

# The Evolution of Polymorphic Compatibility Molecules

Rob J. De Boer

Theoretical Biology, Utrecht University

Several primitive colonial organisms distinguish self from nonself by means of polymorphic compatibility molecules bearing similarity to the major histocompatibility complex (MHC). The evolution of such polymorphisms is generally explained in terms of resistance to parasites. Ignoring parasites, I develop two mathematical models. One model is inspired by slime mold aggregation. The other is a caricature. The results show that parasites are not required to explain polymorphisms. The fact that allorecognition allows organisms to maintain their "genetic identity" (e.g., to preserve evolved adaptations) suffices for the evolution of polymorphisms. The degree of polymorphism that ultimately evolves is equal to the diversity of phenotypes in the population. In these models I consider "cells" growing in particular environments. The "genotype" of a cell codes for an "ecological preference," which determines the growth rate of the cell in the environment, and for a "compatibility molecule." Cells expressing the same compatibility molecule form one "colony." Thus, colonies may be chimeric with respect to the ecological preferences. The reproduction rate of each cell in a colony is weighted by the colony's composition of ecological preferences. These conditions turn out to be sufficient for the evolution of polymorphic compatibility molecules.

## Introduction

Transplantation reactions in vertebrate immune systems are based on highly polymorphic major histocompatibility complex (MHC) molecules. The variability of some of the loci comprising the mouse or human MHC is on the order of 100 alleles (Klein 1986, 1990). It seems likely that all vertebrate species have MHC genes and are able to respond strongly to MHC-disparate (i.e., allogenic) tissue transplants (Klein 1986, 1990). The biological function of MHC molecules in vertebrate immune systems is to present antigens in the form of short peptides to T-lymphocytes (Davis and Bjorkman 1988).

It is not obvious why MHC molecules need to be polymorphic. The rejection of allogenic grafts cannot be an explanation because graft rejection is an experimental artifact and not a natural phenomenon involved in evolutionary selection pressures. A simple benefit of extensive allelic MHC polymorphism is that each individual is expected to inherit different alleles from both parents. This doubles the diversity of MHC molecules on antigen-presenting cells and would enable the host

to present peptides from a larger variety of pathogens (Hughes and Nei 1988). This explanation is insufficient, however, because doubling the MHC diversity can also be achieved by doubling the number of MHC loci. (Also note that increasing the number of MHC molecules can be detrimental for the immune system; Nowak et al. 1992; De Boer and Perelson 1993). Thus, one has to invoke coevolutionary arguments involving mutating pathogens that are most likely to adapt to the most common MHC alleles in the population (Hartl et al. 1975; Bremermann 1980; Hamilton et al. 1990; Lawlor et al. 1990; De Boer 1991).

The situation is markedly different in primitive colonial organisms in which allorecognition is important for colony fusion. Allorecognition may thus be essential for survival and subjected to selection. Indeed, several colonial invertebrates are capable of allorecognition (Burnet 1971). Examples include slime molds (i.e., *Dictyostelium*; Buss 1982), several species of fungi (Todd and Rayner 1980), *Hydra* (Buss et al. 1985), and tunicates (i.e., *Botryllus*; Scofield et al. 1982). All protists and animals known to fuse possess a compatibility system (Buss 1982, 1987; Grosberg 1988). For some species it has been shown that this compatibility system controlling allorecognition is composed of highly polymorphic loci (Scofield et al. 1982; Grosberg and Quinn 1986; Weisman et al. 1990). Fungi and plants also have compatibility systems controlling heterokaryon forma-

Key words: polymorphisms, compatibility molecules, parasites, individuality, colony fusion, self recognition.

Address for correspondence and reprints: Rob J. De Boer, Theoretical Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, Holland. E-mail: rdb@alive.biol.ruu.nl

*Mol. Biol. Evol.* 12(3):494–502. 1995.  
© 1995 by The University of Chicago. All rights reserved.  
0737-4038/95/1203-0014\$02.00

tion in fungi and self-fertilization in plants. Thus, the polymorphic MHC of vertebrates is reminiscent of a preimmune compatibility system that seems to play a role in colony fusion.

The potential costs and benefits for chimeric colonies (i.e., colonies composed of genetically different cells) have been discussed extensively by Buss (1982, 1987) and Bonner (1988). The cost that is most widely considered is that of parasites exploiting the colony. These parasites can be either true, coevolving, foreign pathogens (Hamilton et al. 1990) or of somatic origin (Buss 1982, 1987). Somatic cell parasites are variants arising by mutation within a colony. Allorecognition is supposed to avoid exploitation by parasites because parasites are expected to adapt to (or originate from) the most common compatibility molecule in the population. Thus, the argument is very similar to that explaining polymorphisms in MHC for antigen presentation in vertebrates.

Coevolving parasites play an important role in evolution. Parasites are implicated in the evolution of sex because the genetic heterogeneity due to sexual reproduction allows for resistance to coevolving parasites (Hartl et al. 1975; Bremermann 1980; Hamilton et al. 1990; Howard and Lively 1994). In artificial worlds like Tierra, primitive sexual processes emerge spontaneously whenever parasites appear (Ray 1991). With respect to the problem of MHC polymorphisms, the main question is how diverse the compatibility system should be for sufficient resistance against parasitic variants. It has been argued that the protection against parasitic genes can only account for limited diversity of the compatibility system (Nauta and Hoekstra 1994).

