PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Recognition and Polymorphism in Host-Parasite Genetics

Steven A. Frank

Phil. Trans. R. Soc. Lond. B 1994 **346**, 283-293
doi: 10.1098/rstb.1994.0145

Recognition and polymorphism in host–parasite genetics

STEVEN A. FRANK

Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717, U.S.A.

SUMMARY

Genetic specificity occurs in many host–parasite systems. Each host can recognize and resist only a subset of parasites; each parasite can grow only on particular hosts. Biochemical recognition systems determine which matching host and parasite genotypes result in resistance or disease. Recognition systems are often associated with widespread genetic polymorphism in the host and parasite populations. I describe four systems with matching host–parasite polymorphisms: plant–pathogen interactions, nuclear–cytoplasmic conflict in plants, restriction enzymes in bacterial defence against viruses, and bacterial plasmids that compete by toxin production and toxin immunity. These systems highlight several inductive problems. For example, the observed patterns of resistance and susceptibility between samples of hosts and parasites are often used to study polymorphism. The detectable polymorphism by this method may be a poor guide to the actual polymorphism and to the underlying biochemistry of host–parasite recognition. The problem of using detectable polymorphism to infer the true nature of recognition and polymorphism is exacerbated by non-equilibrium fluctuations in allele frequencies that commonly occur in host–parasite systems. Another problem is that different matching systems may lead either to low frequencies of host resistance and common parasites, or to common resistance and rare parasites. Thus low levels of host resistance or rare parasites do not imply that parasitism is an unimportant evolutionary force on host diversity. Knowledge of biochemical recognition systems and dynamical analysis of models provide a framework for analysing the widespread polymorphisms in host–parasite genetics.

1. INTRODUCTION

Recognition systems often have spectacular genetic polymorphisms. Self-incompatibility or mating-type loci may have 100 or more alleles, each with a distinct label (Fincham *et al.* 1979; Richards 1986). The great diversity of major histocompatibility complex (MHC) alleles is probably maintained by the need to recognize a wide range of parasitic invaders (Potts & Wakeland 1993). The numerous bacterial restriction enzymes that revolutionized molecular biology evolved for specific recognition and defence against parasitic attack (Wilson & Murray 1991).

In this paper, I show that the study of recognition and polymorphism in host–parasite systems requires analysis of five related factors, as follows.

1. The biochemistry of recognition: the interacting molecules of the host and parasite that signal a parasitic invasion to the host's defensive arsenal.
2. The potential genetic diversity: the number of host and parasite genotypes that have different recognition properties.
3. The pattern of specificity: which parasite genotypes can attack which host genotypes.
4. The dynamics of disease: the processes that combine the recognition system and ecology to determine the actual amount of genetic polymorphism at any point in time and space.
5. Inductive problems: the pitfalls of using observed polymorphism and specificity to infer factors 1–4.

I focus on four exemplar systems of recognition and polymorphism: plant–pathogen genetics, nuclear–cytoplasmic conflict in plants, bacterial restriction enzymes that defend against viruses, and bacterial plasmids that compete by toxin production and toxin immunity. In each of these systems the hosts (those attacked) are widely polymorphic for their ability to recognize and resist assault. The parasites (aggressors) are, in turn, widely polymorphic for host-range: the range of host genotypes that they can attack by escaping recognition and resistance. These four examples are perhaps the best understood of systems with reciprocal interactions between genetic polymorphisms of host and parasite. 'Best understood' is, of course, a relative measure. For these examples one can describe a plausible system of recognition and reciprocal interaction to explain the observed polymorphisms, but many details are unknown.

I have two goals. First, I summarize the natural history of these four systems. The facts are interesting, and the examples provide a database in which to search for recurring themes in host–parasite genetics.

Phil. Trans. R. Soc. Lond. B (1994) **346**, 283–293
Printed in Great Britain

(*a*)

	A	*V*
R	−	+
S	+	+

(*b*)

	p_1	p_2
h_1	−	+
h_2	+	−

Figure 1. Resistance (−) and susceptibility (+) between two host alleles and two pathogen alleles. (*a*) Gene-for-gene specificity. (*b*) Matching-allele specificity.

Second, I focus on inductive problems and show that, in host–parasite systems, appearances are often deceptive. I provide examples for the following problems, each of which has led investigators to draw incorrect conclusions from their data.

1. Detectable polymorphism is a poor guide to the actual polymorphism.
2. Detectable polymorphism is a poor guide to the underlying biochemistry of recognition.
3. A low frequency of host resistance and a wide parasite host-range do not imply a lack of coevolutionary pressure.
4. A low level of parasitism does not imply a lack of parasite pressure on host polymorphism.

Dynamic models for each of the individual examples have been published elsewhere; I do not repeat the details here. Instead I emphasize general aspects of host–parasite specificity and evolutionary dynamics that contribute to misleading inferences about recognition and polymorphism.

2. PLANTS AND PATHOGENS

The reciprocal genetic interactions between plants and pathogens have been studied extensively because of the economic importance of crop disease. Recent work has extended the study of plant disease to natural populations. I will focus on two patterns of plant–pathogen genetics in natural populations (Burdon 1987; Frank 1992). (1) Genetic specificity is common. Each host genotype resists only specific pathogen genotypes; each pathogen genotype attacks only specific host genotypes. Many host and pathogen loci interact to determine the success of a pathogen attack. (2) Each host locus has a low frequency of resistance alleles, and each matching pathogen locus has a high frequency of alleles with a broad host-range.

I will show that two distinct models of specificity, gene-for-gene and matching-allele, are consistent with these observations on recognition and polymorphism. These alternative models illustrate the inductive problems of using detectable polymorphisms to infer the actual polymorphism and the biochemistry of specificity.

