.6% on SNPV (Table 1). The difference in mortality between the two tree species was much more pronounced for MNPV than SNPV (Fig. 1). Table A2 reports the mortality rates for treatments at the isolate level.

Table 2 summarizes the relative performances of the GLMs. The best model includes dose-morphotype and morphotype-tree interactions ($w_i = .73$). The next-best model additionally includes a dose-tree interaction ($\Delta_i = 2.5, w_i = .20$). Thus, dose, morphotype, and tree species are all important predictors of mortality, along with dose-morphotype and morphotype-tree interactions, while dose-tree interactions are less likely to impact differences in mortality between tree species. Morphotype appears to be a better predictor of mortality than tree species, as the best eleven models all include morphotype, while only the top eight models include tree species. The

Rank	Model	LOOIC	Δ_i	pLOOIC	log-likelihood	w_i
1	M and T	220.3	0	16.6	−101.8	.84
4	T only	230.3	1.0	23.4	−103.4	.01
5	M only	421.9	201.6	45.2	−188.3	$< 10^{-4}$
6	Neither M nor T	445.3	225.0	60.1	−192.6	$< 10^{-4}$
2	No hierarchy	224.0	3.7	20.9	−101.5	.13
3	Complete hierarchy	227.8	7.5	21.6	−103.1	.02

Table 3: LOOIC table for the BHMs. M and T stand for morphotype and tree species, respectively, in the model names. LOOIC replaces AIC_c , and is calculated as LOOIC $= -2 \log \overline{\mathcal{L}} + 2p$LOOIC, where $\overline{\mathcal{L}}$ is an average likelihood taken over samples from the posterior distributions, and pLOOIC is the effective number of parameters. Weights w_i are calculated as for Table 2.

Rank	Model	AIC_c	Δ_i	K	log-likelihood	w_i
1	M and T	2140.2	0	8	−1061.9	≈ 1
2	M only	2158.7	18.5	4	−1075.3	$< 10^{-4}$
3	T only	2174.3	34.1	4	−1083.1	$< 10^{-4}$
4	Neither M nor T	2182.6	42.4	2	−1089.3	$< 10^{-4}$

Table 4: AIC_c table for the speed-of-kill distribution models. M and T represent morphotype and tree species, respectively, the factors by which the speed-of-kill distributions are grouped for a given model. AIC_c and weights w_i are calculated as for Table 2.

best model predicts higher mortality for MNPV isolates and on Douglas fir, as shown by the positive coefficients in Table 5. The dose-morphotype interaction coefficient β_4 and the tree-species coefficient β_3 are both positive with p-value $< 10^{-4}$, giving statistical significance to the higher mortality for MNPV and Douglas fir. The model predicts a positive slope with dose for all MNPV treatments, consistent with dose-response expectations, but a slightly negative slope for SNPV treatments (Fig. 3a). However, due to the improbability of mortality decreasing with higher viral doses, along with the low sample sizes and mortality in SNPV treatments, it is likely that these negative slopes are the result of statistical noise. Indeed, the estimated dose-response slope for MNPV is significantly positive ($p < 10^{-4}$), while SNPV's slope is insignificantly negative ($p = .17$).

All BHMs converged quickly and were well-mixed, as evidenced by their trace plots and low Gelman-Rubin statistics, and all posteriors appear approximately normal (Fig. A2). The results

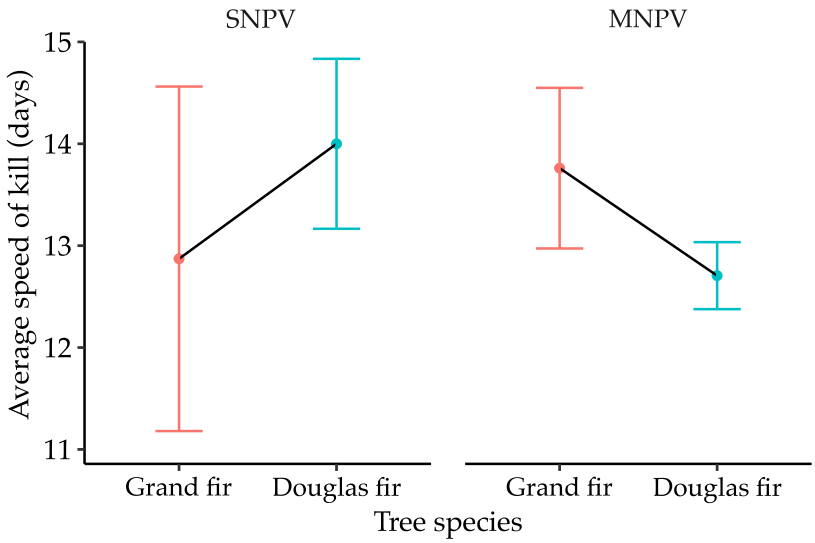

Figure 2: The average number of days from infection to death over all virus-killed larvae, grouped by morphotype and tree species. Points show average speed of kill, and error bars are 95% normal confidence intervals.

from the BHMs are summarized in Table 3. The best model among the BHMs ("Morphotype and tree species", $w_i = .84$) takes both morphotype and tree species into account, while fitting isolates' parameter means separately and standard deviations by morphotype. Compared to the other models with morphotype and tree species as factors ("No hierarchy", $\Delta_i = 3.7, w_i = .13$; and "Complete hierarchy", $\Delta_i = 7.5, w_i = .02$), this further justifies the biologically-based assumptions that (1) a hierarchy in isolates exists, and (2) this hierarchy is more important for parameter standard deviations than parameter means. The other models making these assumptions of hierarchy, but lacking morphotype or tree species as factors, gave significantly worse explanations of the data ($\Delta_i \geq 10.0, w_i \leq .01$). This signifies that both morphotype and tree species are still crucial to predicting mortality in a hierarchical framework. Reassuringly, the best model provides reasonable predictions for most treatments (Fig. 3b).

Point estimates for slope (β) in the best BHM are all positive for MNPV isolates and slightly negative for SNPV isolates (Fig. 3b), similar to the output for the GLMs. The posterior distribution for $\sigma_{\alpha,\text{MNPV}}$ is significantly greater than that of $\sigma_{\alpha,\text{SNPV}}$ by a one-way Welch's t-test ($t = 210.9$, $p < 10^{-4}$), as is the posterior for $\sigma_{\beta,\text{MNPV}}$ over $\sigma_{\beta,\text{SNPV}}$ (but with less significance, $t = 1.7$, $p = .09$), indicating a higher level of variability in mortality among MNPV isolates (Fig. A3).

Negative relationships between dose and average speed of kill were only seen in 3 out of 16 treatments, but this is partly because speed of kill is undefined for treatments with no mortality (Table A2). Thus, speed-of-kill analyses grouped all doses together within treatments. For MNPV, average speed of kill was significantly quicker on Douglas fir than on grand fir (t-test, $t = 2.5$, $p = .02$), but the opposite was true for SNPV ($t = 1.2$, $p = .23$), which had more variation in its speed of kill (Fig. 2). The results from the speed-of-kill distribution models are shown in Table 4, and the best model is illustrated in Fig. 4. Again, the model taking morphotype and tree species into account is the best model ($w_i \approx 1$). This model is followed by the model grouped by morphotype only ($\Delta_i = 18.5, w_i < 10^{-4}$), and then by the model grouped by tree species only ($\Delta_i = 34.1, w_i < 10^{-4}$). Thus, morphotype appears to be a better predictor of speed of kill than tree species.