I study here whether the avoidance of chimeras on the basis of polymorphic compatibility molecules could have a more general explanation than "defense against parasites." In my model I do not allow for any of the above mechanisms for the evolution for polymorphism; that is, there is no overdominance, no mating incompatibility, and no parasites. Instead, I allow for environmental variation which maintains variation in physiological state (e.g., ecological preference). I will show that each physiological phenotype in the model population evolves a unique compatibility allele. Thus, for populations with many different phenotypes, this mechanism accounts for highly polymorphic compatibility loci. The evolution of polymorphisms in this model depends on a balance between the growth of a population in preferred and nonpreferred conditions.

Thus, I formally demonstrate that for populations growing under various environmental conditions, the preservation of a "genetic identity" is a sufficient selection pressure for the evolution of polymorphic compat-

ibility molecules. This result confirms previous speculations that were based on experimental observations with fungi (Todd and Rayner 1980).

My results are also related to the ideas expressed by Varela and Countinho (1991) and Tauber (1994), who argue that a crucial function of the immune system is to define the "self." Their arguments are based on the enormously diverse lymphocyte immune system of variable region molecules. The principal difference is that my results consider compatibility molecules that are polymorphic at the level of the population and not at the level of the individual.

## Problem

Consider a situation in which two strains, say  $a$  and  $b$ , of some species live under two environmental conditions, say A and B. The strains may either migrate between two environments, or one environment may fluctuate temporarily between the two conditions. Let strain  $a$  grow faster than  $b$  in environment A. Similarly, let strain  $b$  grow faster than  $a$  in environment B. Thus, assume that  $a$  has a preference for A and  $b$  for B. The strains grow in colonies that may be either homothallic (i.e., of the form  $aa$  or  $bb$ ,) or chimeric (i.e., of the form  $ab$ ). Now consider the growth of strains in chimeric colonies. In environment A the growth of  $a$  in a chimeric colony  $ab$  is expected to be smaller than that of  $a$  in a homothallic colony  $aa$ . Thus, the growth of  $a$  in environment A is retarded by  $b$ . However, by the same arguments, the growth of  $a$  in a chimeric colony  $ab$  in environment B is expected to be enhanced by the presence of  $b$ . Thus, allorecognition to avoid chimeras would seem beneficial in preferred conditions but detrimental in nonpreferred conditions. Whether strains should in general avoid forming chimeric colonies seems inconclusive on the basis of these verbal arguments. My formal models aim to clarify this issue.

Previous authors have considered related problems. For instance, Buss (1982) and Bonner (1988) argue that cells having gained some advantage by mutation will try to avoid aggregation because other genotypes could "sap their reproductive success." This argument implies that a cell "knows" when it has gained some advantage over its competitors; otherwise, it would be better to allow for aggregation. We make no such distinction between good and bad genotypes: averaged over all environments, all genotypes perform equally well.

## Models

I will consider populations (i.e., strains) with different ecological preferences, and possibly different compatibility molecules, growing colonially in different ecological environments. In each environment the

strains compete with one another. The strain with the highest growth rate (i.e., the strain for which the preference perfectly matches the environment) is expected to outcompete all other strains in the environment. In each environment, colonies are formed by the aggregation of all strains expressing identical compatibility molecules.

The growth of each strain depends on its match between preference and environment and on the total reproduction of the colony it is a member of. The reproduction of the colony depends on its composition. A colony that is entirely composed of strains with a perfect match between their ecological preference and the environment reproduces at a maximum rate. Conversely, colonies that are “contaminated” with non-matching strains have a lower total reproduction. This equally affects all strains in the colony.

In these models I maintain a diversity of ecological preferences by having several environments and several preferences. Strains migrate between the environments, thus allowing for an inflow of maladapted strains in all environments.

### Definitions

Let the strain  $x_{pa}$  be defined as a nondimensional density of a genotype expressing the ecological preference  $p$  and the compatibility allele  $a$ . Further, let  $x_{pa}^e$  denote the density of the strain  $x_{pa}$  in environment  $e$ . Let there be  $n_e = n_p$  environments and preferences, and  $n_a$  compatibility alleles. I define the match  $f(e, p)$  between preference and environment to be linearly dependent on the difference between  $e$  and  $p$ , that is,

$$f(e, p) = \max(0, 1 - \alpha|e - p|), \quad (1)$$

where I make periodic boundary conditions. Compatibility molecules only match when they are transcribed from the same allele  $a$ .

### Logistic Growth

The strains grow logistically (i.e., are inhibited by the total population in the environment). Scaling the maximum per capita growth to one, I let the growth rate be determined by the match function  $f(e, p)$ . Thus, the strains grow according to

$$\frac{dx_{pa}^e}{dt} = x_{pa}^e(f(e, p) - T_e), \quad (2)$$

where  $T_e$  is the total population in environment  $e$ ; that is,

$$T_e = \sum_p \sum_a^{n_p, n_a} x_{pa}^e. \quad (3)$$

Here the equilibrium population size of a strain scales linearly with its growth rate. Thus, the strains with the highest growth rate,  $f(e, p) = 1$ , are expected to outcompete the others. Finally, note that the populations are scaled by setting the maximum population density (i.e. the carrying capacity) to one.

### Caricature Model

Developing a caricature model, I specify that the growth of strains in favorable conditions is hampered by the presence of other strains, whereas that of strains in poor conditions is enhanced by the presence of strains finding these conditions favorable.

