Journal of the Korean Physical Society, Vol. 67, No. 12, December 2015, pp. 2154∼2162

Enhancement of Robustness by Mutation in the Host-parasite Coevolution

Yung-Gyung Kang
School of Computer Engineering, Hanshin University, Osan 18101, Korea

Jeong-Man Park*
Department of Physics, The Catholic University of Korea, Bucheon 14662, Korea

(Received 2 November 2015, in final form 30 November 2015)

We investigate the host-parasite coevolution with mutation in a single-locus system. We use the single-locus quasi-species model with a frequency-dependent fitness landscape and find the phase diagrams for given sets of evolution parameters by means of a linear stability analysis and numerical solutions of the coevolutionary dynamics equations. For a large mutation rate, there is only one internal fixed point. However, for a small mutation rate, within the linear order approximation, we find four steady equilibrium points and one oscillatory equilibrium point. All fixed points show the polymorphism in the host and the parasite populations in the presence of mutation whereas only one oscillatory equilibrium point shows the polymorphism in the absence of mutation. By starting from random initial host-parasite populations, we determine the separatrix and the basins of attraction in evolution parameter space. In comparison with results without mutation, we show that even a very small mutation rate enhances the robustness of the oscillatory equilibrium point; therefore, mutation is significant to maintain diversity in the evolution of finite populations. We also present the corresponding stochastic model for the host-parasite coevolution, and by computer simulation, we find that the stochastic simulation results are consistent with those of the quasi-species model.

PACS numbers: 87.23.Kg, 64.10.+h, 03.50.-z
Keywords: Quasi-species, Host-parasite coevolution, Frequency dependent fitness function
DOI: 10.3938/jkps.67.2154

I. INTRODUCTION

Understanding the dynamics of host-parasite coevolution and its consequences on genetic diversity is a central goal in a wide range of areas such as agriculture, epidemiology, and ecology. Host-parasite interactions are recognized as a major evolutionary force producing biological diversity. Genetic variation for host resistance reduces the probability that an individual parasite will infect an individual host [1]. Conversely, genetic diversity at parasite loci increases the range of potentially-susceptible hosts. A spatial and temporal genetic polymorphism is commonly found in nature at loci involved in host-parasite recognition, such as genes involved in gene-for-gene relationships in plant-parasite interactions [2], genes involved in matching allele relationships in invertebrates [3], or the major histocompatibility complex in vertebrates [4,5].

Due to the inherent complexity of host-parasite coevolutionary processes and the long time scales involved, theoretical models play an important role in the understanding of this process. Two major theoretical models have been presented for the evolutionary dynamics underlying infections in plants and animals. The gene-for-gene (GFG) model is based on data from plant-parasite interactions, especially crop plants [6]. The key feature of the GFG model is that one parasite genotype has universal virulence, which means that it can infect all host genotypes. Under the GFG theory, a cost of virulence is required to keep this virulent genotype from going to fixation. In contrast, the matching-alleles (MA) model is based on self/non-self recognition systems in invertebrates [3]. The key feature of the MA model is that infection (or resistance) requires a specific match between the host and the parasite genotypes at the corresponding locus; therefore, universal virulence is not possible in the MA model. In both the GFG and the MA models, negative indirect frequency-dependent selection (FDS) is assumed to be responsible for the polymorphism when the host and the parasite interact at the multiple loci. In these multi-locus models with negative indirect FDS, host and parasite genotypes have a selective advantage when they are rare in coevolving populations, which leads to cyclic oscillations of diverse genotypes in the host and the parasite populations. In regard to the diversity of genotypes in the host-parasite coevolution, important questions are whether multi-locus interactions

*E-mail: jmanpark@catholic.ac.kr

are necessary to maintain a polymorphism, and whether multi-locus interactions are sufficient to maintain a polymorphism by themselves or are ecological and biological factors required.

In order to address one of the above questions, the necessariness of the multi-locus interaction, we focus on the GFG model with a single-locus interaction and investigate whether a stable polymorphism can be maintained or not in a single-locus GFG model. In the GFG model, resistance is induced if the plant has a resistant (R) gene enabling recognition of a specific parasite avirulent (v) protein [7]. The parasite is not detected by the host, and the resistance is not induced if the host has a susceptible (r) allele or the parasite has a virulent (V) allele. The asymmetry of the GFG interaction implies that in the absence of other factors, there will be an "arms race," as successive pairs of R and V alleles are driven to fixation in the host and the parasite populations, respectively [8, 9]. Accounting for the observed diversity in the host and the parasite populations is a significant challenge. Conditions for the maintenance of a polymorphism in GFG interactions have been studied in a single-locus system with pairs of host R and r genes and parasite V and v genes [10–12]. In the process of coevolution, the selection rate of each allele depends on the frequencies of the other alleles directly and/or indirectly. This implies the existence of direct and/or indirect FDS. In order for a polymorphism to be maintained, it has been reported that there should be a negative direct FDS such as costs of R and V alleles so that the selection rates for the R and the V alleles decrease with increasing frequencies of those alleles [12]. However, while the costs of R and V alleles are necessary to maintain a polymorphism, they are not sufficient to do so in a single-locus GFG interaction system. With regards to the sufficiency and the stability of a polymorphism, we incorporate a mutation in the single-locus GFG model. Here, we investigate whether the existence of mutation stabilizes a polymorphism and is sufficient to maintain a polymorphism.