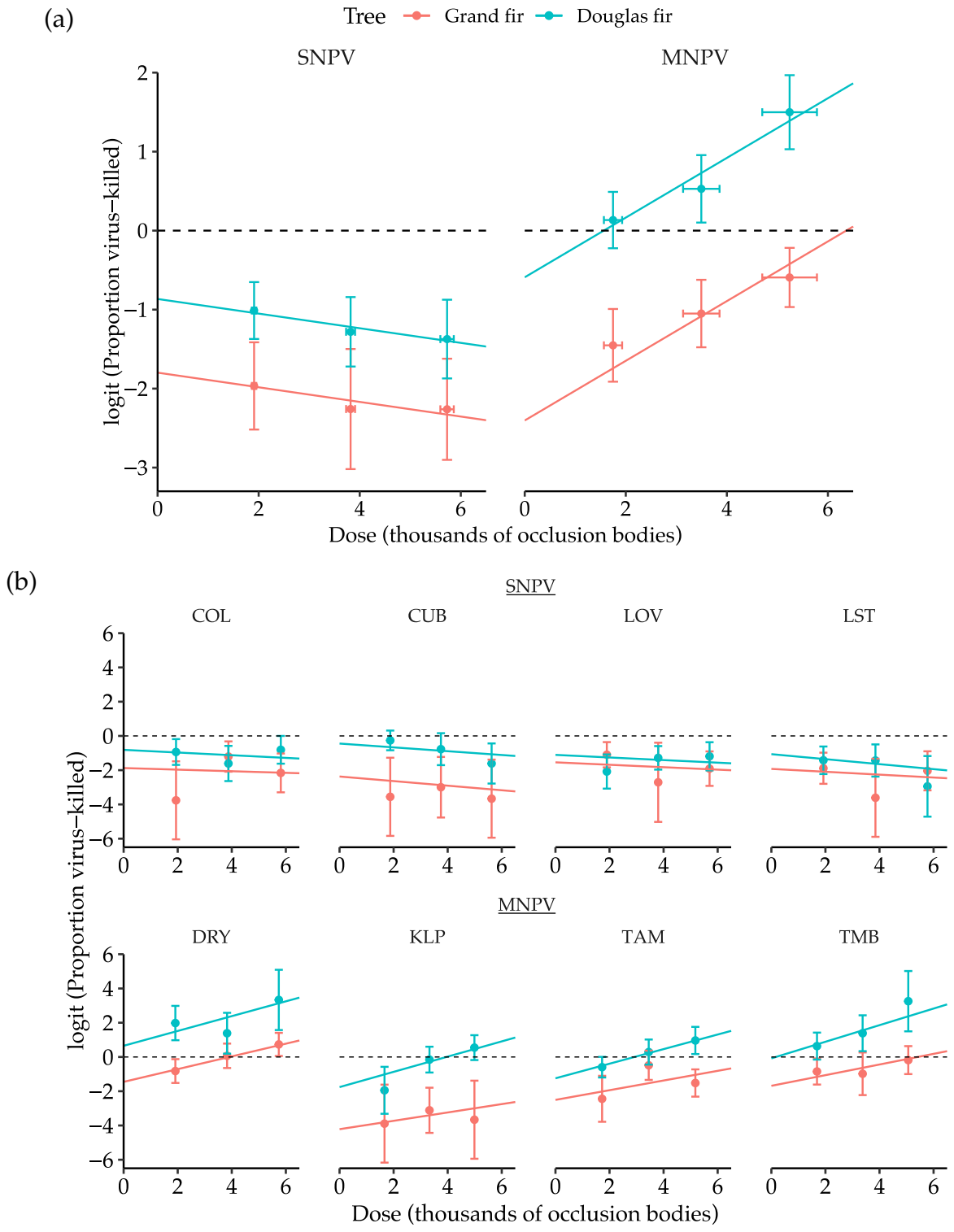

Figure 3: Points show the dose response data on the logistic scale, at the level of (a) morphotype and (b) isolate. The dashed lines at 0 represent 50% mortality. Vertical error bars around a point are 95% binomial confidence intervals for mortality rate, and horizontal error bars are standard 95% normal confidence intervals for mean dose. Lines represent (a) the best GLM, with formula $D \cdot M + M \cdot T$, and (b) the best BHM, "Morphotype and tree species". For treatments with no mortality, it is assumed that a doubled sample size would lead to one virus-killed larva, in order to avoid points at negative infinity.

4 Discussion

As seen in Fig. 1, the tree species on which a DFTM larva feeds, grand fir or Douglas fir, has a noticeable effect on the virulence of NPV in both single- and multi-capsid strains. The results from the GLMs give strong evidence for host-plant influence on NPV performance in the DFTM system (Table 2). The best model's inclusion of a morphotype-tree interaction term supports the existence of a tree-virus interaction, such that tree species affects the virulence of each morphotype in different ways. The model suggests that the virus causes greater mortality on Douglas fir, but the distinction between the two trees is not as great for single-capsid strains as it is for multi-capsid strains. The top model also includes a dose-morphotype interaction, signifying that increasing viral concentration has different effects for SNPV and MNPV. The importance of the dose-morphotype interaction is evident from Fig. 3, where the slopes of the MNPV treatments are strongly positive while the slopes of the SNPV treatments are even slightly negative. Thus, over the range of viral doses studied, a higher concentration of virus in the environment should yield a large increase

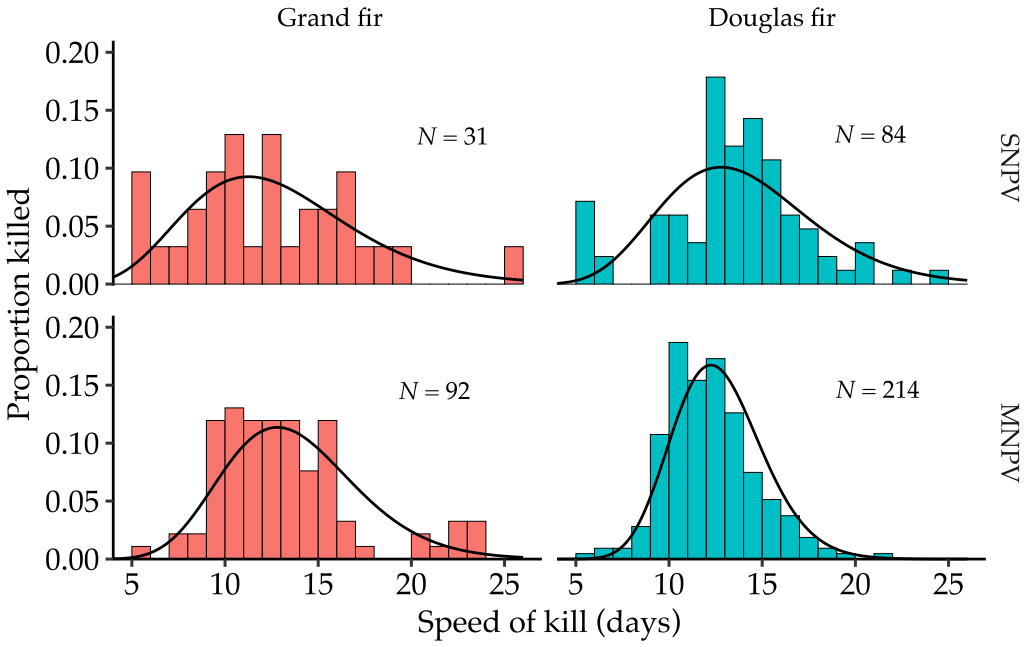

Figure 4: Speed-of-kill distributions grouped by morphotype and tree species. The best speed-of-kill model, consisting of maximum-likelihood gamma distributions grouped by morphotype and tree species, is shown by the black curves over the speed-of-kill distributions. The sample size N is shown for each grouping.