For simplicity, consider two preferences, two environments, and two compatibility molecules; that is, let  $e, p, a \in \{0, 1\}$ . This defines eight populations  $x_{00}^0, x_{01}^0, \dots, x_{11}^1$ . For brevity, I will use a tilde ( $\tilde{\cdot}$ ) to denote the binary complement (i.e.,  $\tilde{1} \equiv 0$  and  $\tilde{0} \equiv 1$ ). Because there are only two environments, the matching function  $f(\cdot)$  can only take two values:  $f(0, 0) = f(1, 1) = 1$  and  $f(0, 1) = f(1, 0) = 1 - \alpha$ . The total density in environment  $e$  is given by equation (3):

$$T_e = x_{00}^e + x_{01}^e + x_{10}^e + x_{11}^e. \quad (4)$$

For strains growing in preferred conditions, I propose

$$\begin{aligned} \frac{dx_{pa}^e}{dt} = & \varepsilon + x_{pa}^e(1 - \beta x_{\tilde{p}a}^e - T_e) \\ & + D(x_{\tilde{p}a}^e - x_{pa}^e), \end{aligned} \quad (5a)$$

for  $e = p$ , and  $p, a \in \{0, 1\}$ ,

where the term  $-\beta x_{\tilde{p}a}^e$  specifies the reduction of the growth of strain  $x_{pa}^e$  due to the presence of the strain with the same compatibility allele  $a$  and the “wrong” preference  $\tilde{p}$ . The last term describes the migration of strain  $x_{pa}$  between the two environments as a standard diffusion process. The term  $\varepsilon$  represents mutation.

For strains growing in nonpreferred conditions, I propose

$$\begin{aligned} \frac{dx_{pa}^e}{dt} = & \varepsilon + x_{pa}^e(1 - \alpha + \beta x_{\tilde{p}a}^e - T_e) \\ & + D(x_{\tilde{p}a}^e - x_{pa}^e), \end{aligned} \quad (5b)$$

for  $e = \tilde{p}$ , and  $p, a \in \{0, 1\}$ ,

where the term  $+\beta x_{\tilde{p}a}^e$  specifies the enhancement of the growth of strain  $x_{pa}^e$  due to the presence of the strain

with the same compatibility allele  $a$  and the “right” preference  $\tilde{p}$ .

Thus, there are eight ODEs and four parameters  $\alpha$ ,  $\beta$ ,  $\varepsilon$ , and  $D$ . The mutation term is always small, (i.e.,  $\varepsilon \leq 0.001$ ) and plays a negligible role. I choose as typical values  $\alpha = 0.2$  (i.e., a reduction in the maximum growth rate of 20%) and  $D = 0.1$  (i.e., a migration of 10%). Both parameters are varied throughout the analysis. I study the effect of  $\beta$  (i.e., the mutual influence of strains with different ecological preferences).

## Results

The caricature model is based on two symmetries. First, summed over both environments, the two strains have the same total growth rate. Second, strains expressing different compatibility alleles but having the same preference have identical growth rates. Thus, the system will always have a symmetric equilibrium in which strains with different compatibility alleles have the same density. Since this should be the same in both environments, this equilibrium corresponds to equal densities for the matching strains  $x_{00}^0 = x_{01}^0 = x_{10}^0 = x_{11}^0$  and equal densities for the nonmatching strains  $x_{10}^1 = x_{11}^1 = x_{00}^1 = x_{01}^1$ . If this equilibrium is stable, evolution is attracted to a situation in which chimeric colonies are not avoided. Due to these equalities, the model can be reduced to two equations in symmetric equilibria. There is one ODE for a matching strain, say  $x$ ,

$$\frac{dx}{dt} = \varepsilon + x[1 - \beta y - 2(x + y)] + D(y - x), \quad (6a)$$

and one for a nonmatching strain, say  $y$ ,

$$\begin{aligned} \frac{dy}{dt} = \varepsilon + y[1 - \alpha + \beta x - 2(x + y)] \\ + D(x - y). \end{aligned} \quad (6b)$$

I numerically confirmed the equilibrium values of this reduced model. The reduced model has only one equilibrium, and it is a stable node.

The full eight-dimensional model is studied by numerical bifurcation analysis. I continue the symmetric state, changing  $\beta$  as a bifurcation parameter (see fig. 1). At a critical value of  $\beta = \beta_C$ , the symmetric state is involved in a pitchfork bifurcation. When  $\beta < \beta_C$ , the system has two stable asymmetric states; otherwise, the symmetric state is stable. In the asymmetric states, there is either  $x_{00}^0 = x_{11}^0 > x_{00}^1 = x_{11}^1$  with  $x_{01}^0 = x_{10}^0 \approx x_{01}^1 = x_{10}^1 \approx 0$  or the complementary counterpart of this:  $x_{01}^0 = x_{10}^0 > x_{01}^1 = x_{10}^1$  with  $x_{00}^0 = x_{11}^0 \approx x_{00}^1 = x_{11}^1 \approx 0$ . Due to these equalities, the model can also be re-

