Local adaptation to biocontrol agents: A multi-objective data-driven optimization model for the evolution of resistance

L. Stevens a,*, D.M. Rizzo b,1

a University of Vermont, Department of Biology, 321 Marsh Life Science Building, Burlington, VT 05405, USA
b University of Vermont, Department of Civil & Environmental Engineering, 219 Votey Building, Burlington, VT 05405, USA

ARTICLE INFO

Article history:
Received 27 August 2007
Received in revised form 13 March 2008
Accepted 25 April 2008
Published on line 12 June 2008

Keywords:
Biological control
Evolution
Resiliency
Meta-population model
Data-driven model
Multi-objective optimization

ABSTRACT

Spatial and temporal variability in the application of biological control agents such as parasites or pathogenic bacteria can cause the evolution of resistance in pest organisms. Because biocontrol will be more effective if organisms are not resistant, it is desirable to examine the evolution of resistance under different application strategies.

We present a computational method that integrates a genetic algorithm with experimental data for predicting when local populations are likely to evolve resistance to biocontrol pathogens. The model incorporates parameters that can be varied as part of pest control measures such as the distribution and severity of the biocontrol agent (e.g., pathogenic fungi). The model predicts the evolution of pathogen defense as well as indirect selection on several aspects of the organism’s genetic system. Our results show that both variability of selection within populations as well as mean differences among populations are important in the evolution of defenses against biocontrol pathogens. The mean defense is changed through the pest organism’s genotype and the variance is affected by components of the genetic system, namely, the resiliency, recombination rate and number of genes.

The data-driven model incorporates experimental data on pathogen susceptibility and the cost of defense. The results suggest that spatial variability rather than uniform application of biological control will limit the evolution of resistance in pest organisms.

#2008 Elsevier B.V. All rights reserved.

1. Introduction

Biocontrol agents such as pathogenic fungi (Tucker and Stevens, 2003) and bacteria (Medvinsky et al., 2007) are often used to control pest organisms such as insects. Models exist for the effect of pathogens on diversity both within and between species (Abell et al., 2005; Laird and Jensen, 2006) and other models consider spatial variation (Medvinsky et al., 2004; Shirley and Rushton, 2005). However, the pest species can evolve resistance to biocontrol agents. In this paper, we examine the evolution of such resistance, along with correlated traits that might co-evolve under different patterns of environmental variation. We develop a model using data on the sensitivity of Japanese beetles, Popilla japonica to the entomopathogenic fungus Metarhizium anisopliae. Insects commonly defend themselves using melanin to encapsulate organisms that breach the cuticle by activation of the enzyme PO. We used data from four populations of Japanese beetles from across Vermont: two were from locations using the bacterial pathogen B. thuringiensis as agents of biological pest control, and two locations did not use biocontrol. Populations from areas with biocontrol exhibited greater PO production and suffered less mortality from pathogen exposure; however, PO levels positively correlate with increased mortality in the absence of parasitism (Tucker and Stevens, 2003). The Japanese beetle system is provides a good model for examin-
Biochemical defenses, such as PO, are one of several means of pathogen defense. Other means of defense include: immune, life history, behavioral and structural defenses. Organisms also have several defense mechanisms including continuous/inducible expression of defenses, qualitative changes in elicitors/receptors, and increasing/decreasing defense quantity (Rigby et al., 2002). Immunity includes both cellular components (e.g., phagocytic blood cells to defend against general pathogens such as bacteria) and the adaptive immune system (e.g. the antibodies generated by vaccines). Only vertebrates have adaptive immunity; however most organisms, including insects, have biochemical and cellular defenses, such as the common biochemical defense involving the activation of the enzyme PO. In several species, quantity of PO is correlated with parasite resistance (e.g., Gillespie et al., 2000) and Japanese beetle populations that vary in previous exposure to biocontrol vary in PO quantity (Tucker and Stevens, 2003). When biochemical or cellular defenses depend on the quantity of the defense, they are likely to have constraints based on physiological, energetic, genetic, and/or hormonal correlations (Zera and Harshman, 2001); for example, producing high quantities of PO can use energy that might otherwise be used for reproduction. Japanese beetle populations with high quantities of PO exhibit a trade-off between high quantities of defense and other aspects of fitness (Tucker and Stevens, 2003). Because trade-offs can exist for pathogen defense, spatial variation in susceptibility to pathogens can evolve as a result of biocontrol efforts making the evolution of pathogen defense a complex optimization problem.