(*a*) *Gene-for-gene models*

During the 1940s and 1950s, H. H. Flor studied the inheritance of specific resistance and virulence factors in flax and its fungal pathogen, flax rust (Flor 1971). The interaction between host and pathogen genotypes turned out to have simple properties that Flor referred to as a 'gene-for-gene' system. In an idealized gene-for-gene system, each pair of resistance and susceptibility alleles in the host has a matching pair of host-range alleles in the pathogen.

Figure 1*a* shows the standard gene-for-gene model (Flor 1956, 1971; Burdon 1987). The host has two phenotypes, resistant (R) and susceptible (S). The pathogen has two phenotypes, avirulent (A) and virulent (V). Plant pathologists use the term 'virulence' for host-range. I follow their convention.

In the gene-for-gene model a host resists attack only when the host–parasite pair has an $R:A$ match. Person & Mayo (1974) refer to this match as a 'stop-signal'. Recent biochemical models suggest that the avirulence allele (A) produces a gene product (elicitor) that can be recognized only by specific host receptors (R). This specific elicitor–receptor recognition induces a non-specific set of host defence mechanisms (Gabriel & Rolfe 1990).

In multilocus interactions each host locus is matched to a single parasite locus. If there is an $R:A$ match between at least one pair of host–parasite loci, then disease does not develop. This is consistent with the model of the $R:A$ match as a stop-signal.

Although the relationship between specific factors is simple in a gene-for-gene system, the total interaction between a host and its pathogen is complex. Flor and others have identified 29 separate host resistance factors in flax, each with a complementary virulence factor in flax rust (Flor 1971; Lawrence *et al.* 1981). Similar gene-for-gene interactions are now known or suspected for over 25 different host–pathogen pairs (Burdon 1987). These systems do not conform exactly to the idealized gene-for-gene assumptions (Christ *et al.* 1987), but these systems do have complementary major-gene interactions between hosts and pathogens.

Several theoretical models have been developed to explain observed polymorphism in systems that are believed to have gene-for-gene specificity (reviewed by Leonard & Czochor 1980; Levin 1983; Burdon 1987; Frank 1992, 1993*a*). These models assume that virulence alleles have a negative effect on fitness that offsets the benefit of wider host-range. This assumption is necessary because, without a fitness cost, the virulence allele would spread to fixation (in figure 1*a*, V has an advantage over A). A model that predicted fixation of virulence could not explain the observed polymorphism of virulence and avirulence alleles (Vanderplank 1968). A similar argument leads to the conclusion that resistance must have a cost.

Models with costs predict that resistance alleles will be rare relative to susceptibility alleles, and virulence alleles will be common relative to avirulence alleles. Observations appear to support the predicted frequencies of these allelic classes (reviewed by Frank 1992, 1993*a*). These observations are not surprising given the assumptions of the gene-for-gene models. A pathogen must avoid recognition at all pairs of matching host–parasite loci in order to succeed. This

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

requires few recognizable avirulence alleles and many costly universal host-range (virulence) alleles. For a host the costs of resistance alleles must be balanced against the low probability of matching rare avirulence alleles. This balance usually leads to a low frequency of the resistance allele at each locus.

Small costs of resistance and virulence lead at equilibrium to widespread multilocus diversity, pathogens that can attack a wide range of hosts, and consequently little effective resistance. In spite of the low level of resistance, the widespread genetic diversity is maintained by reciprocal coevolutionary pressures of host and parasite.

The gene-for-gene model is compelling because it is consistent with several observations. First, the biochemistry of recognition, the elicitor–receptor model, fits the gene-for-gene model in figure 1*a* if universal virulence (V) is the absence of an elicitor and universal susceptibility (S) is the absence of a matching receptor.

Second, the gene-for-gene model, which specifies the types of alleles and their phenotypic effects, is consistent with the observed pattern of phenotypic interactions shown in figure 1*a*. This reasoning may appear circular because the gene-for-gene model was derived from the observed patterns of phenotypic interactions. However, as I show in the next section, a very different model for the genetics of host–parasite interactions would also lead to a similar pattern of phenotypic interactions.

Third, the population genetic observations of rare resistance alleles and common virulence alleles can be explained by gene-for-gene interactions plus small costs of resistance and virulence. There is only limited evidence for such costs, but small and essentially undetectable costs would be sufficient to explain the observed polymorphisms.

(*b*) *Matching-allele models*

The gene-for-gene model is convincing because of the range of observations explained. However, a simple alternative, the matching-allele model, can also explain the same observations (Frank 1993*b*). In a matching-allele model with a single haploid locus, each of the n host alleles causes recognition and resistance to only one of the n parasite alleles (Frank 1991*a*). Thus each host is resistant to $1/n$ of the parasite genotypes, and each parasite can attack $(n-1)/n$ of the host genotypes. Figure 1*b* illustrates the matching-allele model for $n = 2$.

In the matching-allele model each parasite genotype functions as either an avirulence allele or a virulence allele depending on the host genotype. By contrast, the gene-for-gene system always has a universal virulence allele that can attack all host genotypes.

Similarly, each host genotype in the matching-allele model functions as either a resistance or a susceptibility allele depending on the parasite genotype. The classical gene-for-gene system always has a universal susceptible genotype that can be attacked by all parasite genotypes.

In terms of biochemical recognition, matching alleles assume a one-to-one correspondence between elicitors and receptors. In the gene-for-gene model the universal virulence allele does not produce an elicitor that can be recognized by any of the available host receptors. Similarly, the universal susceptibility allele does not produce a receptor that can recognize any of the available pathogen elicitors.

Both the matching-allele and gene-for-gene assumptions are consistent with the elicitor–receptor model for the biochemistry of recognition. Matching-allele and gene-for-gene models differ in the range of elicitors and receptors that may exist at each matching pair of host–parasite loci. Thus biochemical evidence on recognition does not distinguish between the models. However, the phenotypic observations of susceptibility and resistance follow the gene-for-gene model (figure 1*a*) rather than the matching-allele model with $n = 2$ (figure 1*b*). This is not surprising because the gene-for-gene model was derived from the phenotypic pattern in figure 1*a*. Appearances can be deceptive, however. One must consider what type of genetic system and pattern of polymorphism would be inferred from samples of the host and parasite populations. The standard procedure is to isolate some host and parasite lines and then test each host against each parasite for resistance or susceptibility.