Parameter	α	β_1	β_2	β_3	β_4	β_5
Estimate	-1.78	-9.26×10^{-5}	-0.606	0.933	4.70×10^{-4}	0.880
Std. Error	0.305	6.69×10^{-5}	0.406	0.228	9.03×10^{-5}	0.290

Table 5: Parameter values for the best dose-response GLM, with formula $D \cdot M + M \cdot T$ and parameters given by Eq. 1. α is the linear intercept, and β_1-β_5 are slope parameters for dose, morphotype, tree species, dose-morphotype interaction, and morphotype-tree interaction, respectively.

in mortality due to MNPV, while increasing SNPV concentration would not produce a noticeable effect on mortality. A dose-tree interaction is less important, as the best model does not include one, meaning that the effect of increased virus concentration on mortality should be roughly the same on both host-tree species.

Of course, the negative dose-response curves for SNPV are highly unusual (Kennedy et al. 2014), raising concerns about potential experimental issues. However, I suspect that the negative slopes are the product of statistical noise caused by small sample sizes (Table A2) and a generally large amount of variation in mortality rates among isolates of the same morphotype (Cory et al. 1997). It is possible that the range of doses used in the experiment was too low to produce a significant increase in mortality rate for SNPV. Thompson (1966) found that for aerial application of NPV, a dose of around 13,000 occlusion bodies/μL (over twice as much as the largest dose tested in my experiments) is needed to produce an effective control measure for DFTM outbreaks, and even this is minimal compared to the amount of virus that larvae encounter in an actual epizootic (Thompson 1975). Furthermore, the results from the best model show that the positive slopes for MNPV are statistically significant, while the negative slopes for SNPV are not. Experiments conducted after this study confirm that the viral dose needed to produce a noticeable increase in mortality is much higher for SNPV strains than it is for MNPV. Thus, MNPV produced a significant increase in mortality rate over this range of doses while SNPV did not, further supporting the existence of a dose-morphotype interaction.

Another possible reason for the GLM's negative slopes in SNPV treatments is the presence of highly influential data points, i.e. points that disrupt the monotonically increasing trends within an isolate (Fig. A1). Bayesian models and LOOIC are well suited for handling influential observations such as these (Vehtari et al. 2016), as well as influential isolates within a morphotype for hierarchical models. Even with the large amount of variation within a morphotype, the best

hierarchical model outperformed the analogous model without any hierarchy, indicating that an isolate's dose response is better explained when information from its morphotype is included. The results also imply that an isolate's morphotype helps inform the amount of variation in its virulence, but not its mean mortality—as has been previously assumed for this system (Mihaljevic et al. 2020). Encouragingly, out of the four models based on these assumptions, the model that included both morphotype and tree species as factors performed the best by far and made reasonable predictions for almost all data points (Fig. 3b), further supporting the existence of a tree-virus interaction. And the support for the morphotype hierarchy signifies another way by which SNPV and MNPV differ phenotypically.

However, even in a Bayesian framework, mortality rate on its own does not fully capture all of the processes at play after infection. The time it takes for an infected larva to die and release its occlusion bodies into the environment is potentially an important variable in determining the timing and intensity of epizootics. A quicker speed of kill could lead to more waves of infection in a year (Woods and Elkinton 1987); and the longer the virus takes to kill its host, the more time it will have to replicate, thus releasing more occlusion bodies upon the host's death (Kennedy et al. 2014). Comparing the speed-of-kill distribution models, the one that includes morphotype and tree species is again the best model, indicating that the tree-virus interaction also manifests itself in the time it takes the virus to kill its host. Interesting trends can also be noticed from the average speed of kill in Fig. 2. The morphotypes trade off in their speed of kills over the two tree species, with MNPV killing more quickly on Douglas fir (consistent with the higher mortality on Douglas fir) and SNPV killing more quickly on grand fir. The distinct ways in which speed of kill differs between the two morphotypes depending on host tree, along with the correspondences between the geographical distributions of the morphotypes and tree species, strongly suggests that the morphotypes evolved divergent life strategies in part due to host-tree influence.

To my knowledge, no other baculovirus system has been shown to have a differential response to host diet between multiple morphotypes. The DFTM system is unusually well-suited to such an observation due its uncommon combination of multiple host-tree species and multiple morphotypes within its associated baculovirus, all coexisting over large swaths of its range. These results are consequential in that they provide concrete, unprecedented evidence for the evolution of distinct transmission strategies between two morphotypes of the same baculovirus, which can

help explain their coexistence and their conformation to host-plant distribution patterns.

The reciprocal effect of tree species on average speed of kill in Fig. 2 illustrates a potential tradeoff, which could lead to each morphotype being best suited to infect on a different tree species. And while no such tradeoff is readily apparent from the average mortality rate as seen in Fig. 1, it is possible that the weak dose response for SNPV allows it to outperform MNPV at low doses, or that SNPV has a speed-of-kill distribution better suited to its life history. Indeed, SNPV has lower mortality rates on grand fir, but this could be compensated by its faster average speed of kill on grand fir, resulting in higher SNPV transmission on grand fir. Alternatively, the opposite could be true. Theory posits that obligately-killing pathogens can actually benefit from a longer speed of kill when their mortality rate is lower, as with SNPV, due to the benefit of building up virus particles inside the host that can later be released into the environment (Ebert and Weisser 1997).

4.1 Molecular mechanisms

Until now, the focus has been on showing the existence of host-plant effects on virulence, with no attempt to address the physiological mechanisms causing them. While determining molecular-scale processes is beyond the scope of this study, a few educated hypotheses can be offered based on precedent from the literature.

The higher virulence of MNPV isolates could be partly attributed to MNPV strains having more virions per occlusion body than do SNPV strains (Rohrmann 2014). Although virions were not quantified in this study, Martignoni et al. (1971) found that the average MNPV occlusion body contains about 1,600 virions, compared to only 670 virions per SNPV occlusion body (Hughes 1979). Thus, virion count may actually be a better predictor of virulence than occlusion-body count—though the effect of increased virions per occlusion body is likely non-linear due to host immune responses which slough out entire occlusion bodies (Rohrmann 2014).

Larval diet has been shown in other systems to have an effect at every step of the baculovirus's pathogenesis, starting with occlusion bodies entering the midgut, where the high *p*H dissolves their protective coating and releases their virions (Federici 1997). In the gypsy moth system, host plants with lower *p*H have been shown to lower the *p*H of the midgut, and subsequently render the host less susceptible to infection (Keating et al. 1990).

Once virions have been released, they infect midgut epithelial cells, which can be sloughed as a first response to the viral challenge (Washburn et al. 2001). Virions then have to cross the peritrophic membrane, which is highly variable in its permissivity even within individuals of a species (Adang and Spence 1983; Cory et al. 1997). This sloughing and permissivity can be altered by external compounds (Chen et al. 2018; Washburn et al. 1998), which could potentially include foliage secondary compounds. And because midgut sloughing is a density-dependent response with virus concentration, it is possible that decreased levels of sloughing at low doses can help explain the unexpected dose-response curves for SNPV isolates (Kennedy et al. 2014).