Pathogen models have examined uniform and variable environmental effects (Hilker et al., 2006). When selection varies on a spatial scale, adaptation may include not only changes in the mean value of a trait, but also aspects of the variability of the trait; and thus, adaptation is most realistically modeled using multiple traits and a multi-objective optimization approach. Multi-objective optimization problems can be investigated using analytical techniques; however, for more complex situations analytical solutions do not exist. The data-driven nature of evolutionary algorithms make them ideal for analyzing complex multi-objective optimization problems and have many advantages over other optimization techniques for modeling biological adaptation (Foster, 2001) as presented in the following sections.

1.1. Self-adaptation: indirect selection

A feature of evolutionary computation that mimics biological processes is that parameters can “self-adapt” or undergo indirect selection (Eiben et al., 1999). Self-adaptation or the “evolution of evolution” occurs when several parameters are incorporated into the evolving fitness of an individual thereby making them subject to evolution.

Adaptation to biocontrol agents can occur by optimizing the mean and variance of a trait. When biocontrol application is unpredictable, it may be advantageous for organisms to produce offspring with a range of phenotypes. In the model presented here, we simulate the simultaneous evolution of the mean host defense genotype, as well as two parameters of the organism’s genetic system, recombination rate and canalization, that affect trait variability.

1.2. Host defense genotype

Based on empirical studies of plant and animal defenses (Kolb et al., 2001; Pilet et al., 2001; Schneider et al., 2001; Wu et al., 2001), we modeled quantitative (Falconer and Mackay, 1995), polygenic defense genotypes. Most models incorporate haploid organisms (i.e., those having unpaired chromosomes such as bacteria); however, artificial life experiments with haploid organisms that undergo sexual reproduction show that sexual reproduction profoundly influences the evolution of genetic architecture (Misevic et al., 2006). Our model extends previous studies by considering diploid inheritance, which is the more realistic case for animals, plants and fungi.

1.3. Recombination rate

The number of chromosomes, 2n (two copies of each gene, one from the female parent and one from the male), varies among organisms. We based our model on data from a scarab beetle and n = 10, 2n = 20 for most of this group. If recombination (crossing over between homologous chromosomes during the process of gamete formation when cells go from 2n to n) does not occur, the entire intact chromosome inherited from either the female parent or the male parent will end up in the sex cell. With recombination, a novel, hybrid chromosome will result.

The rate of recombination affects the variability of offspring. Recombination rates vary among sexes, populations and species, and can evolve by indirect selection (Uyenoyama and Bengtsson, 1989). Recombination among genes that affect the same trait, such as the multiple defense genes examined in our simulation, affects the ability to form adaptive gene complexes (i.e., genes that work together to perform a particular function) (Palopoli and Wu, 1996). Models have shown that parasites select for recombination in haploid hosts (Hamilton et al., 1990), presumably because recombination generates variability allowing a host to escape attack.