In a preliminary paper, the authors proposed the quasi-species model for the host-parasite coevolution without mutation [13]. In this paper, we propose a quasi-species model based on the GFG interaction for the host-parasite coevolution with mutation to investigate whether a single-locus system with the prey-predator-type negative direct FDS can maintain a polymorphism or multiple GFG loci are necessary for a polymorphism. The effects of mutation on the stability of the polymorphisms was introduced in Leonard's model by Leonard [14] and was investigated only via computer simulations by Segarra [11] and by Tellier and Brown [12]. We investigate the effects of mutation on the stability of the polymorphism analytically by means of a linear stability analysis and find that even for very small mutation rate, the oscillatory equilibrium state, in which both alleles (R and r alleles for hosts and V and v alleles for parasites) co-exist, persists over a larger area in phase space and for a longer time so that the robustness of the oscillatory equilibrium point can be said to be enhanced. Therefore, a mutation is significant for maintaining diversity in the evolution of finite populations. This paper is organized as follows: In Section II, we propose the quasi-species model with the prey-predator-type negative direct FDS. In the following section, a linear stability analysis and numerical solutions are used to determine the phase diagram. Then, we suggest the corresponding stochastic model for the host-parasite coevolution with mutation and present the simulation results in Section IV. The last section is for discussion.

II. FREQUENCY-DEPENDENT COEVOLUTIONARY DYNAMICS

We consider a parasite population causing disease in a host plant population. Both populations are assumed to be haploid (having a single set of unpaired chromosomes) and to have a point mutation. We also assume that the host and the parasite populations are infinitely large and mingle randomly. We propose the quasi-species model for the host-parasite coevolutionary dynamics with a frequency-dependent fitness landscape.

The model assumes a haploid single-locus inheritance in host resistance and parasite virulence similar to that in Jayakar's model [15]. Either a susceptible allele (r) or a resistant allele (R) can be at the host locus and either an avirulent allele (v) or a virulent allele (V) can be at the parasite locus. The frequencies of susceptible and resistant alleles in the host population are x_r and x_R ($x_r + x_R = 1$), respectively. The frequencies of avirulent and virulent alleles in the parasite population are y_v and y_V ($y_v + y_V = 1$), respectively. Similar to typical GFG interactions, the resistance is more effective when the resistant host is attacked by an avirulent parasite than by a virulent parasite. Taking this fact into account, we describe the coevolution dynamics by using the evolution equations for the frequencies of the host and the parasite populations:

$$\begin{aligned} \frac{dx_r}{dt} &= (f_r - \bar{f})x_r + \mu_h(x_R - x_r), \\ \frac{dx_R}{dt} &= (f_R - \bar{f})x_R + \mu_h(x_r - x_R), \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{dy_v}{dt} &= (g_v - \bar{g})y_v + \mu_p(y_V - y_v), \\ \frac{dy_V}{dt} &= (g_V - \bar{g})y_V + \mu_p(y_v - y_V), \end{aligned} \tag{2}$$

where μ_h and μ_p are the mutation rates for the host and the parasite populations, respectively. For simplicity, the same μ_h and μ_p are assumed for forward and backward mutations in both populations ($\mu_h^{r\to R} = \mu_h^{R\to r} = \mu_h$ and $\mu_p^{v\to V} = \mu_p^{V\to v} = \mu_p$). The frequency-dependent fitness of the susceptible host (f_r) and the fitness of the resistant

Table 1. Equilibrium points (EPs), eigenvalues and stability conditions.

Equilibrium points	Eigenvalues	Stability condition
$(0,0)$	$\lambda_1 = -\frac{1}{2}a(1-c_p)(1+c_h)+c_h$	$c_h < \frac{a(1-c_p)}{2-a(1-c_p)}$
	$\lambda_2 = -\frac{1}{4}a(3-c_p)(1-c_h)-dc_p$	stable
$(0,1)$	$\lambda_1 = -\frac{1}{2}(a-2c_h)$	$c_h < \frac{a}{2}$
	$\lambda_2 = \frac{1}{4}a(3-c_p)(1-c_h)+dc_p$	unstable
$(1,0)$	$\lambda_1 = \frac{1}{2}a(1-c_p)(1+c_h)-c_h$	$c_h > \frac{a(1-c_p)}{2-a(1-c_p)}$
	$\lambda_2 = -\frac{1}{4}a(3-4c_p)-dc_p$	$c_p < \frac{3a}{4(a-d)}$
$(1,1)$	$\lambda_1 = \frac{1}{2}(a-2c_h)$	$c_h > \frac{a}{2}$
	$\lambda_2 = \frac{1}{4}a(3-4c_p)+dc_p$	$c_p > \frac{3a}{4(a-d)}$

host (f_R) are given as

$$\begin{aligned} f_r &= A - y_v a_{vr} - (1-c_p) y_V a_{Vr}, \\ f_R &= (1-c_h)(A - y_v a_{vR} - (1-c_p) y_V a_{VR}), \end{aligned} \tag{3}$$

where A is a fitness value of the host in the absence of the parasite and we will set $A = 1$ without loss of generality because it only rescales the time. c_h and c_p ($0 \le c_h, c_p \le 1$) are the relative cost of having the resistant allele at the host locus and the cost of having the virulent allele at the parasite locus, respectively. a_{vr} is the fitness loss of the host with an r-allele due to infection by a parasite with a v-allele, and so forth.