Lastly, once virions have entered a larva's hemolymph (invertebrate equivalent of blood fluid), the larva's hemocytes (blood cells) can remove infected cells (Trudeau et al. 2001; Washburn et al. 1996), a mechanism which has been shown to be crucial in modeling within-host pathogen dynamics and virulence in gypsy moths (Kennedy et al. 2014). It has been shown that differences in a host plant's phenolic compounds (e.g. tannins) can modulate the phenoloxidase cascade underlying the hemocyte's immune response (Lee et al. 2006) and other aspects of NPV's pathogenesis, such as speed of kill (Hoover et al. 1998) or virus production (Ali et al. 2002). Since grand fir and Douglas fir differ appreciably in their concentrations of tannins and other phenolic compounds (Moore et al. 2000), research is needed to investigate how the tree-virus interaction observed in this study is effected at the molecular level by differences in the trees' chemical properties.

The discrepancies between the preliminary transmission data and the mortality rate data for this study may also have a mechanistic basis. Most dramatically, preliminary experiments found transmission rates to be higher on grand fir across morphotypes, while in this study the mortality rates were all higher on Douglas fir (Fig. 1). Many factors can control transmission before infection even occurs. The ones relevant to host-tree species include larval feeding patterns, virus avoidance, and branch geometry, all of which could feasibly contribute to this observed discrepancy. For example, if feeding rates were higher on grand fir (as has been observed) to a sufficient degree, larvae on grand fir would consume virus more quickly and potentially offset the lower mortality rate, leading to higher transmission on grand fir. The differences between transmission and mortality observed for DFTM provide initial evidence that the individual-scale mechanism of mortality cannot be simply extrapolated to the level of between-individual transmission. A more thorough look at the discrepancy between baculovirus transmission and virulence in this system

could be of great benefit to addressing the basic question of how well small-scale mechanisms can explain large-scale processes of disease dynamics.

These results are being shared with the US Forest Service to aid in their DFTM epizootic monitoring efforts. Although further research may alter these conclusions and their applicability to transmission in natural epizootics, I hope that this study may inspire the consideration of host-tree and morphotype distribution in predictive models to more effectively control DFTM outbreaks.

5 Acknowledgements

Many thanks to Greg Dwyer for inspiring this study, and additionally to Katherine Dixon, William Koval, Amy Huang, and Allison Hunter for crucial assistance in experimental setup. Katherine Dixon also provided many valuable insights for the data analysis. This work would not have been possible without the support from the University of Chicago Ecology and Evolution Fellowship as well as the generous access to the US Forest Service's Wenatchee Forestry Sciences Laboratory granted by Carlos Polivka. Thank you to my reviewers Catherine Pfister and J. Timothy Wootton, along with Greg Dwyer and Katherine Dixon, for their constructive comments on earlier drafts.

6 Bibliography

Adang, M. J. and Spence, K. D. 1983. Permeability of the peritrophic membrane of the Douglas fir tussock moth (*Orgyia pseudotsugata*). Comparative Biochemistry and Physiology Part A: Physiology 75:233 – 238.

Ali, M. I., Young, S. Y., Felton, G. W., and McNew, R. W. 2002. Influence of the host plant on occluded virus production and lethal infectivity of a baculovirus. Journal of Invertebrate Pathology 81:158 – 165.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, 2 edition.

Capinera, J. L., Kirouac, S. P., and Barbosa, P. 1976. Phagodeterrency of cadaver components to gypsy moth larvae, *Lymantria dispar*. Journal of Invertebrate Pathology 28:277 – 279.

Chauthani, A. and Claussen, D. 1968. Rearing Douglas-fir tussock moth larvae on synthetic media for the production of nuclear-polyhedrosis virus. Journal of Economic Entomology 61:101–103.

Chen, E., Kolosov, D., O'Donnell, M. J., Erlandson, M. A., McNeil, J. N., and Donly, C. 2018. The effect of diet on midgut and resulting changes in infectiousness of *Ac*MNPV baculovirus in the cabbage looper, *Trichoplusia ni*. Frontiers in Physiology 9:1348.

Cornet, S., Bichet, C., Larcombe, S., Faivre, B., and Sorci, G. 2013. Impact of host nutritional status on infection dynamics and parasite virulence in a bird-malaria system. The Journal of Animal Ecology 83.

Cory, J. S., Hails, R. S., and Sait, S. M. 1997. Baculovirus ecology. In L. K. Miller (ed.), The Baculoviruses. Plenum Press.

de Roode, J. C., Pedersen, A. B., Hunter, M. D., and Altizer, S. 2008a. Host plant species affects virulence in monarch butterfly parasites. Journal of Animal Ecology 77:120–126.

de Roode, J. C., Yates, A., and Altizer, S. 2008b. Virulence-transmission trade-offs and population divergence in virulence in a naturally occuring butterfly parasite. Proceedings of the National Academy of Sciences 105:7489–94.

Duffey, S. S., Hoover, K., Bonning, B. C., and Hammock, B. D. 1995. The impact of host plant on the efficacy of baculoviruses. Reviews in Pesticide Toxicology 3:137–275.

Dwyer, G., Dushoff, J., Elkinton, J. S., and Levin, S. A. 2000. Pathogen-driven outbreaks in forest defoliators revisited: Building models from experimental data. The American Naturalist 156:105–120.

Dwyer, G. and Elkinton, J. 1993. Using simple models to predict virus epizootics in gypsy moth populations. Journal of Animal Ecology 62:1–11.

Dwyer, G. and Elkinton, J. S. 1995. Host dispersal and the spatial spread of insect pathogens. Ecology 76:1262–1275.

Dwyer, G., Firestone, J., and Stevens, T. 2005. Should models of disease dynamics in herbivorous insects include the effects of variability in host-plant foliage quality? The American Naturalist 165:16–31.

Eakin, L., Wang, M., and Dwyer, G. 2015. The effects of the avoidance of infectious hosts on

infection risk in an insect-pathogen interaction. The American Naturalist 185:100–12.

Ebert, D. and Weisser, W. 1997. Optimal killing for obligate killers: The evolution of life histories and virulence of semelparous parasites. Proceedings of the Royal Society B: Biological Sciences 264:985–91.

Elderd, B. D., Rehill, B. J., Haynes, K. J., and Dwyer, G. 2013. Induced plant defenses, host–pathogen interactions, and forest insect outbreaks. Proceedings of the National Academy of Sciences 110:14978–14983.

Federici, B. A. 1997. Baculovirus pathogenesis. In L. K. Miller (ed.), The Baculoviruses. Plenum Press.

Fleming-Davies, A. E., Dukic, V., Andreasen, V., and Dwyer, G. 2015. Effects of host heterogeneity on pathogen diversity and evolution. Ecology Letters 18:1252–1261.

Griego, V. M., Martignoni, M. E., and Claycomb, A. E. 1985. Inactivation of nuclear polyhedrosis virus (*Baculovirus* subgroup A) by monochromatic UV radiation. Applied and Environmental Microbiology 49:709–10.

Hodgson, D. J., Vanbergen, A. J., Hartley, S. E., Hails, R. S., and Cory, J. S. 2002. Differential selection of baculovirus genotypes mediated by different species of host food plant. Ecology Letters 5:512–518.

Hoover, K., Yee, J., Schultz, C., Rocke, D., Hammock, B., and Duffey, S. 1998. Effects of plant identity and chemical constituents on the efficacy of a baculovirus against *Heliothis virescens*. Journal of Chemical Ecology 24:221–252.

Hughes, K. M. 1976. Notes on the nuclear polyhedrosis viruses of tussock moths of the genus *Orgyia* (Lepidoptera). The Canadian Entomologist 108:479–484.