Here is a reasonable method of classification (see figure 2). (i) Find the host genotype that resists the highest proportion of parasites in the sample. Label that host genotype R for resistant. Only those hosts that resist exactly the same set of parasites are classified as R. (ii) Label all other hosts as S for susceptible. (iii) Label all parasites that cannot attack host genotype R as A for avirulent. (iv) Label all other parasites as V for virulent.

Figure 2 shows that, after following this procedure, one has a classification that is similar to the gene-for-gene system in figure 1a. In figure 2, the h_1 allele is classified as R and the matching p_1 allele is classified as A. What is the frequency of host–parasite pairs that would be misclassified by a gene-for-gene model if the

		A	V		
		p_1	p_2	p_3	p_4
R	h_1	−	+	+	+
S	h_2	+	−	+	+
	h_3	+	+	−	+
	h_4	+	+	+	−

Figure 2. Resistance (−) and susceptibility (+) in a matching allele model with four alleles. It is assumed that the $h_1 : p_1$ match has a frequency greater than or equal to any other diagonal match. Thus $h_1 : p_1$ is arbitrarily labelled as the $R : A$ match for the procedure outlined in the text.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

true system were a matching-allele model with n alleles? The R and A alleles were defined strictly by their response in the sample, so there can be no errors in any host–parasite pair in which the host is classified as R or the parasite as A. All errors must occur when a host–parasite pair, classified as $S : V$, yields a resistant reaction rather than the predicted susceptible response. These errors occur only on the diagonal elements of the $(n-1) \times (n-1)$ submatrix of $S : V$ in figure 2. The frequency of errors in the entire table is $(n-1)/n^2$ at equilibrium. Surprisingly, the data fit the gene-for-gene pattern more closely as the number of alleles, n, increases. (Details and further discussion are in Frank (1993*b*).)

The matching-allele model is consistent with the elicitor–receptor model of biochemical specificity and with the gene-for-gene pattern of phenotypic specificity. The next issue concerns the observed patterns of rare resistance alleles, common virulence alleles, and a low phenotypic frequency of resistance. In the matching-allele model each host allele resists only $1/n$ of the alternative pathogen alleles. The observed frequency of host alleles classified as resistance (R) will tend to be low, and the observed frequency of pathogen alleles classified as virulence (V) will tend to be high (figure 2). Thus the matching-allele model can explain the observed polymorphism and frequency of resistance without invoking costs of resistance and virulence.

These models illustrate three problems (Frank 1993*b*). First, detectable polymorphism is a poor guide to the actual polymorphism. Second, detectable polymorphism is a poor guide to the underlying biochemistry of recognition. For example, if p_3 and p_4 were absent from a sample (figure 2), then host alleles h_3 and h_4 would be grouped as a single allelic type that is susceptible to all pathogen genotypes. This conclusion is misleading about the actual polymorphism and about the specificity of recognition. The third problem is that a low frequency of host resistance and a wide parasite host-range do not imply a lack of coevolutionary pressure. In both gene-for-gene and matching-allele models the frequency of resistance tends to be low even though the widespread polymorphism is maintained by coevolutionary interactions between host and parasite.

These inductive problems are not surprising, particularly after they have been illustrated with simple examples. None the less, they are common mistakes. The next section provides further evidence that detectable polymorphism is misleading when used to infer actual polymorphism and the biochemistry of recognition.

3. CYTOPLASMIC MALE STERILITY IN PLANTS

Most organisms inherit mitochondrial DNA from their mothers, with no input from their fathers. By contrast, most other genetic material is obtained equally from the mother and father. Typically these different modes of transmission, matrilineal versus biparental, have no consequences for the direction of evolutionary change favoured by selection. For example, efficient respiration increases both matrilineal and biparental transmission.

The allocation of resources to sons and daughters affects matrilineal and biparental transmission differently. Traits that enhance the production of daughters at the expense of sons always increase the transmission of matrilineally inherited genes. For example, in some hermaphroditic plants the mitochondrial genes may inhibit pollen development and simultaneously enhance the production of seeds (Edwardson 1970; Hanson 1991). Selection of genetic variants in the mitochondria would favour complete loss of pollen production in exchange for a small increase in seed production because the mitochondrial genes are transmitted only through seeds (Lewis 1941).

Reallocation of resources from pollen to seeds can greatly reduce the transmission of nuclear genes because biparental transmission depends on the sum of the success through seeds and pollen. Thus there is a conflict of interest between the mitochondrial (cytoplasmic) and nuclear genes over the allocation of resources to male (pollen) and female (ovule) reproduction (Gouyon & Couvet 1985; Frank 1989). Consistent with this idea of conflict, nuclear genes often restore male fertility by overcoming the male-sterility effects of the cytoplasm.

The nuclear–cytoplasmic conflict is very similar to a host–parasite system: there is antagonism over resources for reproduction, cytoplasmic (parasite) genes determine the host-range for exploitation, and cytoplasmic genes interact with nuclear (host) resistance genes to determine the specificity of the interaction. Cytoplasmic inheritance influences the patterns of 'parasite' transmission but, on the whole, the genetics and population dynamics are typical of host–parasite interactions (Gouyon & Couvet 1985; Frank 1989; Gouyon *et al.* 1991).

The reduction of pollen caused by cytoplasmic genes is called cytoplasmic male sterility (CMS). Laser & Lersten (1972) list reports of CMS in 140 species from 47 genera across 20 families. More than half of these cases occurred naturally, about 20% were uncovered by intraspecific crosses, and the rest were observed in interspecific crosses. Moreover, this listing is an underestimate of the true extent of CMS because detecting a cytoplasmic component to a male sterile phenotype requires genetic analysis of polymorphism (see below).