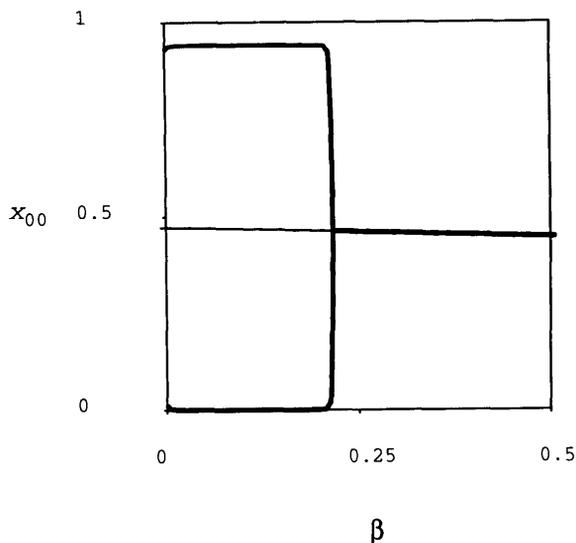


FIG. 1.—The pitchfork bifurcation of the symmetric steady state of the caricature model. I plot the equilibrium density of  $x_{00} = x_{00}^0 + x_{00}^1$  as a function of the degree of the mutual influence  $\beta$ . For the parameter setting  $\alpha = 0.2$ ,  $D = 0.1$ ,  $\varepsilon = 10^{-6}$ , the symmetric equilibrium, in which phenotypes are not identified by a unique compatibility molecule, becomes unstable around  $\beta_C = \beta \approx 0.22$ . For  $\beta < \beta_C$ , two asymmetric stable steady states exist in which the different phenotypes are identified by a compatibility molecule, that is,  $x_{00} \approx 0.9$  and  $x_{01} \approx 0$  or  $x_{00} \approx 0$  and  $x_{01} \approx 0.9$ .

duced in asymmetric equilibria. There is one ODE for a matching strain, say  $x$ ,

$$\frac{dx}{dt} = \varepsilon + x[1 - (x + y)] + D(y - x), \quad (7a)$$

and one for a nonmatching strain, say  $y$ ,

$$\frac{dy}{dt} = \varepsilon + y[1 - \alpha - (x + y)] + D(x - y). \quad (7b)$$

I numerically confirmed the equilibrium values of this reduced model. The reduced model has only one equilibrium, and it is a stable node. Because in asymmetric states half of the strains are virtually zero, the equilibrium densities of the other strains are roughly twice as large as those of the symmetric state (see fig. 1).

In figure 2 I demonstrate the robustness of the results by a parameter sensitivity analysis. For  $D = 0.01$ ,  $0.1$ ,  $1$  I compose a two-dimensional diagram having  $\alpha$  and  $\beta$  as bifurcation parameters. Thus, figure 2 depicts the location of the pitchfork bifurcation in parameter space. First, observe that the bifurcation is not restricted to a small parameter domain. Second, see that  $\beta_C$  almost linearly depends on  $\alpha$  and hardly depends on  $D$ .

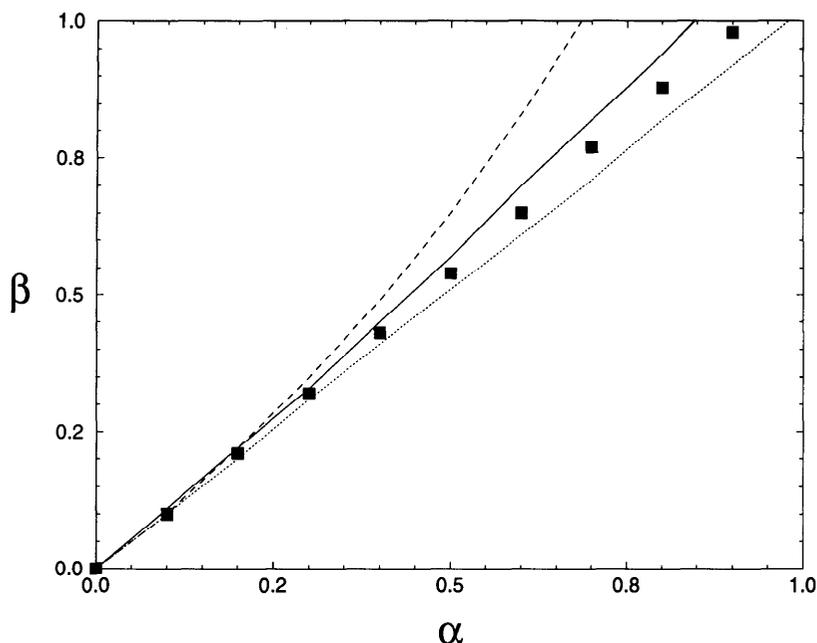


FIG. 2.—Two parameter continuations of the pitchfork bifurcation of the caricature model. For various values of  $D$ , I plot  $\beta_C$  as a function of  $\alpha$ . The dotted line is for  $D = 0.01$ , the solid line for  $D = 0.1$ , and the dashed line for  $D = 1$ . The black squares are the values predicted by equation (8c) where  $x + y$  is the numerical solution of equation (7) for  $D = 0.1$ .

An intuitive interpretation of the pitchfork bifurcation is that at the critical point  $\beta_C$ , the growth in non-preferred conditions exceeds that in preferred conditions. If most of the growth is in nonpreferred conditions, it is favorable to allow for chimeric colonies. The maximum growth rate in nonpreferred condition is

$$1 - \alpha + \beta x, \quad (8a)$$

where  $x$  is the density of a matching strain. The maximum growth rate in preferred condition is

$$1 - \beta y, \quad (8b)$$

where  $y$  is the density of a nonmatching strain. Thus, we write for  $\beta_C$

$$1 - \alpha + \beta_C x = 1 - \beta_C y, \quad \text{or} \quad \beta_C = \frac{\alpha}{x + y}, \quad (8c)$$

where  $x$  and  $y$  are the densities of a strain in preferred and nonpreferred conditions, respectively. Since each strain grows in both conditions,  $x + y$  represents the density of each genotype. In figure 1 we saw that below the critical point the equilibrium strain density is of order 1. Thus, this predicts that  $\beta_C \approx \alpha$ . Additionally, because below the critical point the equilibrium population values hardly depend on  $\beta$ , we may solve equation (7) to find approximate values for  $x + y$  in equation (8c). We

numerically confirm these intuitive results by the black squares in figure 2. These are numerical solutions of equations (7) and (8c) for various values of  $\alpha$ . Given the approximation for  $x$  and  $y$  around the critical point, these solutions correspond well with the two-dimensional bifurcation results.