Hughes, K. M. 1979. Some interactions of two baculoviruses of the Douglas-fir tussock moth (Lepidoptera: Lymantriidae). The Canadian Entomologist 111:521–523.

Hughes, K. M. and Addison, R. 1970. Two nuclear polyhedrosis viruses of the Douglas-fir tussock moth. Journal of Invertebrate Pathology 16:196 – 204.

Hunter, F. and Elkinton, S. 2000. Effects of synchrony with host plant on populations of a spring-feeding lepidopteran. Ecology 81:1248–1261.

Keating, S. T., Schultz, J. C., and Yendol, W. G. 1990. The effect of diet on gypsy moth (*Lymantria dispar*) larval midgut *p*H, and its relationship with larval susceptibility to a baculovirus. Journal of Invertebrate Pathology 56:317 – 326.

Keating, S. T., Yendol, W. G., and Schultz, J. C. 1988. Relationship between susceptibility of gypsy moth larvae (Lepidoptera, Lymantriidae) to a baculovirus and host plant foliage constituents. Environmental Entomology 17:952–958.

Kennedy, D. A., Dukic, V., and Dwyer, G. 2014. Pathogen growth in insect hosts: Inferring the importance of different mechanisms using stochastic models and response-time data. The American Naturalist 184:407–423.

Lange, B., Reuter, M., Ebert, D., Muylaert, K., and Decaestecker, E. 2014. Diet quality determines interspecific parasite interactions in host populations. Ecology and Evolution 4.

Lee, K., Cory, J., Wilson, K., Raubenheimer, D., and Simpson, S. 2006. Flexible diet choice offsets protein costs of pathogen resistance in a caterpillar. Proceedings of the Royal Society B: Biological Sciences 273:823–9.

Levin, S. A. 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur Award lecture. Ecology 73:1943–1967.

Maclauchlan, L., Hall, P., Otvos, I., and Brooks, J. 2009. An integrated management system for the Douglas-fir tussock moth in southern British Columbia. Journal of Ecosystems and Management 10.

Martignoni, M. 1999. History of TM BioControl-1: The first registered virus-based product for control of a forest insect. American Entomologist 45:30–37.

Martignoni, M. E., Iwai, P. J., and Breillatt, J. P. 1971. Heterogenous buoyant density in batches of viral nucleopolyhedra. Journal of Invertebrate Pathology 18:219–226.

Mason, R. and Wickman, B. 1991. Integrated pest management of the Douglas-fir tussock moth. Forest Ecology and Management 39:119–130.

Mihaljevic, J. R., Polivka, C. M., Mehmel, C. J., Li, C., Dukic, V., and Dwyer, G. 2020. An empirical test of the role of small-scale transmission in large-scale disease dynamics. The American Naturalist 195:616–635.

Moore, J. A., Mika, P. G., and Shaw, T. M. 2000. Root chemistry of mature Douglas-fir differs by habitat type in the interior Northwestern United States. Forest Science 46:531–536.

Morris, O. N. 1963. A nuclear polyhedrosis of *Orgyia pseudotsugata*: Causative agent and histopathology. Canadian Journal of Microbiology 9:899–900.

Otvos, I. S., Cunningham, J. C., and Friskie, L. M. 1987. Aerial application of nuclear polyhedrosis virus against Douglas-fir tussock moth, *Orgyia pseudostugata* (McDunnough) (Lepidoptera: Lymantriidae). 1. Impact in the year of application. The Canadian Entomologist 119:697–706.

Parker, B. J., Elderd, B. D., and Dwyer, G. 2010. Host behaviour and exposure risk in an insect-pathogen interaction. Journal of Animal Ecology 79:863–870.

R Core Team 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rohrmann, G., McParland, R., Martignoni, M., and Beaudreau, G. 1978. Genetic relatedness of two nucleopolyhedrosis viruses pathogenic for *Orgyia pseudotsugata*. Virology 84:213 – 217.

Rohrmann, G. F. 2014. Baculovirus nucleocapsid aggregation (MNPV vs SNPV): An evolutionary strategy, or a product of replication conditions? Virus Genes 49:351–357.

Sadler, T., Ward, V., Glare, T., and Kalmakoff, J. 1998. Examination of New Zealand's endemic *Wiseana* nucleopolyhedrovirus by analysis of the viral polyhedrin gene. Archives of Virology 143:2273–88.

Schaupp, Willis C., J., Costello, S. L., and Ciesla, W. M. 2008. The history of Douglas-fir tussock moth in Colorado and Wyoming. Technical Report R2-67, USDA Forest Service.

Scott, D. W. and Spiegel, L. 2002. One and two year follow-up evaluation of TM BioControl-1 treatments to suppress Douglas-fir tussock moth in the Blue Mountains for Northeastern

Oregon and Southeastern Washington. Technical Report BMPMSC-02-02, USDA Forest Service, Pacific Northwest Region.

Shikano, I., McCarthy, E. M., Elderd, B. D., and Hoover, K. 2017. Plant genotype and induced defenses affect the productivity of an insect-killing obligate viral pathogen. Journal of Invertebrate Pathology 148:34 – 42.

Singh, T. and Satyanarayana, J. 2009. Insect outbreaks and their management, pp. 331–350. In R. Peshin and A. K. Dhawan (eds.), Integrated Pest Management: Innovation-Development Process: Volume 1. Springer Netherlands, Dordrecht.

Stan Development Team 2018. Stan Modeling Language Users Guide and Reference Manual.

Thompson, C. G. 1966. Tests to determine effectiveness of different virus dosages and concentrations, p. 27. In Results of test to develop operational procedures for controlling Douglas-fir tussock moth with aerial application of polyhedral virus spray. USDA Agricultural Research Service and Forest Service, Portland, OR.

Thompson, C. G. 1975. Comparison of environmental stresses between field and insectary populations, p. 158. In M. D. Summers, R. Engler, L. A. Falcon, and P. V. Vail (eds.), Baculoviruses for Insect Pest Control: Safety Considerations. American Society of Microbiology, Washington, D.C.

Trudeau, D., Washburn, J., and Volkman, L. 2001. Central role of hemocytes in *Ac*MNPV pathogenesis in *Heliothis virescens* and *Helicoverpa zea*. Journal of Virology 75:996–1003.

Vehtari, A., Gelman, A., and Gabry, J. 2016. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27:1413–1432.

Washburn, J., Kirkpatrick, B., Haas-Stapleton, E., and Volkman, L. 1998. Evidence that the stilbene-derived optical brightener M2R enhances *Autographa californica* M nucleopolyhedrovirus infection of *Trichoplusia ni* and *Heliothis virescens* by preventing sloughing of infected midgut epithelial cells. Biological Control 11:58 – 69.

Washburn, J., Kirkpatrick, B., and Volkman, L. 1996. Insect protection against viruses. Nature 383:767.

Washburn, J., Wong, J., and Volkman, L. 2001. Comparative pathogenesis of *Helicoverpa zea* S nucleopolyhedrovirus in noctuid larvae. The Journal of General Virology 82:1777–84.

Williams, H., Monge-Monge, K., Otvos, I., Reardon, R., and Ragenovich, I. 2011. Genotypic variation among Douglas-fir tussock moth nucleopolyhedrovirus (*Op*NPV) isolates in the western United States. Journal of Invertebrate Pathology 108:13–21.

Woods, S. A. and Elkinton, J. S. 1987. Biomodal patterns of mortality from nuclear polyhedrosis virus in gypsy moth (*Lymantria dispar*) populations. Journal of Invertebrate Pathology 50:151 – 157.