Wild populations of CMS maintain several distinct cytoplasmic genotypes (cytotypes). Each cytotype is capable of causing male sterility by an apparently different mechanism because each is susceptible to a particular subset of nuclear restorer alleles. Nuclear restorer alleles are typically polymorphic at several loci, with each allele specialized for restoring pollen fertility when associated with particular cytotypes. The observations are summarized in Frank (1989), Couvet *et al.* (1990) and Koelewijn (1993).

CMS has reciprocal genetic specificity of nucleus and cytoplasm and widespread polymorphism. The basic questions of recognition and polymorphism are similar

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

to those in other host–parasite systems. What is the relationship between the polymorphism that is detectable in a sample of plants and the actual distribution of genetic variation? How does the detectable number of genotypes compare with the potential range of polymorphic alternatives? What can be learned about the biochemistry of specificity from the detectable polymorphism?

The genetics of CMS have been studied by segregation ratios from crosses. These data are used to infer the number of cytotypes, the number of nuclear loci, the specificity of nuclear–cytoplasmic interactions in determining phenotype, and the amount of polymorphism in the sample. Without biochemical evidence of specificity, there is no other way to begin an analysis. However, the work is tedious and the segregation analysis is more intuition than algorithm. A final conclusion that, for example, there are at least three but perhaps many more nuclear loci involved provides little information beyond the fact that the multilocus genetics of specificity are complex.

The problem of inferring specificity occurs even if the underlying biochemistry is simple. To study this problem, I built a computer simulation with coevolving nucleus and cytoplasm. I set the number of loci and the specificity of the interaction. I studied a variety of specificities based on the stop-signal theories of plant–pathogen genetics. The idea is that a cytoplasmic male sterility allele interferes with a step in the pathway of pollen development. A matching nuclear allele restores pollen fertility by preventing the action of a specific cytoplasmic allele. (See Frank (1989) for detailed assumptions and a summary of the relevant biological observations.)

In the model I studied most intensively, I assumed that each cytoplasm has two loci. At each of these loci there are four alternative alleles, which are coded as the set $\{000, 001, 010, 100\}$. The haploid cytoplasmic genotype for the two loci is written by concatenating the pair of alleles, yielding a string with six bits, for example 010100. Each cytoplasmic 1 represents a different way in which the cytoplasm can interfere with pollen development. There are six matching nuclear loci. At each diploid locus the 1 allele can act as a restorer and the 0 allele has no effect, with 1 dominant to 0. The phenotype for the six loci can be written as a string with six bits, for example 101111. Each restorer locus is specific for one of the six possible methods of cytoplasmic interference in pollen production. A plant is male sterile if the cytoplasm has a single specificity not matched by a restorer. In terms of the bit-strings, if the cytoplasm has one or more unmatched 1s, then the plant is male sterile. The interaction between the example cytoplasmic and male sterility genotypes would yield male sterility because the 1 at the second (from the left) cytoplasmic locus is unmatched by its nuclear restorer locus.

Figure 3 illustrates some problems of inferring polymorphism and specificity. The columns in the left box show three different cytotypes (the cytoplasmic 'loci' do not recombine). The rows show the four nuclear phenotypes that occur when there is polymorphism at the first and sixth loci. Perfectly executed crossing experiments and analysis yield the classification of male sterile (female) and male fertile (hermaphrodite) phenotypes in the figure, and the following conclusions. (i) There appear to be only two cytotypes because nuclear polymorphism in the sample cannot separate cytotypes B and C. Additional samples with more nuclear polymorphism would increase the inferred number of cytotypes. (ii) The first nuclear restorer locus appears to act independently of cytotype. This locus would be classified as a purely nuclear male sterility factor until a cytotype with a 0 in the first position was discovered. (iii) The system would be classified as purely nuclear control if the cytotype A were absent.

The right-hand box in figure 3 emphasizes a different set of problems. (i) Cytotype D appears to be a fully male-fertile cytotype. Additional nuclear polymorphism will change this inference about cytoplasmic effects. (ii) It appears that cytotype A requires simultaneous effects from two nuclear loci to restore pollen fertility. Additional cytoplasmic polymorphism would show that the nuclear loci have independent effects.

These examples and the further details in Frank (1989) highlight three problems of induction. First, the detectable polymorphism in a sample underestimates the actual polymorphism in the sample. Second, the inferred genetic system based on detectable polymorphism is a poor guide to the true specificity. Third, even the actual polymorphism in a sample is a poor guide to the number of loci and the specificity of the interaction. This occurs because nonequilibrium fluctuations in allele frequencies cause many alleles to be rare or absent at any particular point in time and space.

Why does it matter if the detectable and actual polymorphism differ in a sample? There is no problem if the goal is pattern description: a locus with no effect in the current context can be ignored when describing the current pattern. But suppose the goal is to explain why polymorphism is maintained and why the frequencies of genotypes and male sterile plants vary widely across space. The chain of reasoning is: detectable polymorphism → inferred specificity → model to explain evolutionary dynamics → predictions about polymorphism and spatial variation → comparison with detectable polymorph-

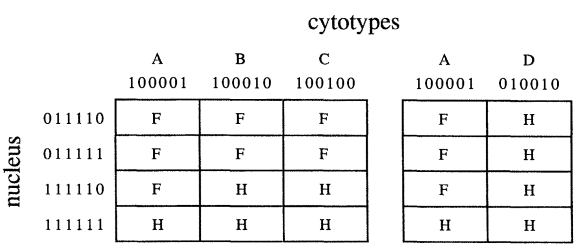

Figure 3. Hermaphrodites (H) and females (F) determined by interaction between nuclear and cytoplasmic genotypes. An individual is female (male sterile) if the cytotype has at least a single 1 matched to a nuclear 0.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

ism. This seems sufficiently circular that it should converge on the truth. However, a model based on inferred specificity is almost certain to be the wrong model, so any match of prediction with observation is probably spurious.