Summarizing, I conclude that generally polymorphic compatibility molecules are favorable. When  $\beta < \beta_C$  (i.e., when the mutual influence is sufficiently small), there are two asymmetric stable steady states in which each ecological preference is characterized by a unique compatibility allele. The system thus converges to equation (7) in which the strains hardly influence each other. If  $\beta > \beta_C$ , the only equilibrium that an evolutionary process can attain is one in which there is no association between the ecological preference and a compatibility allele. The system converges to equation (6) in which strains do influence each other by forming chimeric colonies. The reason for this is that when  $\alpha$  is small or  $\beta$  is large, the strains expand most in the non-preferred conditions. Then it is beneficial to form chimeric colonies. Thus, when the mutual influence  $\beta$  is sufficiently small, we expect evolution to attain an attractor where populations are polymorphic such that colony fusion is prevented.

#### Aggregation Model

Slime molds like *Dictyostelium discoideum* are widely studied because of their primitive mechanism for

complex pattern formation. These slime molds have a cellular growth phase and a colonial reproduction phase. The classical data on the *D. mucoroides* (Buss 1982) reveal that these protists have polymorphic compatibility molecules and that they refrain from colonial fusion with incompatible strains. Thus, they are capable of self/non-self discrimination based on an MHC-like system.

My “aggregation” model is inspired by slime molds with a cellular growth phase and colonial reproduction. My aim is not to develop a model of slime mold aggregation but to illustrate that the results on polymorphism also hold for an example system that is somewhat closer to biological reality. Additionally, I demonstrate that these models can account for a high degree of polymorphism.

The critical feature of the aggregation model is that it explicitly accounts for the formation of colonies. First, the total reproductive output of a colony depends on its phenotypic composition. Thus, colonies composed of well-adapted strains have a larger reproductive output than those contaminated with poorly adapted strains. Second, the model assumes that the members of a colony equally share the total reproductive output. Thus, poorly adapted strains profit from the reproduction of the better-adapted strains in their colony.

For the cellular growth phase, I employ equation (2) to implement the logistic competition for some resource. Thus, I numerically solve equation (2) from  $t = 0$  until  $t = L$ , where  $L$  is defined as the length of the growth phase. In each environment, the strains with the maximal growth rate  $f(e, p) = 1$  (and, hence, maximal carrying capacity) are expected to outcompete the others. The degree to which the others are outcompeted depends on the length of the growth phase  $L$ . Thus, the fraction of nonmatching strains in each environment can be increased by decreasing  $L$ .

At the end of the growth phase, colonies are formed in each environment. Hence, all strains form colonies at  $t = L$ . I simply aggregate all strains expressing the same compatibility molecules in each environment. Thus, the strains form colonies composed of one compatibility molecule and possibly of many ecological preferences. I propose that the total number of germs produced per colony is proportional to the weighted sum of all its members. The total number of germs produced by a colony with compatibility allele  $a$  in environment  $e$  is

$$G_a^e = r \sum_p^{n_p} f(e, p) x_{pa}^e, \quad (9a)$$

where it is assumed that genotypes produce germs proportional to their match  $f(e, p)$  between preference and

environment. The allocation of these germs to the individual members of the colony is proportional to their densities and does not depend on their preferences. I thus assume that the total number of germs depends on the “support” each member gives to the colony. Whether or not, or how much, an individual member reproduces might be due to a chance process that is only influenced by the fraction of that member in the colony. Thus, for each member of the colony, the number of germs is

$$g_{pa}^e = G_a^e \frac{x_{pa}^e}{\sum_p x_{pa}^e}, \quad (9b)$$

where the latter term is simply the fraction of member  $x_{pa}^e$  in its colony.

In the caricature model of equation (5), a term  $\beta$  represents the mutual influence of the strains. In equations (9a) and (9b), the production of germs depends on the match function  $f(e, p)$ . Thus, the reproductive output of a particular strain is increased by the other strains with a better match and decreased by strains with a worse match.

The colonial reproduction (i.e., eq. [9b]), provides the initial condition for the next growth phase. I allow for migration between the environments by assuming that the germs  $g_{pa}^e$  produced per environment are distributed evenly over all environments. Thus, the initial condition for the cellular growth phase is

$$x_{pa}^e(0) = \frac{1}{n_e} \sum_e^{n_e} g_{pa}^e. \quad (10)$$

On top of these deterministic processes, the model allows for stochastic mutation. I assume that the production of germs  $g_{pa}^e$  is subjected to mutation. The probability of a mutational event is proportional to the number of germs produced (i.e., to  $g_{pa}$ ). When a mutation occurs, a few germs (i.e.,  $\varepsilon$ ) obtain another genotype by randomly selecting another value of either  $p$  or  $a$ . Thus, at the reproduction stage, for each genotype that reproduces, a random number  $0 < R < 1$  is drawn from a uniform distribution, that is,

$$\begin{aligned} &\text{if } R < g_{pa}^e P_m, \\ &\text{then } g_{\widehat{pa}}^e = g_{pa}^e + \varepsilon \quad \text{and} \quad g_{pa}^e = g_{pa}^e - \varepsilon, \end{aligned} \quad (11)$$

where  $g_{\widehat{pa}}^e$  is defined by equation (9b),  $P_m$  determines the probability of mutation, and  $g_{\widehat{pa}}^e$  denotes a mutant having a random value for either  $p$  or  $a$ .