Zwart, M., Hemerik, L., Cory, J., de Visser, J. A., Bianchi, F., Van Oers, M., Vlak, J., Hoekstra, R., and Van der Werf, W. 2009. An experimental test of the independent action hypothesis in virus-insect pathosystems. Proceedings of the Royal Society B: Biological Sciences 276:2233–42.

A Appendix

A.1 Code and data availability

All code and data can be found at https://github.com/freedmanari/undergrad-thesis.

A.2 Prior construction

For those models taking both morphotype and tree species into account—"Morphotype and tree species" (the best model), "No hierarchy", and "Complete hierarchy"—priors were taken from the GLM with all interaction terms, $D \cdot M \cdot T$. The reason for choosing the most complicated GLM to represent these models is that they assign different intercepts and slope parameters to each combination of isolate and tree species, and a GLM can only accomplish this with a triple interaction term. Mean parameter values, the μ_α and μ_β, are taken from the GLM's predictions before being transformed by the logistic-link function. For example, if the GLM is written as

$$\begin{aligned}\mathbb{P}\{\text{virus-killed}\} = \text{logit}^{-1}\big[&\alpha + \beta_1 \cdot D + \beta_2 \cdot M + \beta_3 \cdot T + \beta_4 \cdot (D \cdot M) + \beta_5 \cdot (M \cdot T) \\ &+ \beta_6 \cdot (D \cdot T) + \beta_7 \cdot (D \cdot M \cdot T)\big] \end{aligned} \tag{A1}$$

(where $M = 0$ or 1 for SNPV or MNPV, and $T = 0$ or 1 for grand or Douglas fir), then

$$\mu_{\alpha,\text{DRY,GF}} = \alpha + \beta_2, \tag{A2}$$

since dose is not considered for α intercept terms ($D = 0$), DRY belongs to the MNPV morphotype ($M = 1$), and GF means grand fir ($T = 0$). On the other hand,

$$\mu_{\beta,\text{TMB,DF}} = \beta_1 + \beta_4 + \beta_6 + \beta_7, \tag{A3}$$

since β slope terms only consist of GLM terms relating to dose, TMB belongs to the MNPV morphotype (M=1), and DF means Douglas fir (T=1). If the μ_α and μ_β are considered to be morphotype-level parameters instead, as they are for the "Complete hierarchy" model, then they are indexed by morphotype rather than isolate, and so $\mu_{\alpha,\text{DRY,GF}}$ in Eq. A2 and $\mu_{\beta,\text{TMB,DF}}$ in Eq. A3 would be replaced by $\mu_{\alpha,\text{MNPV,GF}}$ and $\mu_{\beta,\text{MNPV,DF}}$, respectively. In this case, since parameter means are modeled hierarchically, the μ_α and μ_β inform additional intermediate mean parameters η_α and η_β, such that $\eta_{\alpha,h,j} \sim \mathcal{N}(\mu_{\alpha,h,j}, |\mu_{\alpha,h,j}|)$ and $\eta_{\beta,h,j} \sim \mathcal{N}(\mu_{\beta,h,j}, |\mu_{\beta,h,j}|)$.

For these models, the σ_α and σ_β, indexed by morphotype, take their corresponding prior means, τ_α and τ_β, from the same GLM. Let $\sigma_{\text{SNPV}}(x)$ and $\sigma_{\text{MNPV}}(x)$ be the standard deviations for GLM parameter x in a version of Eq. A1 where either SNPV is the base morphotype ($M = 0$ for SNPV

and 1 for MNPV, the default) or MNPV is the base morphotype ($M = 0$ for MNPV and 1 for SNPV). Then the τ_α and τ_β were taken to be

$$\begin{aligned}
\tau_{\alpha,\text{SNPV}} &= \sqrt{\left(\sigma_{\text{SNPV}}(\alpha)\right)^2 + \left(\sigma_{\text{SNPV}}(\beta_3)\right)^2} \\
\tau_{\alpha,\text{MNPV}} &= \sqrt{\left(\sigma_{\text{MNPV}}(\alpha)\right)^2 + \left(\sigma_{\text{MNPV}}(\beta_3)\right)^2} \\
\tau_{\beta,\text{SNPV}} &= \sqrt{\left(\sigma_{\text{SNPV}}(\beta_1)\right)^2 + \left(\sigma_{\text{SNPV}}(\beta_6)\right)^2} \\
\tau_{\beta,\text{MNPV}} &= \sqrt{\left(\sigma_{\text{MNPV}}(\beta_1)\right)^2 + \left(\sigma_{\text{MNPV}}(\beta_6)\right)^2},
\end{aligned} \tag{A4}$$

in order to take into account the variability in parameters from both tree species. The parameters $\sigma_{\alpha,h}$ and $\sigma_{\beta,h}$ were then taken to have prior distributions that were normal about $\tau_{\alpha,h}$ and $\tau_{\beta,h}$, such that $\sigma_{\alpha,h} \sim \mathcal{N}(\tau_{\alpha,h}, 3\tau_{\alpha,h})$ and $\sigma_{\beta,h} \sim \mathcal{N}(\tau_{\beta,h}, 3\tau_{\beta,h})$, truncated to be positive. The multiplier 3 for the standard deviations was chosen by experimenting with a range of multipliers and picking the one that gave the best average LOOICs for all BHMs tested.

The other BHMs—"Morphotype only", "Tree species only", and "Neither morphotype nor tree species"—took their priors in an analogous manner from GLMs with the formulae $D \cdot M$, $D \cdot T$, and D, respectively.

The BHMs and their priors are summarized in Table A1.

A.3 Supplementary tables and figures

Morphotype and tree species	Morphotype only
$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right)$	$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_i + \beta_i x_n)\right)$
$\alpha_{i,j} \sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_{\alpha,h})$ $\beta_{i,j} \sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_{\beta,h})$	$\alpha_i \sim \mathcal{N}(\mu_{\alpha,i}, \sigma_{\alpha,h})$ $\beta_i \sim \mathcal{N}(\mu_{\beta,i}, \sigma_{\beta,h})$
$\sigma_{\alpha,h} \sim \mathcal{N}(\tau_{\alpha,h}, 3\tau_{\alpha,h})$ $\sigma_{\beta,h} \sim \mathcal{N}(\tau_{\beta,h}, 3\tau_{\beta,h})$	$\sigma_{\alpha,h} \sim \mathcal{N}(\tau_{\alpha,h}, 3\tau_{\alpha,h})$ $\sigma_{\beta,h} \sim \mathcal{N}(\tau_{\beta,h}, 3\tau_{\beta,h})$
Tree species only	**Neither morphotype nor tree species**
$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right)$	$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_i + \beta_i x_n)\right)$
$\alpha_{i,j} \sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_{\alpha})$ $\beta_{i,j} \sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_{\beta})$	$\alpha_i \sim \mathcal{N}(\mu_{\alpha,i}, \sigma_{\alpha})$ $\beta_i \sim \mathcal{N}(\mu_{\beta,i}, \sigma_{\beta})$
$\sigma_{\alpha} \sim \mathcal{N}(\tau_{\alpha}, 3\tau_{\alpha})$ $\sigma_{\beta} \sim \mathcal{N}(\tau_{\beta}, 3\tau_{\beta})$	$\sigma_{\alpha} \sim \mathcal{N}(\tau_{\alpha}, 3\tau_{\alpha})$ $\sigma_{\beta} \sim \mathcal{N}(\tau_{\beta}, 3\tau_{\beta})$
No hierarchy	**Complete hierarchy**
$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right)$	$y_n \sim \text{Bin}\left(s_n, \text{logit}^{-1}(\alpha_{i,j} + \beta_{i,j} x_n)\right)$
$\alpha_{i,j} \sim \mathcal{N}(\mu_{\alpha,i,j}, \sigma_{\alpha,i})$ $\beta_{i,j} \sim \mathcal{N}(\mu_{\beta,i,j}, \sigma_{\beta,i})$	$\alpha_{i,j} \sim \mathcal{N}(\eta_{\alpha,h,j}, \sigma_{\alpha,h})$ $\beta_{i,j} \sim \mathcal{N}(\eta_{\beta,h,j}, \sigma_{\beta,h})$
$\sigma_{\alpha,i} \sim \mathcal{N}(\tau_{\alpha,i}, 3\tau_{\alpha,i})$ $\sigma_{\beta,i} \sim \mathcal{N}(\tau_{\beta,i}, 3\tau_{\beta,i})$	$\eta_{\alpha,h,j} \sim \mathcal{N}(\mu_{\alpha,h,j}, \lvert\mu_{\alpha,h,j}\rvert)$ $\eta_{\beta,h,j} \sim \mathcal{N}(\mu_{\beta,h,j}, \lvert\mu_{\beta,h,j}\rvert)$
	$\sigma_{\alpha,h} \sim \mathcal{N}(\tau_{\alpha,h}, 3\tau_{\alpha,h})$ $\sigma_{\beta,h} \sim \mathcal{N}(\tau_{\beta,h}, 3\tau_{\beta,h})$