The mean fitness of the host population ($\bar{f}$) is defined as

$$\bar{f} = f_r x_r + f_R x_R. \tag{4}$$

The frequency-dependent fitness of the avirulent parasite (g_v) and the fitness of the virulent parasite (g_V) are given as

$$\begin{aligned} g_v &= a_{vr}x_r + (1-c_h)a_{vR}x_R - x_r b_{rv} - (1-c_h)x_R b_{Rv} - d, \\ g_V &= (1-c_p)(a_{Vr}x_r + (1-c_h)a_{VR}x_R - x_r b_{rV} - (1-c_h)x_R b_{RV} - d), \end{aligned} \tag{5}$$

where d is the natural death rate of the parasite in the absence of the host, a_{vr} is the fitness gain of the parasite with a v-allele infecting the host with an r-allele, and b_{rv} is the fitness loss of the parasite with a v-allele due to the resistance of the host with an r-allele, and so forth. The mean fitness of the parasite population ($\bar{g}$) is defined as

$$\bar{g} = g_v y_v + g_V y_V. \tag{6}$$

For analytic simplicity, we set

$$\begin{pmatrix} a_{vr} & a_{vR} \\ a_{Vr} & a_{VR} \end{pmatrix} = \begin{pmatrix} \frac{a}{2} & 0 \\ a & \frac{a}{2} \end{pmatrix}, \quad a = 0.8, \tag{7}$$

and

$$\begin{pmatrix} b_{rv} & b_{rV} \\ b_{Rv} & b_{RV} \end{pmatrix} = \begin{pmatrix} \frac{b}{2} & 0 \\ b & \frac{b}{2} \end{pmatrix}, \quad b = \frac{a}{2}. \tag{8}$$

For given sets of dynamics parameters ($A = 1$, $a = 0.8$, $b = a/2$, and two d values of 0 and 0.1), we solve the coevolution equations in the phase space of $\{c_p, c_h\}$ within the linear stability analysis for large and small mutation region in the next section. The results for other sets of parameters will be investigated in future studies.

III. LINEAR STABILITY ANALYSIS AND NUMERICAL SOLUTIONS

The system of the coevolution equations of the quasi-species model with a frequency-dependent fitness landscape is given as non-linear differential equations:

$$\begin{aligned} \frac{dx_r}{dt} &= \left(f_r - (f_r x_r + f_R(1-x_r))\right) x_r + \mu_h(1-2x_r), \\ \frac{dy_v}{dt} &= \left(g_v - (g_v y_v + g_V(1-y_v))\right) y_v + \mu_p(1-2y_v), \end{aligned} \tag{9}$$

where

$$f_r = 1 - \frac{1}{2}a y_v - a(1-c_p)(1-y_v),$$

$$
\begin{aligned}
f_R &= (1-c_h)\left(1-\frac{1}{2}a(1-c_p)(1-y_v)\right),\\
g_v &= \frac{1}{2}ax_r-\frac{1}{2}bx_r-b(1-c_h)(1-x_r)-d,\\
g_V &= (1-c_p)\left(ax_r+\frac{1}{2}a(1-c_h)(1-x_r)-\frac{1}{2}b(1-c_h)(1-x_r)-d\right).
\end{aligned}
\tag{10}
$$

In the absence of mutation, Eq. (9) has four trivial equilibrium points: $(x_r^{ep},y_v^{ep})=(0,0),(0,1),(1,0)$, and $(1,1)$ and a single internal equilibrium point: $(x_r^{ep},y_v^{ep})=(x^*,y^*)$, with

$$
\begin{aligned}
x^* &= \frac{a(3-c_p)(1-c_h)+4dc_p}{a[3(c_p-c_h)+c_pc_h]},\\
y^* &= -\frac{a(1-c_p)(1+c_h)-2c_h}{a(c_p-c_h+c_pc_h)}.
\end{aligned}
\tag{11}
$$

The stabilities of the four trivial equilibrium points are determined by the signs of the eigenvalues (λ_1 and λ_2) of the system of the coevolution equations in the linear stability analysis:

$$
\frac{d}{dt}\delta x=\lambda_1\delta x,\quad \frac{d}{dt}\delta y=\lambda_2\delta y,
\tag{12}
$$

where $\delta x=x_r-x_r^{ep}$ ($\delta y=y_v-y_v^{ep}$) if $x_r^{ep}=0$ or x^* ($y_v^{ep}=0$ or y^*) and $\delta x=x_r^{ep}-x_r$ ($\delta y=y_v^{ep}-y_v$) if $x_r^{ep}=1$ ($y_v^{ep}=1$). If either $\lambda_1>0$ or $\lambda_2>0$, then the equilibrium points are locally unstable. The conditions that determine the stabilities of the four trivial equilibrium points are shown in Table 1.