## Results

I study the aggregation model (eqq. [2, 9–11]) for an initial condition with all strains expressing the same

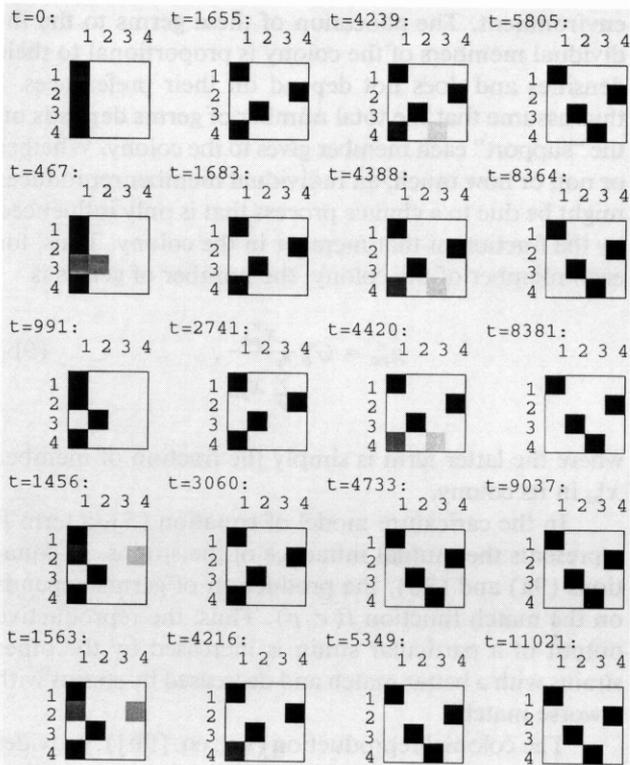


FIG. 3.—Genotype frequencies of the aggregation model. Each panel represents a snapshot taken whenever a mutation significantly changes the frequency distribution. In each panel the rows depict the preferences  $1 \leq n_p = n_e \leq 4$ . The columns depict the compatibility alleles  $1 \leq n_a \leq 4$ . Parameters:  $L = 10$ ,  $P_m = 0.01$ ,  $r = 0.1$ ,  $\varepsilon = 0.01$ , and  $\alpha = 0.2$ . At  $t = 467$ ,  $x_{32}$  has attained a new compatibility allele  $a = 2$ . This mutant  $x_{32}$  has successfully replaced  $x_{31}$  at  $t = 991$  and persists throughout the simulation. At  $t = 1,456$ ,  $x_{21}$  has attained a new compatibility allele  $a = 4$ . This mutant  $x_{24}$  persists throughout the simulation. At  $t = 2,741$ ,  $x_{11}$  attains the compatibility allele  $a = 2$ . Because  $a = 2$  is expressed already by  $x_{32}$ , this mutant fails to expand. At  $t = 4,216$ ,  $x_{41}$  attains a new compatibility allele  $a = 3$ . Now each preference is uniquely identified by a compatibility allele, (i.e., by  $a = 1, 4, 2$ , and  $a = 3$ , respectively). This configuration is stable to further perturbations (compare  $t = 11,021$  with  $t = 4,733$ ). The system was integrated with the variable time-step order four Runge Kutta integrator provided by Press et al. (1988).

compatibility allele,  $a = 1$ , and with random values for the preferences and environments. The simulation results depicted in figure 3 show that each ecological preference establishes a unique compatibility allele. The panels are snapshots of the genotype frequencies. Snapshots are taken whenever a mutation gives rise to a new frequency distribution. Each square in figure 3 indicates in a gray scale the density of one genotype  $x_{pa}$  summed over all environments. The columns in each panel denote the compatibility alleles (i.e.,  $1 \leq a \leq 4$ ), the rows denote the preferences (i.e.,  $1 \leq p \leq 4$ ). After 4,733 generations (see panel 14), the four preferences have selected the compatibility alleles 1, 4, 2, and 3, respec-

tively. From then on, the distribution remains fixed. Thus, each preference is identified by a unique marker. If we double the number of preferences, environments, and compatibility alleles, we obtain similar results. I have checked this up to  $n_p = n_e = n_a = 32$ ; that is, there is a population attaining 32 different compatibility alleles (see fig. 4).

For simplicity, the per capita growth in the model is independent of the size of the colony. It is possible to allow for faster growth in larger colonies by multiplying equation (9b) with a standard saturation function of the colony size (i.e., of  $\sum_p^n x_{pa}$ ). This give similar results; that is, in the more complicated model, each preference also evolves a unique compatibility allele (not shown). There are two technical differences, however. First, because in the more complicated model large colonies do better than small ones, mutants appearing in small numbers typically fail to expand. Thus, in the more complicated model, I had to modify equation (11). The second difference appears when the simple model allows for more compatibility alleles than ecological preferences (i.e., when  $n_a > n_p$ ). In the simple model, in which the effect of colony size is neutral, every phenotype selects a unique set of compatibility alleles. In the more complicated model, in which large colonies do better than small ones, every phenotype selects a single unique compatibility allele. I think the latter is more realistic.