1.4. Canalization (i.e., resiliency)

Seventy percent of the variation in PO activity of the caterpillar Spodoptera littoralis is estimated to result from environmental factors (Cotter and Wilson, 2002). Several authors have addressed the ability of genotypes, genetically identical at fertilization, to become variable during development (Kaplan and Cooper, 1984; Boldogkoi, 2004) and distinguish between phenotypic plasticity, canalization and adaptive coin flipping. Phenotypic plasticity has been defined as the ability to produce different phenotypes in response to environmental cues (e.g., the number of leaves on a tree). The ability of an organism to produce a constant, stable phenotype regardless of environmental circumstances has been termed canalization. In the numerical computation and artificial life literature, the term...
resiliency has a similar meaning (Soule, 2006). Canalization, where the phenotype is entirely determined by genotype, and adaptive coin flipping are important parameters in our model. Canalization and resiliency are the extremes of a continuum. The term adaptive coin flipping emphasizes the random nature of the phenotypic trait and in our model is analogous to the pure error component of model fitting. Compared to a lack-of-fit term, resiliency describes variability in the phenotype that cannot be fit by any model. The increased distribution from resiliency allows parents to produce offspring with a wider variability as may be advantageous when environments vary in time or space.

Spatial structure in plant and insect biomass was reported from reaction–diffusion models examining the fate of recessive resistance genes of insect pests to transgenic plants carrying insecticidal genes from B. thuringiensis (Medvinsky et al., 2004). The biomass of plants with the insecticidal genes was determined by the duration of the insect reproductive period (Medvinsky et al., 2004, 2005). Additional modeling efforts have revealed that when resistance genes are dominant, the insect can cause extinction of the plant under some situations (Medvinsky et al., 2005). Adding complexity to the model revealed that temporal dynamics result in the coexistence of chaotic attractors and limit cycles for a single set of parameter values (Medvinsky et al., 2006). Transgenic plant biomass may ultimately depend on insect diffusion fluxes generated by the spatial heterogeneity of transgenic plants and the insects susceptible and resistant to their insecticides (Medvinsky et al., 2007).

In the model presented here, we extend this approach to examine the evolution of resistance within the insect pest populations. We ask two questions. First, does variation in the application of biocontrol agents result in the evolution of pathogen defense, recombination rate and canalization? Second, how do defense genotype, recombination and canalization interact to optimize adaptation to biocontrol pathogens in variable environments? The answers to these questions are discussed with respect to developing application procedures for biocontrol agents that minimize the likelihood that pest organisms will evolve pathogen defenses.

### 2. The data-driven model

Adaptation to bio-driven pathogen can be affected by the prevalence of infection (the number of individuals in a population that are infected), the intensity of infection (the number of parasites per infected host), the number of genes that contribute to the defense, as well as host population size.

We identify uniformly distributed pathogens as those that infect all hosts equally and consider the case where all individual have 50 parasites per host as well as a case where all individuals have 100 parasites. When more genes contribute to a trait, there is more potential for variability due solely to genetic factors, so consideration of the number of genes contributing to a phenotype is important. In small populations, random events are more important and small populations may be overwhelmed by random processes and thus, unable to evolve to local selective pressures. We consider each of these in turn.

### 2.1. Model background

Genetic algorithms were originally developed to solve optimization problems, using the concept of evolution as the basis for repetitive cycles of adaptation of a plan or design (Goldberg, 1989). The algorithm starts with a number of “parent” strings analogous to possible design solutions (i.e., organisms with various quantities of PO) that encode a set of parameters and selects the best-adapted strings to reproduce. Their offspring become parents and, in turn, produce further “generations” (Fig. 1). The biological processes of recombination and mutation generate diversity among the solutions. Over many generations the solutions, on which the algorithm operates, evolve toward an optimal design. We refer to our model as an evolutionary algorithm. It is a combination of a genetic algorithm and other techniques from evolutionary computation (Bäck et al., 1997) that results in a data-driven, multi-objective, optimization algorithm. The data-driven model uses data from laboratory experiments to predict when organisms are likely to evolve pathogen defenses. The model predicts both the result of direct selection (on the quantity of pathogen defense), indirect selection (on other aspects of an organism’s genetic system that influence within population variation for the trait). By considering both the trait under selection as well as other components of the genetic system, we allow both the mean and the variance of the trait to evolve.