Table A1: The Bayesian hierarchical models. Models above the double line are those for which only parameter standard deviations are specified hierarchically, as suggested by the virus's biology, with morphotype and tree species as potential factors. α, β are the linear-fit parameters; μ_α, μ_β are their means; σ_α, σ_β are their standard deviations; and τ_α, τ_β are the prior means for σ_α, σ_β. For a treatment $n \in \{1, \ldots, 48\}$, y_n is the number virus-killed, s_n is the sample size, and x_n is the viral dose. All instances of h, i, or j represent indexing by morphotype, isolate, or tree species, respectively. Models below the double line test the hierarchy assumptions, with either no morphotype hierarchy ("No hierarchy") or a hierarchy for parameter means as well as standard deviations ("Complete hierarchy"). η_α, η_β are additional intermediate mean parameters for the "Complete hierarchy" model, to simulate the hierarchy in standard deviations. These last two models are based off of the "Morphotype and tree species" model from above the double line, with differences shown in red.

	SNPV, Grand fir											
	COL			CUB			LOV			LST		
Dose	1,938	3,876	5,814	1,875	3,750	5,625	1,900	3,800	5,700	1,925	3,850	5,775
Virus-killed	0	6	3	0	1	0	9	0	4	5	0	3
Total	22	26	29	18	21	20	36	8	31	38	19	26
Average SOK	NA	8.7	16.0	NA	8.0	NA	12.8	NA	19.25	12.4	NA	12.3
	SNPV, Douglas fir											
	COL			CUB			LOV			LST		
Dose	1,938	3,876	5,814	1,875	3,750	5,625	1,900	3,800	5,700	1,925	3,850	5,775
Virus-killed	9	4	8	20	6	3	4	10	7	7	5	1
Total	32	24	26	46	19	18	36	46	30	36	26	20
Average SOK	13.1	14.8	13.8	14.7	13.5	14.7	14.0	14.0	10.6	15.1	15.8	16.0
	MNPV, Grand fir											
	DRY			KLP			TAM			TMB		
Dose	1913	3826	5739	1663	3326	4989	1725	3450	5175	1688	3376	5064
Virus-killed	11	15	25	0	2	0	2	8	7	9	3	10
Total	36	29	37	25	47	20	25	21	39	30	11	22
Average SOK	12.8	15.0	13.6	NA	10.0	NA	8.0	17.4	15.7	12.3	13.3	12.5
	MNPV, Douglas fir											
	DRY			KLP			TAM			TMB		
Dose	1913	3826	5739	1663	3326	4989	1725	3450	5175	1688	3376	5064
Virus-killed	29	12	28	2	12	19	16	16	21	17	16	26
Total	33	15	29	16	26	30	45	28	29	26	20	27
Average SOK	11.7	12.1	12.4	14.0	13.4	13.9	14.4	13.9	13.4	12.9	12.5	10.8

Table A2: Mortality rate and average speed of kill (SOK) for every treatment, each of which consists of one of eight isolates, one of two tree species, and one of three doses. Dose is given in number of occlusion bodies, Virus-killed refers to the number of virus-killed larvae in a treatment, Total is the sample size of the treatment, and Average SOK is the average number of days from infection to death for virus-killed larvae in a treatment (or NA for treatments with no mortality). See Fig. A1 for a graphical representation of the mortality rate data.

Isolate	Morphotype	Site Name	State	Latitude	Longitude
COL	SNPV	Cheyanne Mountain	CO	38.739185	-104.880892
CUB	SNPV	Cub Creek	WA	47.082654	-121.231302
LOV	SNPV	Lovell Valley	ID	47.307357	-116.980771
LST	SNPV	Lost River	WA	48.65235	-120.50683
DRY	MNPV	Dry Camp	NM	35.215163	-106.407929
KLP	MNPV	Klipchuck Campground	WA	48.597559	-120.513011
TAM	MNPV	Tamarack Flats	ID	44.38744922	-116.2111798
TMB	MNPV	Tussock Moth Biocontrol-1	CA	41.964121	-120.421626

Table A3: Collection location and morphotype for each viral isolate used in the infection experiments.