The ideal is to build a model of dynamics based on the biochemistry of specificity. That choice is not yet available because the biochemistry is unknown. At this point a reasonable guess about specificity, such as matching alleles, is perhaps more likely to provide a true model than an approach that begins with detectable polymorphism. An approach that insists on starting with detectable polymorphism inevitably leads to introductory statements such as (Connor & Charlesworth 1989): 'The genetics of male-sterility in gynodioecious [mixed hermaphrodite and male-sterile] species has not been easily or often solved ... For studies of gynodioecism the deficiency lies in genetic solutions to actual problems and not in the need for theoretical models'. If only it were so easy! Based on their own genetic solutions they conclude their paper with

> The solution offered here for male-sterility control in *C. selloana* based on up to three complementary loci with recessive alleles, extends the number and kind of control systems for male-sterility in flowering plants. The results suggest, too, that in the original population probably many other loci are segregating for recessive sterility factors. This agreement with the results of Van Damme for dicotyledonous *Plantago lanceolata* is therefore very striking. It is as yet unclear why gynodioecious populations have so many sterility, or restorer, loci polymorphic.

The weakness of these conclusions is typical of many papers that focus exclusively on segregation analysis to determine the genetic control of male sterility. To paraphrase: the genetics are complicated, many loci are involved, and we don't know why.

Theory has provided plausible explanations of polymorphism (Charlesworth 1981; Delannay *et al.* 1981; Gouyon & Couvet 1985; Frank 1989; Gouyon *et al.* 1991). Further progress will probably require biochemical analysis of recognition and specificity between competing nuclear and cytoplasmic genes.

4. BACTERIA AND THEIR VIRAL PARASITES

I have emphasized the importance of recognition and specificity in host–parasite genetics. Bacterial communities are promising model systems because the biochemistry of recognition is relatively easy to study when compared with CMS or plant–pathogen interactions. In this section I describe the natural history of bacterial restriction–modification (R–M) enzymes, which are used to defend against viral attack. I discuss a surprising outcome of bacterial–viral coevolution, in which parasitic attack leads to widespread genetic polymorphism in the bacterial defence system and the near-extinction of the parasites.

Bacteria have a simple recognition-based immunity system that protects them from invasion by foreign DNA (Wilson & Murray 1991). There are two components to the system. Restriction enzymes cut DNA molecules that carry a particular sequence of nucleotides. Modification enzymes recognize the same nucleotide sequence but, instead of cutting the DNA, these enzymes modify the recognition site in a way that protects that molecule from restriction. A bacterial cell's own DNA is modified, otherwise the restriction enzymes would cut the DNA and kill the cell.

R–M enzymes are known for over 200 different recognition sites (Kessler & Manta 1990; Roberts 1990). Circumstantial evidence suggests that defence against bacteriophage viruses has been a powerful force promoting diversity. First, R–M can protect host cells from invading phages (Luria & Human 1952; Arber 1965). Secondly, phages that develop in bacteria with a particular R–M type are modified for the associated recognition sequence. These modified phages can attack other bacteria of the same R–M type, but are sensitive to restriction by different R–M systems. Rare R–M types are favoured because few phages will be modified for their recognition sequence. This frequency-dependent selection promotes diversity of R–M as a defence against phages (Levin 1986, 1988). Thirdly, phages carry a variety of antirestriction mechanisms (Kruger & Bickle 1983; Sharp 1986; Korona *et al.* 1993). For example, many phages lack particular R–M recognition sequences. The probability of having these recognition sequences is very high if no selective pressure is acting on sequence composition.

The circumstantial evidence favours phage-mediated selection as an explanation for R–M diversity. However, direct studies of interactions between phages and bacteria suggest that bacteria resist phage attack by modifying the receptor sites at which phages adsorb and enter the cell (Lenski 1984, 1988; Lenski & Levin 1985). In these studies R–M apparently has little effect on the long-term dynamics of phage and bacteria; this results suggests that R–M diversity may be maintained by processes other than phage-mediated selection in stable communities (Korona & Levin 1993).

Laboratory studies of phages and bacteria maintained in chemostats provide repeatable observations about coevolution between phages and bacteria (Lenski 1988; Korona & Levin 1993). Phages and bacteria are mixed to begin the experiment. No matter what the short-term dynamics, the bacteria usually evolve a set of surface receptors that resist attack by phages. These modified receptors may reduce host growth rate because the receptors used by phages are typically the site for uptake of important nutrients.

With the appearance of receptor-based resistance, the community settles to a balance of resistant bacteria with reduced growth and sensitive bacteria with phage-induced mortality (Levin *et al.* 1977). In these communities, resistant bacteria typically outnumber sensitive bacteria, and there is a small phage population supported on the sensitive strain (Lenski 1988).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The outcome of evolution in laboratory communities can be summarized as follows. If receptor-resistance is rare during the early phases of the experiment, R–M may provide some defence against phage. As the experiment proceeds, receptor-based resistance becomes common, phage become rare, and R–M loses its selective advantage.

The observations from natural populations provide conflicting evidence about the role of R–M. On the one hand, phages are rare in natural isolates (Scarpino 1978) and receptor-based resistance is common (Lenski & Levin 1985). These observations support the view that phage-mediated selection is a very weak force in the maintenance of R–M diversity. On the other hand, phages often carry anti-restriction mechanisms, suggesting that R–M is an important selective force on phage and that, in turn, phages probably influence R–M diversity.

Levin (1986, 1988; Korona & Levin 1993) suggested that the conflicting evidence can be explained by a model in which R–M is advantageous in colonizing new habitats where phages are common and receptor-based resistance for the local phage has not yet evolved. As the newly established community matures, receptor-based resistance spreads and eventually dominates. Thus R–M diversity is maintained by cycles of selection that occur during colonization.

I suggested an alternative model to explain the apparent contradiction that, in some mature communities, phages are rare and R–M systems are diverse (Frank 1994*a*). My model showed that variation in R–M diversity is itself a direct cause of community structure.