The linearized equations near the internal equilibrium point (x^*,y^*) are given as

$$
\begin{aligned}
\frac{d}{dt}\delta x &= -\frac{[a(3-c_p)(1-c_h)+4dc_p]\,[a(3-4c_p)+4dc_p]\,(c_p-c_h+c_pc_h)}{2a\,[3(c_p-c_h)+c_pc_h]^2}\delta y,\\
\frac{d}{dt}\delta y &= \frac{(a-2c_h)\,[a(1-c_p)(1+c_h)-2c_h]\,[3(c_p-c_h)+c_pc_h]}{4a(c_p-c_h+c_pc_h)^2}\delta x,
\end{aligned}
\tag{13}
$$

and the eigenvalues are

$$
\lambda_{1,2}=\pm\sqrt{\frac{[a(3-c_p)(1-c_h)+4dc_p]\,[a(3-4c_p)+4dc_p]\,(a-2c_h)\,[a(1-c_p)(1+c_h)-2c_h]}{8a^2\,[3(c_p-c_h)+c_pc_h]\,(c_p-c_h+c_pc_h)}}.
\tag{14}
$$

The internal equilibrium point is unstable if the quantity inside the square root is positive while it becomes a center-fixed point, which is enclosed by limit cycles, when the quantity inside the square root is negative. By analyzing the system of the linearized coevolution equations and using the linear stability conditions, we determine the phase diagrams in $\{c_p,c_h\}$ space for a fixed value of d. The phase space is divided into four regions by three boundary equations:

$$
c_h = \frac{a(1-c_p)}{2-a(1-c_p)},
\tag{15}
$$

$$
c_p = \frac{3a}{4(a-d)},
\tag{16}
$$

$$
c_h = \frac{a}{2}.
\tag{17}
$$

These analytically-obtained phase diagrams are shown in Fig. 1, along with the simulation results. The three boundary equations are denoted by solid lines whereas the boundaries from the stochastic simulations are denoted by dotted lines. Equation (15) separates region A from regions B and D, Eq. (16) separates region B from regions C and D, and Eq. (17) separates regions C and D. Three regions (A, B, and C) show monomorphic compositions in the host and the parasite populations with the R-allele host and the V-allele parasite ($(x_r^{ep},y_v^{ep})=(0,0)$, region A), with the r-allele host and the V-allele parasite ($(x_r^{ep},y_v^{ep})=(1,0)$, region B), and with the r-allele host and the v-allele parasite ($(x_r^{ep},y_v^{ep})=(1,1)$, region C). There is no region with the R-allele host and the v-allele parasite ($(x_r^{ep},y_v^{ep})=(0,1)$). In region D where the internal equilibrium point becomes a center-fixed point, both the host

Table 2. Unshifted and Shifted equilibrium points.

Unshifted EPs	Shifted EPs : (x_r^{ep}, y_v^{ep})
$A : (0,0)$	$A^\mu : (A_x\mu_h, A_y\mu_p)$
$X : (0,1)$	$X^\mu : (X_x\mu_h, 1 - X_y\mu_p)$
$B : (1,0)$	$B^\mu : (1 - B_x\mu_h, B_y\mu_p)$
$C : (1,1)$	$C^\mu : (1 - C_x\mu_h, 1 - C_y\mu_p)$
$D : (x^*, y^*)$	$D^\mu : (x^* + D_x\mu_p, y^* + D_y\mu_h)$

and the parasite populations show a polymorphism. In the host population, the frequencies of the r-allele host and the R-allele host oscillate periodically. Also, in the parasite population, the frequencies of the v-allele parasite and the V-allele parasite oscillate periodically with the same periodicity as that of the host population.

In the presence of a mutation, the analysis of the coevolutionary dynamics equations becomes very complicated. Thus, we consider two extreme region: a large mutation rate and a small mutation rate. For a large mutation rate, there is only one fixed point: $(x_r^{ep}, y_v^{ep}) = (1/2, 1/2)$ because we assumed the same mutation rate for forward and backward mutations in both populations. Otherwise, it would have been $(x_r^{ep}, y_v^{ep}) = \left(\mu_h^{R\to r}/(\mu_h^{R\to r} + \mu_h^{r\to R}), \mu_p^{V\to v}/(\mu_p^{V\to v} + \mu_p^{v\to V}\right)$. Mutations regularly generate both alleles in the host and the parasite populations and make both populations polymorphic.

For a small mutation rate, we solve the coevolutionary dynamics equations up to the first order in each mutation rate and find five equilibrium points which are shifted by the mutation rates from those for without mutation. These five shifted equilibrium points are shown with the unshifted ones in Table 2.