### 2.2. The data

The model developed here illustrates the use of quantitative data on a costly biochemical defense, PO, from Japanese beetles, P. japonica. We develop a model of a quantitative defense that has physiological, energetic, genetic, and/or hormonal constraints. These constraints establish a functional relationship between defense phenotype, parasite prevalence and intensity, and fitness that can be empirically measured.

### 3. Evolutionary algorithm

The simulated organisms have quantitative, polygenic defenses, are diploid (two strings), and show defense quantities that range from 0 to 12 units. We examined a situation with 6 genes each with 2 alleles for a total of 12 alleles; and a second situation with 12 genes and 24 alleles.

#### 3.1. Initialize population

In Fig. 1, we show an example of the 2 strings of a diploid individual, each with 6 genes (for a total of 12 alleles designated 0 and 1); each type 1 allele contributes 1 unit to the individual’s genotype enabling the genetic component of the defense phenotype to range between 0 and 12 units. For 12 genes, allele values are 0 and 0.5; the contribution from each gene diminishes, but the number of possible genotypes increases. An individual’s defense quantity is calculated assuming each of the genes and alleles has an equal effect. For example, the defense genotype is the sum of the allele values and for the individual shown in Fig. 1 is 5; while
recombination rate and canalization are the average of the allele values, $r = 0.2$ and $e = 0.45$, respectively.

Individuals inherit from each parent and pass on to their progeny one allele for each defense gene, a recombination rate allele and a canalization allele. The initial frequency of each defense allele is 0.5. Recombination rates of complete linkage ($r = 0$, a gene complex on a single chromosome), independent assortment ($r = 0.5$) and an intermediate case ($r = 0.25$) were examined for population sizes of 50 and 500 individuals.

In our simulations, canalization varied, $0 / C20 e / C20 1$, so that 0–100% of the phenotype was randomly determined. The parameter, restricted to increments of 0.1, was initialized as a uniform random value between 0 and 1 that determined the proportion of the genotype that was due to non-genetic factors. For $r = 0$ and 0.5, all individuals had the same recombination rate. In the $r = 0.25$ simulations, the recombination rate underwent self-adaptation. For $r = 0.25$, individual recombination rates followed a random uniform distribution in increments of 0.05 with an initial population mean recombination rate of 0.25. This parameterization allows us to investigate how the scale of spatial variation in prevalence and intensity affect the evolution of the mean and variance in the defense phenotype. We track the evolution of the genotype that influences the mean value of the defense trait as well as recombination and canalization, which affect the phenotypic variance in defense.

### 3.2. State of infection and calculation of genotype

In test cases with “uniform” biocontrol application, all individuals in a population had 0, 50, 60 or 100 parasites. With “variable” biocontrol, 50 or 60% of the individuals (prevalence of 0.5 and 0.6) had 100 parasites per host (intensity of 100 parasites per host). For the variable infection treat-
ments, an individual was infected with 100 parasites if a uniformly distributed random number between 0 and 1 is less than the population prevalence. These data show an inflection point at 0.57; populations with prevalence < 0.57 should evolve 0 defenses, and those with prevalence > 0.57 should evolve to 12 units of defense.

3.3. Determine individual fitness using empirical data

The defense phenotype is a combination of the defense genotype and canalization. Canalization determines the extent to which identical genotypes become variable though development. An individual’s phenotype determines its fitness; and this fitness score determines the likelihood that the individual will be selected for the next generation. The fitness function is based on empirical data of Japanese beetle mortality as a function of severity of infection with the entomopathogenic fungus M. anisopliae (Tucker and Stevens, 2003).