The goal is to explain the community mixture of phages, bacteria with receptor-based resistance to the phages, and bacteria with only R–M defence. Computer studies showed that the abundances of the different phages and bacteria settle to steady values in a mature community, so I describe the results in terms of the community equilibrium.

The main conclusions are shown in figure 4. There are N different bacterial R–M types. Each R–M type has a matching subpopulation of phage that was born in that type; the matching phages have modified DNA that protects them from the defences of that R–M type. The equilibrium abundance of phage modified for each of the R–M types is p^*, thus the total abundance of phage is Np^*. The abundance of each R–M type is h^*, the total abundance is Nh^*. The abundance of the bacteria with receptor-based resistance to the phage is h_r. The last important parameter is d, which is most strongly affected by the amounts of nutrients available for bacterial growth (Frank 1994*a*).

The effects of increasing R–M diversity on community composition can be divided into three stages. Each stage is labelled by a circled number in the panels of figure 4.

1. As the number of R–M types increases from $N = 1$, the abundances of phages and phage-sensitive R–M types increase and the abundance of receptor-resistant bacteria decreases. A rare R–M type always invades an equilibrium community with phages because none of the phages are specialized (modified) for the new type. Each new R–M type increases to the point where it maintains its own phage subpopulation that limits the further spread of that R–M type. Phage limitation of each R–M type has two consequences. First, each new R–M type causes an approximately linear increase in the total abundance of phages and R–M types. Second, the resources taken by each new R–M type reduce the abundance of the receptor-resistant population but do not interfere with other R–M types.

2. The receptor-resistant bacteria are eventually driven to extinction when a sufficient number of R–M types have accumulated. Novel R–M types can continue to invade. Each new type causes a reduction in the phage population. This reduction is probably caused by the high proportion of phage deaths that result when the phages invade and are restricted by bacteria for which the phage DNA is unmodified. As N rises, the proportion of bacteria that are resistant to each particular phage increases, $(N-1)/N$.

3. Further increase in the number of R–M types drives the phages to extinction. At this point, each phage is matched to such a small proportion $(1/N)$ of the bacterial population that phage death rate exceeds the rate of new births. The stable community at the transition between stages 2 and 3 supports a diversity of R–M types but no phages.

The role of habitat quality can be seen by comparing the left panel, with relatively low quality, and the right panel, with relatively high quality.

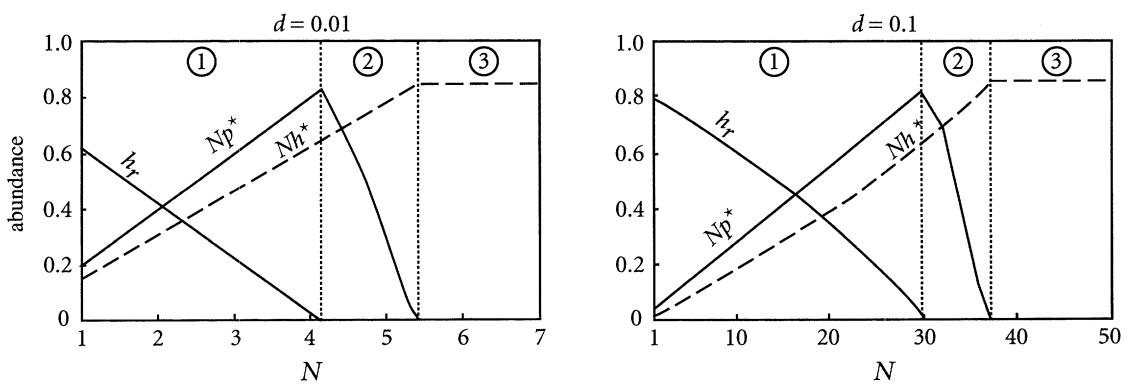

Figure 4. Community composition as a function of the number of R–M types, N.

Richer habitats favour the receptor-resistant bacteria because the competition for nutrients is weaker and the reduced nutrient uptake imposes a lower cost. Thus richer habitats require a greater number of R–M types to drive the receptor-resistant bacteria to extinction.

The interesting aspect of the model is that R–M diversity can have a strong influence on community composition. For a fixed level of nutrients, an increase in R–M diversity causes a shift from a community with few phages and relative dominance by receptor-resistant bacteria to a community in which receptor-based resistance is rare and phages are common. A further increase in R–M diversity can drive the phages to extinction. In a laboratory experiment this sequence would lead to an endpoint that, at first sight, would seem strange: a community of diverse R–M types but an absence of the selective pressure (the phages).

Host diversity drives the parasite to extinction because of the biochemistry of recognition and specificity. Each parasite can attack only $1/N$ of the R–M types. As N increases, the number of hosts available to each phage drops to the point where deaths exceed births, and the phages are driven to extinction.

5. BACTERIOCINS AND ALLELOPATHY

In this section I outline the natural history of another interesting bacterial system. The problem concerns competition between bacterial strains rather than what is usually thought of as a host–parasite interaction. However, the genetic specificity of attack and defence promotes widespread polymorphism in much the same way as in host–parasite systems. Many of the inference problems mentioned above apply to this system, but for the sake of brevity I limit myself to description.

Bacteria often carry plasmids that encode a bacterial toxin (bacteriocin) and immunity to that toxin (Reeves 1972; Hardy 1975; Lewin 1977). Immunity works by neutralizing the toxin after it has entered the cell. Bacteria may also be resistant to bacteriocins because they lack a compatible receptor through which the toxin can enter the cell.