For the equilibrium point (EP) A^μ, we find

$$\begin{aligned} A_x &= \frac{2}{a(1-c_p)(1+c_h) - 2c_h}, \\ A_y &= \frac{4}{a(3-c_p)(1-c_h) + 4dc_p}. \end{aligned} \tag{18}$$

While A_y is always positive, A_x is positive only in region A in Fig. 1. The steady state corresponding to the EP A^μ shows the polymorphic population structures in both populations. In the host population, the frequencies of a susceptible allele and a resistant allele are $x_r^{ep} = A_x\mu_h$ and $x_R^{ep} = 1 - A_x\mu_h$, respectively. In the parasite population, the frequencies of an avirulent allele and a virulent allele are $y_v^{ep} = A_y\mu_p$ and $y_V^{ep} = 1 - A_y\mu_p$, respectively. (In the absence of a mutation, the state corresponding to the EP A is monomorphic in both populations.) The polymorphism in the EP A^μ is solely due to the mutation process. Mutations regularly generate rare alleles (in this case the r-allele for the host and the v-allele for the parasite) in the host and the parasite populations and make both populations polymorphic. The rare-allele frequency in the host (parasite) population is proportional to the mutation rate of the host (parasite) population. This state cannot invade regions B and D because A_x becomes negative so that the frequency of a susceptible allele in the host population becomes negative in regions B, C, and D. For the EP X^μ, we find

$$\begin{aligned} X_x &= \frac{2}{a - 2c_h}, \\ X_y &= -\frac{4}{a(3-c_p)(1-c_h) + 4dc_p}. \end{aligned} \tag{19}$$

Because X_y is always negative, the frequency of an avirulent allele in the parasite population becomes greater than 1; therefore, the state corresponding to the EP X^μ is impossible to realize in the linear stability analysis.

For the EP B^μ, we find

$$\begin{aligned} B_x &= -\frac{2}{a(1-c_p)(1+c_h) - 2c_h}, \\ B_y &= \frac{4}{a(3-4c_p) + 4dc_p}. \end{aligned} \tag{20}$$

Both B_x and B_y are positive only in region B in Fig. 1. The state corresponding to the EP B^μ also shows polymorphic population structures in both populations. (In the absence of a mutation, the state corresponding to the EP B is monomorphic in both populations.) Mutations regularly generate an R-allele in the host population and a v-allele in the parasite population, and the rare-allele frequency in the host (parasite) population is proportional to the mutation rate of the host (parasite) population. This state cannot invade regions A, C, and D because B_x becomes negative so that the frequency of a susceptible allele becomes greater than 1 in region A and B_y becomes negative so that the frequency of an avirulent allele becomes negative in regions C and D.

For the EP C^μ, we find

$$\begin{aligned} C_x &= -\frac{2}{a - 2c_h}, \\ C_y &= -\frac{4}{a(3-4c_p) + 4dc_p}. \end{aligned} \tag{21}$$

Both C_x and C_y are positive only in region C in Fig. 1. The state corresponding to the EP C^μ also shows polymorphic population structures in both populations. (In the absence of a mutation, the state corresponding to the EP C is monomorphic in both populations.) Mutations regularly generate an R-allele in the host population and a V-allele in the parasite population, and the rare-allele frequency in the host (parasite) population is proportional to the mutation rate of the host (parasite) population. This state cannot invade regions B and D because C_y becomes negative so that the frequency of an avirulent allele becomes greater than 1 in region B and C_x becomes negative so that the frequency of a susceptible allele becomes greater than 1 in region D.

For the EP D^μ, we find

$$D_x = \frac{4(c_p - c_h + c_p c_h)\left[a - 2c_h + a(1-c_p)(1+c_h) - 2c_h\right]}{\left[3(c_p - c_h) + c_p c_h\right](a - 2c_h)\left[a(1-c_p)(1+c_h) - 2c_h\right]},$$
$$D_y = \frac{2\left[3(c_p - c_h) + c_p c_h\right]\left[a(3-4c_p) + 4dc_p + a(3-c_p)(1-c_h) + 4dc_p\right]}{(c_p - c_h + c_p c_h)\left[a(3-4c_p) + 4dc_p\right]\left[a(3-c_p)(1-c_h) + 4dc_p\right]}. \quad (22)$$

There is no positive-definite constraints on D_x and D_y, and the state corresponding to the EP D^μ also shows polymorphic, periodically-oscillating population structures in both populations just as the state corresponding to the EP D is polymorphic and periodically oscillating in both populations in the absence of a mutation. Unlike the other four EP cases, there are two interesting features for the EP D^μ. First, the frequency shift in the host population is proportional to the mutation rate of the parasite, and the frequency shift in the parasite population is proportional to the mutation rate of the host. Thus, the mutation rate of the parasite controls the evolution of the host population and vice versa. Second, the periodically-oscillating state of the EP D^μ can invade regions A, B, and C. On the boundaries determined by Eqs. (15), (16), and (17), the denominators of D_x and/or D_y vanish, and the frequency shifts look diverging. However, that is an artifact due to the expansion of the solutions in terms of the small mutation rate up to only the first order. If we keep all terms in the expansion, we will have finite frequency shifts on the boundaries. By analytic continuation of the frequency shifts from the inside of the EP D^μ region to the outside of the EP D^μ region, the periodically-oscillating state of the EP D^μ can invade regions A, B, and C. The detailed theoretical analysis around the boundaries to figure out how far the periodically-oscillating state of the EP D^μ can invade regions A, B, and C is very complicated. Thus, instead of the theoretical analysis, in the next section, we propose a stochastic model to investigate the phase diagram around the boundaries.