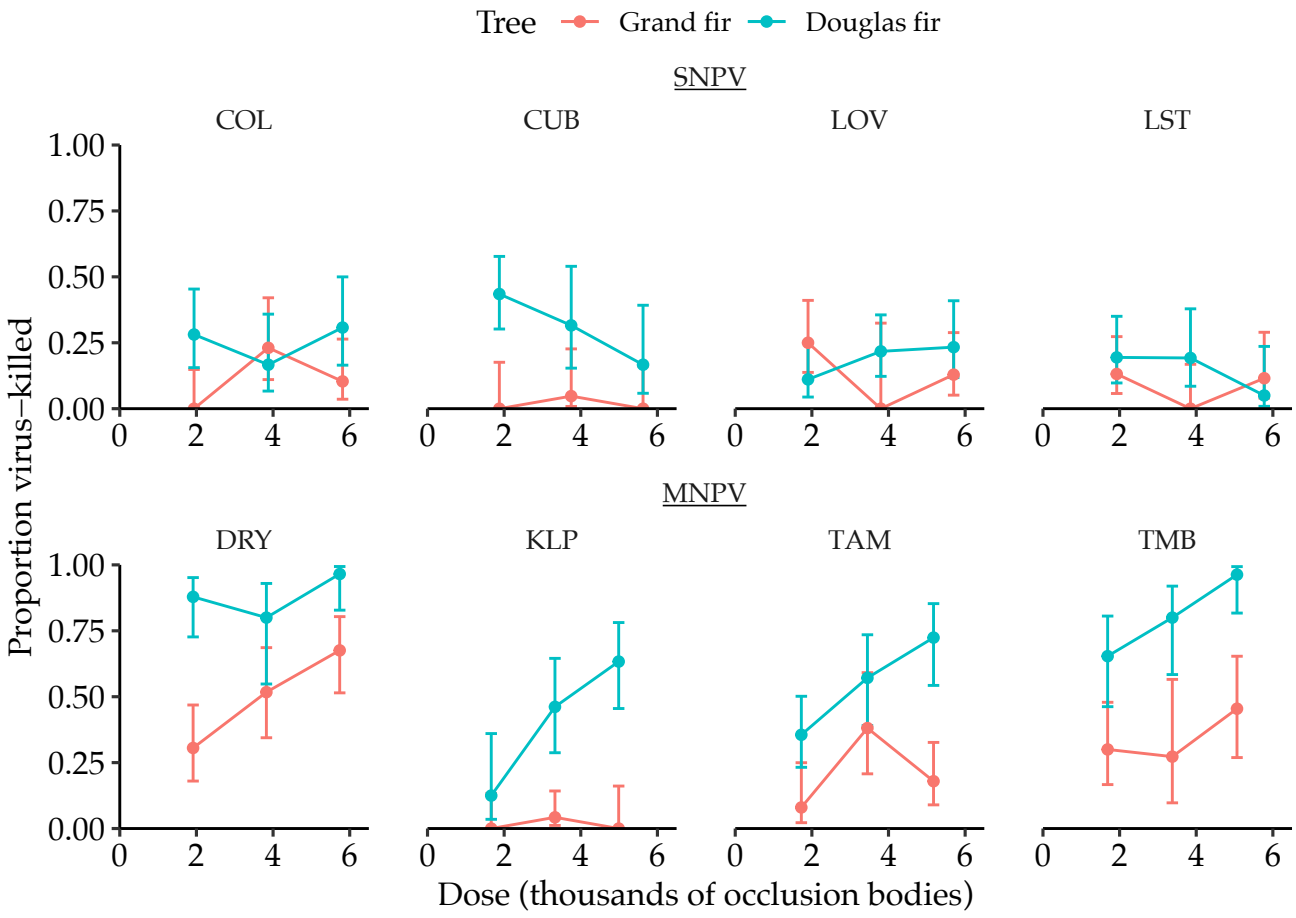

Figure A1: Dose-response data for each isolate. Each point represents the mortality rate for a single treatment. Error bars show 95% binomial confidence intervals.

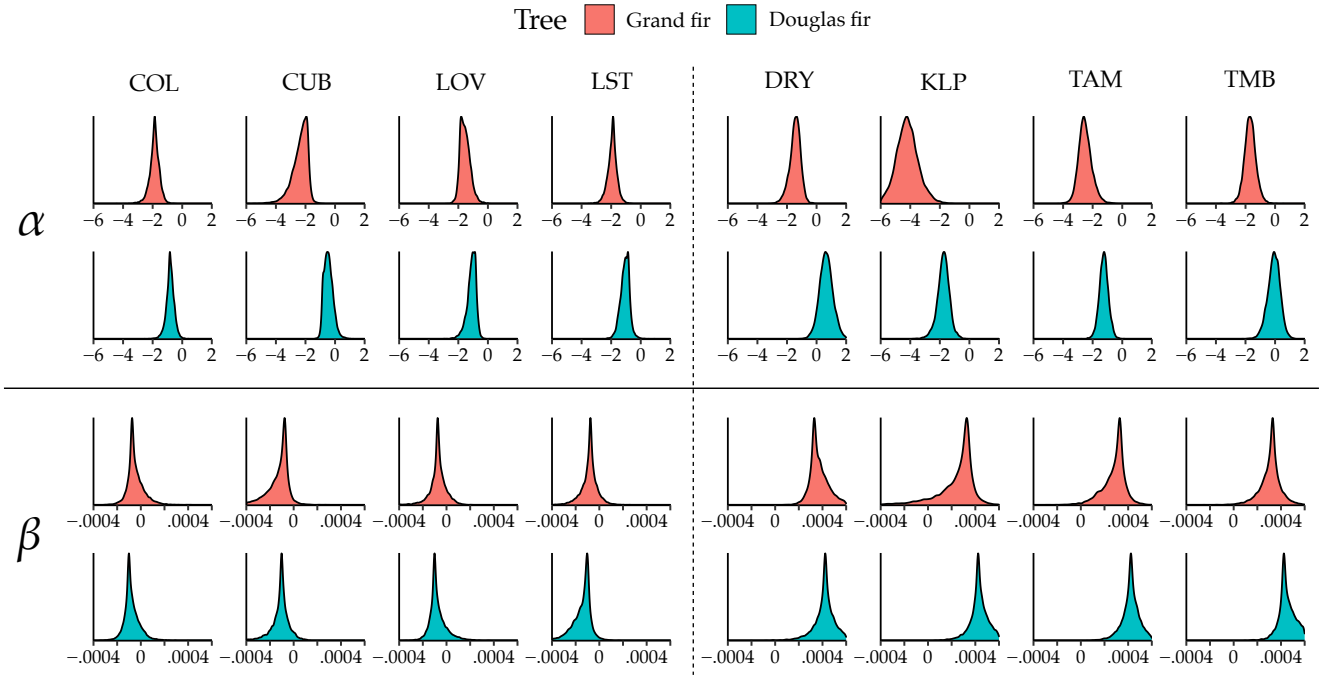

Figure A2: Posterior distributions for all linear-fit parameters $\alpha_{i,j}$ and $\beta_{i,j}$ from the best BHM ("Morphotype and tree species"), indexed by isolate (columns) and tree species (rows/color). The top two rows are the α parameters while the bottom rows are the β parameters, and columns on the left are for SNPV isolates while columns on the right are MNPV.

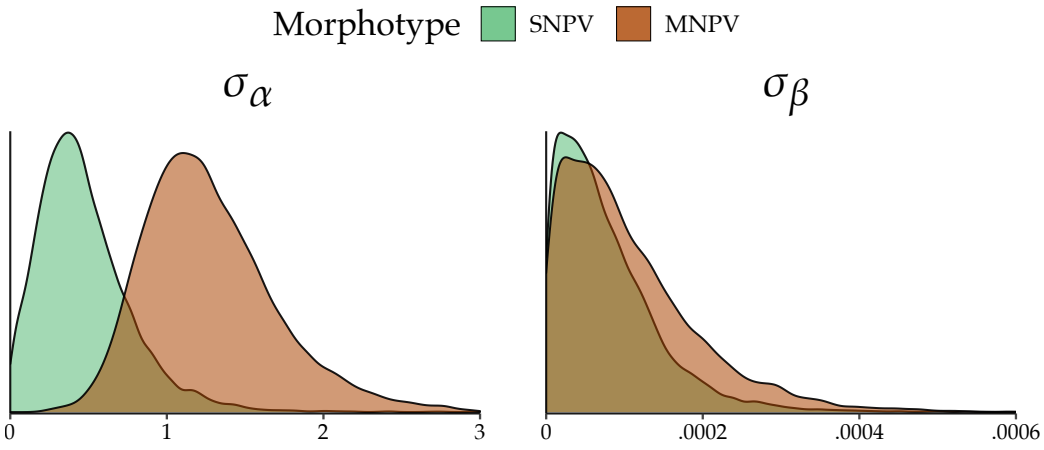

Figure A3: Posterior distributions for $\sigma_{\alpha,h}$ and $\sigma_{\beta,h}$, the standard deviations for $\alpha_{i,j}$ and $\beta_{i,j}$, respectively, from the best BHM ("Morphotype and tree species"). Posteriors for $\sigma_{\alpha,\text{SNPV}}$ and $\sigma_{\beta,\text{SNPV}}$ are shown in green, while $\sigma_{\alpha,\text{MNPV}}$ and $\sigma_{\beta,\text{MNPV}}$ are shown in brown.