3.4. Selection of parents and production of gametes

Individuals were selected stochastically to mate and reproduce. The probability of being chosen was proportional to their rank fitness (highest fitness = highest rank). Each time an individual is selected as a parent, it produces a gamete. When \( r = 0 \), one string of defense alleles (Fig. 1) is randomly chosen for the gamete. When \( r = 0.5 \), one allele at each defense locus is randomly selected. When \( 0 < r < 0.5 \), a uniform random number is generated for each defense gene. For the first gene, if the random number is lower than the recombination rate, the allele from the first string is inserted into the gamete; if the random number is higher, the allele from the second string is chosen. If the second random number is greater than the recombination rate, the same string contributes the second defense allele, otherwise, the allele comes from the alternative string. The process continues for each defense gene. The alleles for recombination, \( r \), and canalization, \( e \), are independently chosen with probability 0.5.

3.5. Mating

Gametes randomly combine to generate the individuals of the next generation. The cycle continues as the new population undergoes assignment of infection and calculation of fitness as described above.

4. Results

Conventional standards for genetic algorithms were followed: 10 replicates were run until no genetic variation occurred (Herrera and Lozano, 2001). The final defense phenotype, genotype, canalization and recombination rate (\( r = 0.25 \) treatments) were recorded. The dependent variables were analyzed by ANOVA using JMP statistical software (SAS, 1995). A full factorial model was used with main effects: host population size (50 or 500), state of infection (prevalence: \( U = \) uniform treatment; \( V = \) variable treatment), intensity of infection, and recombination on the evolution of (a) defense phenotype, (b) defense genotype and (c) canalization. Data are for six genes and host population size of 500. Data for 12 genes are similar. Data (not shown) for population size of 50 indicate more variation and less precision in reaching the optima.

4.1. Phenotypic differences among treatments

Treatments that vary in simulated biocontrol application vary in evolved defense phenotypes, demonstrating that the algorithm successfully models evolution of pathogen defense (Fig. 2a). For each dependent variable, the significant effects and interaction terms are presented (Table 1). In small \( N = 50 \) populations, genetic drift causes populations to vary around the optimum: the average is the same; however, there is more variation among the mean defense phenotypes of small populations than for larger populations. Because the difference in the evolved phenotypes between the population size treatments was in the variation among treatments rather than the mean value of treatments, we discuss and show results of only the large population size treatments and focus on effects of the other variables. Of these other variables, the defense phenotype most closely achieves the optimum in large populations with uniform prevalence, extreme (0 or 100) intensity, and high rates of recombination. The number of genes does not affect the evolved phenotypes.

4.2. Genotypic evolution

Simulation results indicate recombination rate, prevalence, intensity and number of defense genes had a significant effect on the evolved genotype (Fig. 2b). When compared to the results for phenotypic evolution, the \( F \) and \( P \) values for the
main effects are generally smaller and more interactions terms are significant (Table 1).

4.3. Evolution of canalization

In our model, canalization, the degree to which random factors, such as those that cause differences among identical twins or genetically homogeneous strains, evolves by indirect selection (Table 1) on the evolution of canalization. Overall, canalization tends to be higher in uniform environments. In variable environments, especially with the 0.5 recombination rate, phenotypes are determined less by genes and more by random factors so that a parent of any genotype can produce progeny with high fitness (Fig. 2c).

4.4. Recombination rate

Analysis of the \( r = 0.25 \) treatments shows that not all variables evolve by indirect selection. Even though recombination rate has a significant effect on the evolution of the phenotype, genotype, and canalization, it does not evolve under indirect selection (Table 1). Of the 11 effects in the recombination rate ANOVA (column 5, Table 1), only one three-way interaction is marginally significant (\( P = 0.034 \)).

5. Discussion

The data-driven model presented here gives insight into how environmental variability in biocontrol application could affect the evolution of pathogen defenses in pest populations. Many organisms live in heterogeneous environments with the result that selection can act on the mean and variance of a trait. For many species interactions, the fitness impact of infection increases with the intensity of infection (Meagher and Dudek, 2002; Rosa et al., 2002). Our model incorporates such empirical data to predict when and how populations will locally adapt to biological control pathogens and predicts that local adaptation will be affected by the variability in the prevalence and intensity of biocontrol agents. As described in the methods section, our data measures the fitness effect in populations with high or low resistance to pathogens and examined the effect of five doses of fungal spores on beetle survival. Thus, when developing biocontrol strategies, data on pathogen prevalence can be insufficient to predict evolution of resistance among populations, and information on the severity of infection may be required.