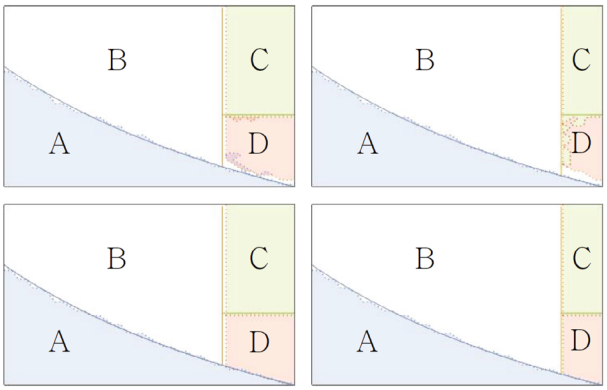

Fig. 1. (Color online) Phase diagrams obtained from the linear stability analysis (solid lines) and from stochastic simulations (dotted lines) in cost space $\{c_p, c_h\}$. In all figures, the horizontal axis denotes the virulent allele cost c_p for parasites running from 0 to 1, and the vertical axis denotes the resistent allele cost c_h for hosts running from 0 to 1. Figures in the upper (lower) row are phase diagrams without mutation (with mutation rate $\mu = 10^{-5}$). Figures in the left (right) column are phase diagrams with the death rate $d = 0$ ($d = 0.1$). Other parameters are $a = 0.8$ and $b = 0.4$.

IV. STOCHASTIC MODEL AND SIMULATION RESULTS

Even for the single-locus interaction host-parasite coevolution, analyzing the dynamical characteristics by using direct integration of the coevolution dynamics equations in the presence of mutation, Eq. (9), is very complicated. In order to analyze the dynamical characteristics, we use a stochastic model corresponding to the host-parasite coevolution equations, which was proposed in our preliminary paper [13]. There, we investigated the host-parasite coevolution without mutation by using both direct integration of the coevolution dynamics equations and simulation of the corresponding stochastic model, and we confirmed that the proposed stochastic model was consistent with direct integration. Thus, we modified the previous stochastic model to include the mutation process. The stochastic model can also be used for a multi-locus interaction system with point mutation, recombination, and horizontal gene transfer, as well as for a single-locus interaction system with mutation.

We consider single-locus host and parasite populations of size 1000 ($N_h = 1000, N_p = 1000$). First, we prepare random initial host and parasite populations. For the reproduction process, we use the same expressions for the frequency-dependent fitness values of the host and the parasite individuals as in Eq. (10). We calculate the fitness values for the r-allele (f_r) and the R-allele (f_R). If $f_r > f_R$, we produce one r-allele host and remove one R-allele host randomly, and if $f_r < f_R$, we produce one R-allele host and remove one r-allele host randomly. Similarly, we calculate the fitness values for the v-allele (g_v) and the V-allele (g_V). If $g_v > g_V$, we produce one v-allele parasite and remove one V-allele parasite randomly, and if $g_v < g_V$, we produce one V-allele parasite and remove one v-allele parasite randomly. For the mutation process, we choose one individual randomly from the host population and mutate its allele from an r-allele to an R-allele or vice versa with the mutation rate μ_h.