Many distinct bacteriocin types are found within a population. A type is defined by its susceptibility to a set of toxin-producing test strains. With n test strains, there are 2^n possible types. Epidemiological studies frequently use bacteriocin typing to identify and follow pathogenic strains of bacteria. These studies provide information about the diversity of bacteriocin production and susceptibility in populations. For example, Chhibber *et al.* (1988) summarize data on the number of isolates, test strains, and bacteriocin susceptibilities for ten studies of *Klebsiella pneumoniae*. The smallest number of observed types occurred in a study with 200 isolates, four test strains, and 11 types of a possible $2^4 = 16$; the most occurred in a study with 553 isolates, seven test strains, and 64 types of a possible $2^7 = 128$. Similar levels of diversity have been reported for a variety of species (Gaston *et al.* 1989; Senior & Vörös 1989; Rocha & Uzeda 1990; Traub 1991; Riley & Gordon 1992).

The patterns of resistance and susceptibility are determined by the distribution of toxins, specific immunity, and receptor-based resistance. There appears to be widespread diversity for all three components. Only one model has attempted to explain bacteriocin diversity (Frank 1994*b*). That model emphasized the diversifying effects of varying habitat quality: toxins are relatively more effective in resource-rich habitats. The model provides a basis for further studies, but does not explain the diversity of receptor-based resistance and specific immunity found in natural populations.

Although this system is perhaps the least well understood of my four examples, it may also be the most promising for future studies. Recent molecular work has provided sequences for genes involved in toxin production (Riley 1993*a*,*b*). As molecular libraries are developed for different toxin, immunity and receptor alleles, it will be possible to measure directly the distribution of polymorphism without the confusion of inferring genotype from complex phenotypic interactions. In addition, laboratory experiments can be used to test predictions about the ecological and genetic interactions that determine the distribution of diversity.

6. DISCUSSION

Coevolutionary systems are difficult to study because the biochemistry of recognition and the associated host–parasite specificity cannot be inferred from phenotypic interactions and detectable polymorphism. For example, in cytoplasmic male sterility the same nuclear polymorphisms may be classified as controlling male sterility without cytoplasmic interaction, as working in pairs to restore fertility for a particular cytoplasm, or as having a one-to-one specificity with a range of cytotypes. These different classifications depend on the cytoplasmic polymorphism available to classify the phenotypes of the nuclear alleles.

Another point of potential confusion is that intense parasite pressure can lead to low frequencies of host resistance. This occurs, for example, in a matching-allele model where each of n host alleles recognizes and resists only $1/n$ of the parasite alleles. An increase in the number of alleles, n, reduces the chance that any individual host will match a particular parasite. Thus parasite pressure causes diversification of both host and parasite, with low frequencies of host resistance and widespread parasitism.

A different outcome occurs when the specificity causes each host type to resist all but one of the parasite types. For example, in the simple bacterial restriction–modification model, each of the N viral types can attack only $1/N$ of the hosts. As the number of types increases, the viral host-range shrinks. If N is sufficiently large, the virus is driven to extinction. Thus parasite pressure causes widespread polymorphism in the host and, eventually, a great reduction in the abundance of parasites.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

In each of my four examples I have simplified the natural history and the available theory in order to emphasize these general points about recognition and polymorphism. For example, in restriction–modification systems the viruses have an array of anti-restriction mechanisms, including a lack of the DNA sequence cut by restriction enzymes (Kruger & Bickle 1983; Sharp 1986). In this case the biochemistry of recognition is clear, but very little is known about the distribution of polymorphism in natural populations. Recent surveys by Korona *et al.* (1993) show that natural isolates of the viruses carry a diverse array of defences against restriction. Much more work of this kind is needed, with emphasis on joint studies of biochemistry and natural diversity.

On the theoretical side, there is often a call for greater realism to match the complexity of host–parasite genetics. More realism, by itself, will provide little insight. The simplified theories that I have presented emphasize general trends among these complex systems. The trends may not hold, but they must be replaced by equally simple tendencies if there is to be any general science of host–parasite genetics. Of course, each individual system will have its own peculiar natural history and dynamics that must be studied independently. However, the uniqueness of each system should not obscure the fact that there are only a limited number of ways for hosts to recognize parasites, and only a limited number of patterns for matching host resistance to parasite host-range.

I have ignored evolutionary dynamics and spatial variation in order to focus on recognition and polymorphism. It is clear that non-equilibrium fluctuations in time and space play a crucial role in host–parasite genetics. Many properties of dynamics have been studied with models (Seger 1992). Two predictions may be useful for a wide variety of host–parasite interactions.

First, the capacity of the parasite population for explosive growth is often the most important factor controlling the magnitude of fluctuations in population sizes and in gene frequencies (May & Anderson 1983; Frank 1991*a*, 1993*a*). Fast-growing parasites often cause local epidemics. These epidemics may be started by a limited set of genotypes that happens to avoid the local host resistance. These few pathogen genotypes favour a shift in the host population to resistant genotypes. The net effect is changing population sizes and gene frequencies in each local population, with spatially isolated populations cycling through epidemics and genotypes at different times (Gouyon & Couvet 1985; Frank 1989, 1991*b*, 1993*a*; Burdon *et al.* 1989, 1990; Thompson & Burdon 1992). Locally novel parasite genotypes that cause epidemics are likely to arrive by migration from other patches, and host resistance to counter the new parasites also may be introduced by migration. Thus local extinctions of genotypes and subsequent recolonizations from other patches can drive a continual space-time turnover of gene frequencies. By contrast, slow-growing diseases are likely to have more stable population sizes and gene frequencies, with greater diversity in each population and less spatial variation (Frank 1991*a*, 1993*a*).

The second prediction is that an increasing number of genotypes in a host–parasite interaction causes greater space-time variation (Seger 1988; Frank 1989, 1991*a*; Hamilton *et al.* 1990). For example, in a matching-allele model with only two alternative types ($n = 2$), each local population is likely to have both of the host resistance alleles. This prevents a parasite genotype from spreading rapidly because the host population has a specific resistance factor for all parasite genotypes. By contrast, when there are many allelic variants, a local population of hosts is likely to be missing one or more variants. The matching parasite can invade and spread rapidly, driving out the other parasite genotypes and strongly favouring the introduction and spread of the matching host allele. Thus, the greater the number of possible genotypes, the greater the potential for extensive space-time variation. An increase in the number of possible genotypes may also cause a decline in local variation by enhancing the tendency for rapid turnover of locally successful genotypes.