Evolution is a non-linear, multi-objective optimization process. An important contribution of our model is the analytical consideration of evolution in a diploid organism as a multi-objective, optimization process, including selection on canalization and recombination rate to affect variation. Examination of complete linkage (\( r = 0 \)) and free recombination (\( r = 0.5 \)) showed the recombination rate significantly affects evolution (Table 1). In spite of its effect on adaptation, recombination rate did not evolve (right column, Table 1) showing that a trait with a statistically significant effect on evolution may not evolve. Perhaps because diploid hosts experience recombination by segregation and independent assortment, as well as by crossing over, selection for recombination rate in this instance is reduced.

Canalization evolves when biocontrol application is uniform and high (Fig. 2c, 0.6 and 1, U). In contrast, if biocontrol agents are applied such that populations experience variability, the evolutionary response will be to expand the phenotypic distribution from a given genotype (Fig. 2c, 0.5 and 0.6, V). It is interesting that even in uniform environments, canalization approaches, but does not become 1. Canalization is linked to plasticity, the ability of organisms to modulate phenotypes based on environmental cues (Scheiner and...
Environments that favor low canalization would be likely to favor the evolution of plasticity and the current model could be expanded to predict when plasticity might evolve. Our model also simulates additive defense alleles and genes; however epistatic (i.e., non-linear) interactions between genes are likely to exist. Models mimicking the evolution of microorganisms using non-additive effects have shown that drastic phenotypic differences evolve in identical environments, that is multiple schemes of survival and reproduction had similar fitness (White and Adami, 2004). With respect to our results, it may be that epistasis could expand the combinations of defense genotypes, canalization and recombination that evolved in the different treatments.

Finally, we note that when there are no parasites, selection will be most effective when canalization is small. The fact that our model showed the evolution of both low genotypic values and low canalization in the absence of parasites suggests there can be evolution to increase the efficacy of evolution. A similar ability to fine-tune the outcome of selection by indirect selection on components of the genetic system has been reported by Soule (2006). Populations that initially evolved to a low, broad fitness peak evolved to become more resilient, and the population shifted to a higher, narrower peak.

Having populations that differed in both their history of biocontrol application and the production of biochemical defenses provided an opportunity to estimate and parameterize the three-dimensional, non-linear, functional relationship between pathogen dose, defense quantity and fitness. These biological data motivated our choice of a data-driven model. This represents a different approach compared to mathematical models with physics-based formulas. Evolutionary algorithms are excellent tools for examining the complex nature of these interacting factors.

In summary, our study presents a data-driven, multi-objective, optimization model to make predictions about the evolution of defenses against biocontrol agents. Although the factors included in the model were shown by others to be important in evolution, modeling of a quantitative defense, as a diploid, quantitative trait using empirical data is new. We show that pest organisms potentially have a multitude of ways to respond to biocontrol agents and genetic systems could evolve, not only to maximize fitness through the mean and variance of a trait, but also to increase the effectiveness of selection. These results can assist in making research-based management decisions.

Acknowledgement

Work was supported by VT EPSCoR NSF EPS Grant #0701410.

References


Computational remarks

The code implemented for MATLAB Version 6.1 Release 12.0 is copyrighted and available from the authors by sending a request to L. Stevens at http://lxsteven@uvm.edu. Permission is granted for the use in individual study. Publication or commercial use is explicitly not allowed without permission of the authors. The authors and the University of Vermont provide this code on an “as is” basis and accept no responsibility for errors, mistakes, or misrepresentations that may occur as a result of its use. A complete disclaimer is provided at the beginning of the program. Our use of MATLAB is for product identification purposes and does not represent an endorsement of the product.