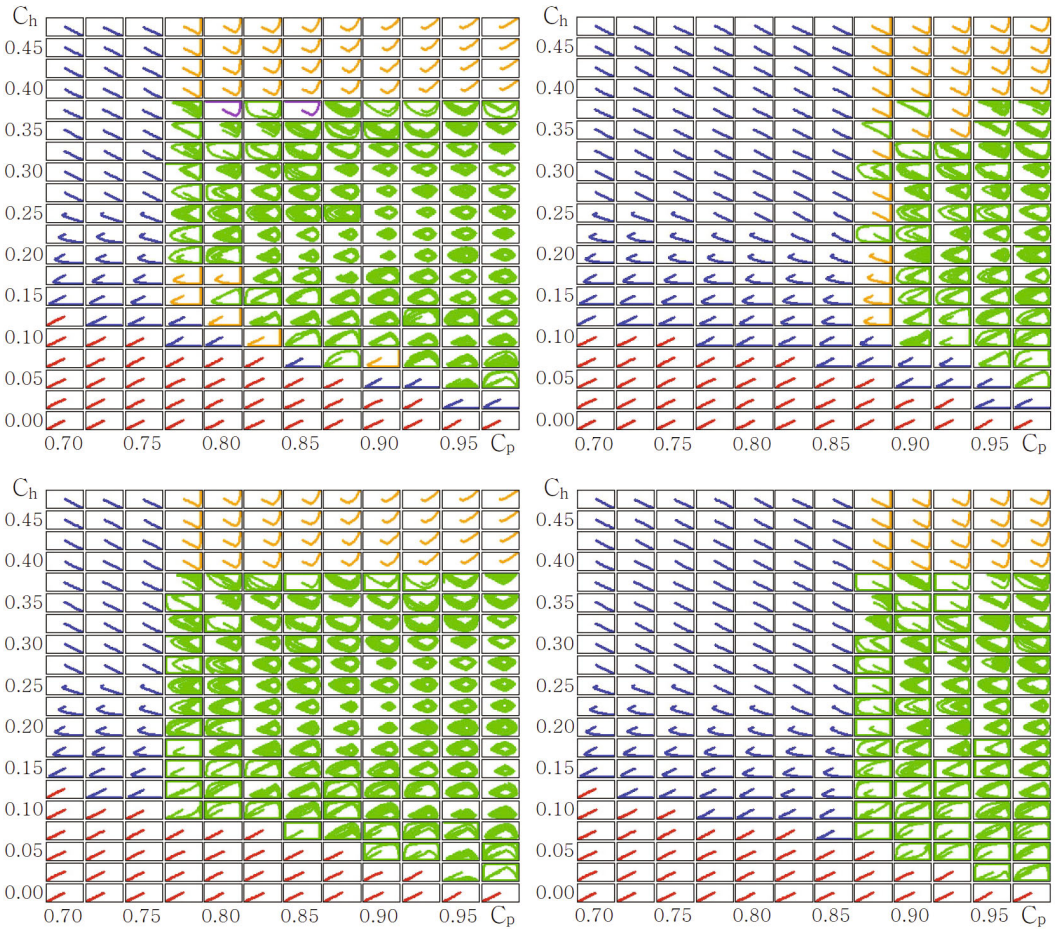

Fig. 2. (Color online) Dynamics results from stochastic simulations around boundaries of region D. Figures in the upper (lower) row are phase diagrams without mutation (with mutation rate $\mu = 10^{-5}$). Figures in the left (right) column are phase diagrams with the death rate $d = 0$ ($d = 0.1$). Other parameters are $a = 0.8$ and $b = 0.4$. In all small figures, the horizontal axis denotes the frequency of the r-allele in the host population, x, running from 0 to 1, and the vertical axis denotes the frequency of the v-allele in the parasite population, y, running from 0 to 1. Each small figure shows the dynamic trajectory of the point (x, y) for given cost values (c_p, c_h) starting from a random initial condition.

Similarly, we choose one individual randomly from the parasite population and mutate its allele from a v-allele to a V-allele or vice versa with the mutation rate μ_p. One generation corresponds to 1000 steps of the combined production and mutation processes, and we simulate the model for 1000 generations. By performing simulations for various cost values of the resistance and the virulence without a mutation rate and with a mutation rate $\mu_h = \mu_p = 10^{-5}$ for two values of $d = 0, 0.1$, we obtained the phase diagrams in $\{c_p, c_h\}$ space.

The phase diagrams from the stochastic simulations are shown in Figure 1, along with those from the linear stability analysis. Boundaries from the stochastic simulations are denoted by dotted lines whereas analytically-obtained boundaries are denoted by solid lines. The phase diagrams obtained from the stochastic simulations look similar to those obtained from the linear stability analysis of the host-parasite coevolution equations. However, the boundaries around the region D from stochastic simulations without mutation (the upper-row figures in Fig. 1) are much rougher than those from stochastic simulations with mutation, $\mu = 10^{-5}$ (the lower-row figures in Fig. 1). In the upper-left figure, which corresponds to simulations without mutation with a death rate $d = 0$, phases B and C invaded region D between regions A and D. In addition, the phase X, which is unstable under the linear stability analysis, invaded region D between regions C and D. Similarly, in the upper-right figure, which corresponds to simulations without mutation with a death rate $d = 0.1$, phases B and C invaded region D. On the contrary, in both figures in the lower row, which correspond to simulations with a mutation rate $\mu = 10^{-5}$ and with a death rate $d = 0$ (left) and $d = 0.1$ (right), region D is much more robust against invasion by phases B and C, as well as invasion by the unstable phase X. In order to emphasize the behaviors around the boundaries of the region D, we show the simulation dynamics results for representative points in the four regions in

Figure 2. The phase diagrams and the dynamics results are qualitatively consistent with the theoretical analysis, but quantitatively they are slightly different due to the genetic drift effect in finite-size populations. In any finite population, there will be continued stochastic variation in the frequencies of alleles. Stochastic variation will eliminate alleles with very small frequencies by chance at some point and eventually stop oscillations in the state of the EP D^{μ} to drive the population from polymorphism towards monomorphism. Thus, in the absence of a mutation, an analytically-stable oscillating state of the EP D^{μ} is driven to either the $(0,1)$ state of the EP B or the $(1,1)$ state of the EP C near boundaries where at least one allele becomes rare. On the other hand, in the presence of mutation, mutations regularly generate rare alleles to keep the state of the EP D^{μ} oscillating and both population compositions polymorphic.