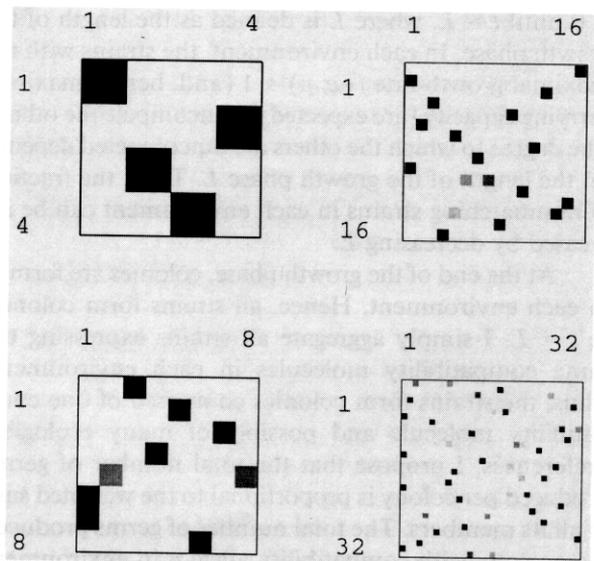


FIG. 4.—Genotype frequencies of the aggregation model. See the legend of fig. 3 for further explanation. In panels a–d I increase the number of preferences and environments  $n_p = n_e$  and the number of compatibility alleles  $n_a$  to show that the model allows for large heterogeneity. Panels a–c represent stationary distributions; panel d is a transient. a, For  $n_p = n_e = n_a = 4$ , the distribution at  $t = 11,079$ . b, For  $n_p = n_e = n_a = 8$ , the distribution at  $t = 3,904$ . c, For  $n_p = n_e = n_a = 16$ , the distribution at  $t = 19,957$ . d, For  $n_p = n_e = n_a = 32$ , the distribution at  $t = 12,220$ .

## Discussion

Supposing that there are compatibility molecules which enable multicellular colonies to prevent fusion with other colonies, the results here demonstrate that for any genotype competing with other genotypes under various environmental conditions, it may be advantageous to express a marker uniquely identifying itself. My interpretation of these results is that unique compatibility molecules allow genotypes to prevent contamination with maladapted genotypes (i.e., allow genotypes to preserve their "genetic identity"). I show that this is favorable whenever the total reproduction in good conditions, in which fusion reduces fitness, is larger than the total reproduction in worse conditions, in which fusion tends to enhance fitness. It is obvious that the mechanism proposed here is very general and is not restricted to the slime molds and colonial tunicates to which the models are most related.

Although I have excluded explicit parasites from the models, one could argue that the different strains are parasites of one another. A nonmatching strain growing in a colony that is largely composed of matching strains is some sort of parasite sapping the colony's reproductive success. However, the same strain (i.e., genotype) growing in preferred conditions is a host whose reproductive success is decreased by the same sort of parasites. Thus, my results can also be interpreted in terms of the protection against parasites (Hartl et al. 1975; Bremermann 1980; Buss 1982, 1987; Hamilton et al. 1990; Howard and Lively 1994). My novel contribution is that the model can account for a high degree of polymorphism because there are many different parasites when there are many different ecological genotypes. In other words, I show that polymorphism at the level of adaptations to environmental circumstances can account for a similar degree of polymorphism at the level of adhesion molecules.

The control of cell or colony fusion is related to the problem of sexual reproduction. Sex is the mixing and random segregation of two genetic identities. In ciliates (Grell 1973), basidiomycetes (Day 1978), and angiosperms (De Nettancourt 1977), sex is controlled by a system of incompatibility types. In addition to this compatibility system, these organisms may have binary (i.e., male/female) mating systems (Hurst and Hamilton 1992). The slime mold *Physarum polycephalum* will only mate when there is incompatibility at three polymorphic loci (Kawano et al. 1987). A more recent example are mice that apparently chose mates on the basis of MHC haplotypes (Potts et al. 1991). It is tentative to speculate that the control of sex and the control of fusion, which we have studied here, are based on the same mechanism. My results suggest that genetic fusion

(i.e., sex) is generally advantageous when environmental conditions are poor. In such circumstances, organisms could employ polymorphic compatibility molecules to avoid inbreeding.

The compatibility molecules that we have considered are related to cell adhesion molecules (CAMs). In embryonic development, CAMs play a crucial role because a small number of CAMs, with different binding specificities, allows cells to coordinate movement into differentiated tissues (Edelman 1987, 1989). The neural adhesion molecule (N-CAM) that has been sequenced is a member of the immunoglobulin superfamily (Cunningham et al. 1987). This is a family of transmembrane molecules including immunoglobulin, the T-cell antigen receptor, MHC, CD3, CD4, CD8, and several other molecules comprising the vertebrate immune system (Williams and Barclay 1988). The primordial function of members of this superfamily seems to be homophilic interactions controlling the behavior of cells in a multicellular organism (Williams and Barclay 1988). There is weak evidence that N-CAM is related to slime mold adhesion molecules (Matsunaga and Mori 1987), but well-defined CAMs have so far only been isolated in vertebrates (Edelman 1987).

Because CAMs are a member of the immunoglobulin superfamily and because vertebrate MHC molecules are generally polymorphic, it seems likely that the vertebrate immune system originally evolved from a somatic compatibility system. Thus, the results here would imply that the immune system evolved from a system of self/non-self discrimination, the function of which was the protection of any individual's genetic identity. The role the immune system plays in the defense against parasites could thus be acquired much later in evolution.

## LITERATURE CITED

- BONNER, J. T. 1988. The evolution of complexity by means of natural selection. Princeton University Press, Princeton, NJ.
- BREMERMANN, H. J. 1980. Sex and polymorphisms as strategies in host-pathogen interactions. *J. Theor. Biol.* **87**:671-702.
- BURNET, F. M. 1971. "Self-recognition" in colonial marine forms and flowering plants in relation to the evolution of immunity. *Nature* **232**:230-235.
- BUSS, L. W. 1982. Somatic cell parasitism and the evolution of somatic tissue compatibility. *Proc. Natl. Acad. Sci. USA* **79**:5337-5341.
- . 1987. The evolution of individuality. Princeton University Press, Princeton, NJ.
- BUSS, L. W., J. L. MOORE, and D. R. GREEN. 1985. Auto-reactivity and self-tolerance in an invertebrate. *Nature* **313**:400-402.
- CUNNINGHAM, B. A., J. J. HEMPERLY, B. A. MURRAY, E. A. PREDIGER, R. BRACKENBURY, and G. M. EDELMAN. 1987.

- Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. *Science* **236**:799–806.
- DAVIS, M. M., and P. J. BJORKMAN. 1988. T cell antigen receptor genes and T cell recognition. *Nature* **334**:395–402.
- DAY, P. R. 1978. Evolution of incompatibility. Pp. 67–80 in M. N. SCHWALB and P. G. MILES, eds. *Genetics and morphogenesis in the Basidiomycetes*. Academic Press, New York.
- DE BOER, R. J. 1991. Allelic variations of human TCR V gene products. *Immunol. Today* **5**:169–169.
- DE BOER, R. J., and A. S. PERELSON. 1993. How diverse should the immune system be? *Proc. R. Soc. Lond. B* **252**:171–175.
- DE NETTANCOURT, D. 1977. *Incompatibility in angiosperms*. Springer, Berlin.
- EDELMAN, G. M. 1987. CAMs and Igs: cell adhesion and the evolutionary origins of immunity. *Immunol. Revs.* **100**: 11–45.
- . 1989. Topobiology. *Sci. Am.* **260**:76–88.
- GRELL, K. G. 1973. *Protozoology*. Springer, Berlin.
- GROSBERG, R. K. 1988. The evolution of allorecognition specificity in clonal invertebrates. *Q. Rev. Biol.* **63**:377–412.
- GROSBERG, R. K., and J. F. QUINN. 1986. Kin recognition and colony fusion in the colonial ascidian *Botryllus schlosseri*. *Nature* **322**:456–459.
- HAMILTON, W. D., R. AXELROD, and R. TANESE. 1990. Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl. Acad. Sci. USA* **87**:3566–3573.
- HARTL, D. L., E. R. DEMPSTER, and S. W. BROWN. 1975. Adaptive significance of vegetative incompatibility in *Neurospora crassa*. *Genetics* **81**:553–569.
- HOWARD, R. S., and C. M. LIVELY. 1994. Parasitism, mutation accumulation and the maintenance of sex. *Nature* **367**:554–557.
- HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitutions at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**:167–170.
- HURST, L. D., and W. D. HAMILTON. 1992. Cytoplasmic fusion and the nature of the sexes. *Proc. R. Soc. Lond. B* **247**:189–194.
- KAWANO, S., T. KURIOWA, and R. W. ANDERSON. 1987. A third multiallelic mating type locus in *Physarum polycephalum*. *J. Gen. Microbiol.* **133**:2539–2546.
- KLEIN, J. 1986. *Natural history of the major histocompatibility complex*. Wiley, New York.
- . 1990. *Phylogeny of the major histocompatibility complex*. Munksgaard, Copenhagen.
- LAWLOR, D. A., J. ZEMMOUR, P. D. ENNIS, and P. PARHAM. 1990. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. *Ann. Rev. Immunol.* **8**:23–63.
- MATSUNAGA, T., and N. MORI. 1987. The origin of the immune system: the possibility that immunoglobulin superfamily molecules and cell adhesion molecules of chicken and slime mold are all related. *Scand. J. Immunol.* **25**:485–495.
- NAUTA, M., and R. F. HOEKSTRA. 1994. Evolution of vegetative incompatibility in filamentous ascomycetes I: deterministic models. *Evolution* (in press).
- NOWAK, M. A., K. TARCZY-HORNOCH, and J. M. AUSTYN. 1992. The optimal number of MHC molecules in an individual. *Proc. Natl. Acad. Sci. USA* **89**:10896–10899.
- POTTS, W. K., C. J. MANNING, and E. K. J. WAKELAND. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature* **352**:619–621.
- PRESS, W. H., B. P. FLANNERY, S. A. TEUKOLSKY, and W. T. VETTERLING. 1988. *Numerical recipes in C*. Cambridge University Press, New York.
- RAY, T. S. 1991. An approach to the synthesis of life. Pp. 371–408 in C. LANGTON, C. TAYLOR, J. D. FARMER, and S. RASMUSSEN, eds. *Artificial life II*. Santa Fe Institute Studies in the Sciences of Complexity. Vol. X. Addison-Wesley, Redwood City, Calif.
- SCOFIELD, V. L., J. M. SCHLUMPBERGER, L. A. WEST, and I. L. WEISSMAN. 1982. Protochordate allorecognition is controlled by a MHC-like gene system. *Nature* **295**:499–502.
- TAUBER, A. I. 1994. The immune self: theory or metaphor? *Immunol. Today* **15**:134–136.
- TODD, N. K., and A. D. M. RAYNER. 1980. Fungal individualism. *Sci. Prog., Oxf.* **66**:331–354.
- VARELA, F. J., and A. COUTINHO. 1991. Second generation immune networks. *Immunol. Today* **12**:159–166.
- WEISMANN, I. L., Y. SAITO, and B. RINKEVICH. 1990. Allorecognition histocompatibility in a protochordate species: is the relationship to MHC somatic or structural? *Immunol. Revs.* **113**:227–241.
- WILLIAMS, A. F., and A. N. BARCLAY. 1988. The immunoglobulin superfamily—domains for cell surface recognition. *Ann. Rev. Immunol.* **6**:381–405.

STANLEY A. SAWYER, reviewing editor

Received June 20, 1994

Accepted November 21, 1994