V. DISCUSSION

We have investigated the host-parasite coevolution by using the quasi-species model based on the gene-for-gene interaction with a frequency-dependent fitness landscape. In the model, we assumed that the host and the parasite populations had a haploid single-locus inheritance in the host resistance and the parasite virulence and that both populations were monocyclic with and/or without mutation.

In the absence of mutation, there are five equilibrium points (EPs), among which four are trivial corner EPs and one is an internal EP. Three of the four trivial corner EPs and the internal EP are stable locally for some sets of dynamics parameters, and one of the four trivial corner EPs is always unstable. Three of the four trivial corner EPs result in a monomorphism in the host and the parasite populations, and the internal EP gives a polymorphism in both populations.

In the presence of mutation, we have also obtained five EPs, among which four are shifted corner EPs and one is a shifted internal EP. Again, three of the four shifted corner EPs and the shifted internal EP are stable locally for some sets of dynamics parameters, and one of the four shifted corner EPs is always unstable. All four stable, shifted EPs show a polymorphism in both populations. This means that a multi-locus is not necessary to maintain a polymorphism in host-parasite coevolution. Contrary to the previous result for the GFG interaction system in which a polymorphism is not possible in the single-locus model, the polymorphism can be maintained in the quasi-species model with a frequency-dependent fitness landscape even for a single locus.

In the phase diagrams, states corresponding to the three stable, shifted corner EPs cannot invade other regions. However, the periodically-oscillating state corresponding to the internal shifted EP can invade other regions so that the region of the EP D^{μ} may expand under the linear stability analysis as we take into account mutation. Although the phase diagrams look similar qualitatively, the dynamics around the boundaries of region D are quite different from those in the absence of mutation. To investigate the dynamics around the boundaries, we proposed the stochastic model corresponding to the quasi-species model for host-parasite coevolution with a frequency-dependent fitness landscape. By using stochastic simulations, we determined the phase diagrams in cost $\{c_p, c_h\}$ space. The results from the stochastic simulations with mutation are consistent with those from the linear stability analysis of the quasi-species model. However, we found some discrepancies in the stochastic simulation without mutation along the boundaries of the region D. In region D near the boundaries, the frequency of one allele becomes very small in the oscillating state in both the host and the parasite populations, and stochastic variations in a finite population eliminate this rare allele to drive the population towards a monomorphic state. In the presence of mutation, mutations regularly generate rare alleles to keep the oscillation going and populations polymorphic. This implies that even a very small mutation rate enhances the robustness of the oscillating polymorphic states.

Although we have analyzed single-locus monocyclic host-parasite coevolution, extending it to multi-locus interactions, to polycyclic interactions, and to both is straightforward. However, the coevolution equations for the coevolution system extended to the multi-locus and/or to the polycyclic interaction become too formidably complicated to be analyzed analytically. On the other hand, because the evolutionary stochastic process is so simple, the stochastic model for the single-locus interaction can be easily extended to the multi-locus interaction, provided the proper frequency-dependent fitness values are determined for all types of the host and the parasite individuals.

ACKNOWLEDGMENTS

YGK was supported by a Hanshin University research grant in 2015. JMP was supported by the Catholic University of Korea research fund 2015 and by the Basic Science Research Program through the National Research Foundation of Korea (Grant No. NRF-2013R1A1A2006983).

REFERENCES

[1] R. M. May and R. M. Anderson, Parasitology **100**, S59 (1990).
[2] P. H. Thrall, J. J. Burdon, and A. Young, J. Ecol. **89**, 736 (2001).
[3] R. K. Grosberg and M. W. Hart, Science **289**, 2111 (2000).

[4] V. Apanius, D. Penn, P. R. Slev, L. R. Ruff, and W. K. Potts, Crit. Rev. Immunol. **17**, 179 (1997).
[5] A. V. S. Hill, Annu. Rev. Genomics Hum. Genet. **2**, 373 (2001).
[6] H. H. Flor, Adv. Genet. **8**, 29 (1956).
[7] J. L. Dangl and J. D. G. Jones, Nature **411**, 826 (2001).
[8] J. Bergelson, G. Dwyer, and J. J. Emerson, Annu. Rev. Genet. **35**, 469 (2001).
[9] E. B. Holub, Nat. Rev. Genet. **2**, 516 (2001).
[10] A. Sasaki, Proc. R. Soc. Lond. Ser. B Biol. Sci. **267**, 2183 (2000).
[11] J. Segarra, Phytopathology **95**, 728 (2005).
[12] A. Tellier and J. K. M. Brown, Proc. R. Soc. Lond. Ser. B Biol. Sci. **274**, 809 (2007).
[13] Y-G. Kang and J-M. Park, J. Korean Phys. Soc. **64**, 322 (2014).
[14] K. J. Leonard, Ann. N. Y. Acad. Sci. **287**, 207 (1977).
[15] S. D. Jayakar, Theor. Popul. Biol. **1**, 140 (1970).