Each of my examples and general conclusions focuses on host–parasite systems with simple genetic specificities. By contrast, many hosts defend against parasites by inducing a variety of non-specific defences and by acquiring specific recognition with antibody selection. It is difficult to piece together a complete evolutionary story for recognition and polymorphism for these systems because there is rarely sufficient information about both host and parasite. The appeal of my four examples is their simplicity and the natural history data that suggest how host–parasite coevolution works.

There is much interest in the coevolution of different kinds of host–parasite and plant–herbivore interactions (Wakelin & Blackwell 1988; Harvell 1990*a*,*b*; Crawley 1992; Fritz & Simms 1992). However, for these systems it is difficult to achieve even the limited realism of my examples. Some recent models have laid the groundwork for the coevolutionary genetics of quantitative traits (Seger 1992; Saloniemi 1993; Frank 1993*d*, 1994*c*) and systems in which hosts induce non-specific chemical and structural defences in response to attack (Clark & Harvell 1992; Frank 1993*c*). Further progress will benefit greatly from a few systems in which specificity and polymorphism are described for both the host and the parasite.

For comments on the manuscript I thank R. M. Bush and M. A. Riley's discussion group at Yale: M. Feldgarden, Y. Tan, J. Wernegreen and E. Wright. My research is supported by NSF grant DEB-9057331 and NIH grants GM42403 and BRSG-S07-RR07008.

REFERENCES

Arber, W. 1965 Host controlled modification of bacteriophage. *A. Rev. Microbiol.* **19**, 365–368.

Burdon, J.J. 1987 *Diseases and plant population biology*. Cambridge University Press.

Burdon, J.J., Brown, A.H.D. & Jarosz, A.M. 1990 The

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

spatial scale of genetic interactions in host–pathogen coevolved systems. In *Pests, pathogens and plant communities* (ed. J. J. Burdon & S. R. Leather), pp. 233–247. Oxford: Blackwell Scientific Publications.

Burdon, J.J., Jarosz, A.M. & Kirby, G.C. 1989 Pattern and patchiness in plant–pathogen interactions – causes and consequences. *A. Rev. Ecol. Syst.* **20**, 119–136.

Charlesworth, D. 1981 A further study of the problem of the maintenance of females in gynodioecious species. *Heredity, Lond.* **46**, 27–39.

Chhibber, S., Goel, A., Kapoor, N., Saxena, M. & Vadehra, D.V. 1988 Bacteriocin (klebocin) typing of clinical isolates of *Klebsiella pneumoniae*. *Eur. J. Epidemiol.* **4**, 115–118.

Christ, B.J., Person, C.O. & Pope, D.D. 1987 The genetic determination of variation in pathogenicity. In *Populations of plant pathogens: their dynamics and genetics* (ed. M. S. Wolfe & C. E. Caten), pp. 7–19. Oxford: Blackwell Scientific Publications.

Clark, C.W. & Harvell, C.D. 1992 Inducible defenses and the allocation of resources – a minimal model. *Am. Nat.* **139**, 521–539.

Connor, H.E. & Charlesworth, D. 1989 Genetics of male-sterility in gynodioecious *Cortaderia* (Gramineae). *Heredity, Lond.* **63**, 373–382.

Couvet, D., Atlan, A., Belhassen, E., Gliddon, C., Gouyon, P.H. & Kjellberg, F. 1990 Co-evolution between two symbionts: the case of cytoplasmic male-sterility in higher plants. *Oxf. Surv. Evol. Biol.* **7**, 225–249.

Crawley, M.J. (ed). 1992 *Natural enemies: the population biology of predators, parasites and diseases*. Oxford: Blackwell Scientific.

Delannay, X., Gouyon, P.-H. & Valdeyron, G. 1981 Mathematical study of the evolution of gynodioecy with cytoplasmic inheritance under the effect of a nuclear restorer gene. *Genetics* **99**, 169–181.

Edwardson, J.R. 1970 Cytoplasmic male sterility. *Bot. Rev.* **36**, 341–420.

Fincham, J.R.S., Day, P.R. & Radford, A. 1979 *Fungal genetics*, 4th edn. Berkeley: University of California Press.

Flor, H.H. 1956 The complementary genic systems in flax and flax rust. *Adv. Genet.* **8**, 29–54.

Flor, H.H. 1971 Current status of the gene-for-gene concept. *A. Rev. Phytopathol.* **9**, 275–296.

Frank, S.A. 1989 The evolutionary dynamics of cytoplasmic male sterility. *Am. Nat.* **133**, 345–376.

Frank, S.A. 1991*a* Ecological and genetic models of host–pathogen coevolution. *Heredity, Lond.* **67**, 73–83.

Frank, S.A. 1991*b* Spatial variation in coevolutionary dynamics. *Evol. Ecol.* **5**, 193–217.

Frank, S.A. 1992 Models of plant–pathogen coevolution. *Trends Genet.* **8**, 213–219.

Frank, S.A. 1993*a* Coevolutionary genetics of plants and pathogens. *Evol. Ecol.* **7**, 45–75.

Frank, S.A. 1993*b* Specificity versus detectable polymorphism in host–parasite genetics. *Proc. R. Soc. Lond.* B **254**, 191–197.

Frank, S.A. 1993*c* A model of inducible defense. *Evolution* **47**, 325–327.

Frank, S.A. 1993*d* Evolution of host–parasite diversity. *Evolution.* **47**, 1721–1732.

Frank, S.A. 1994*a* Polymorphism of bacterial restriction-modification systems: the advantage of diversity. *Evolution.* (In the press.)

Frank, S.A. 1994*b* Spatial polymorphism of bacteriocins and other allelopathic traits. *Evol. Ecol.* **8**, 369–386.

Frank, S.A. 1994*c* Coevolutionary genetics of hosts and parasites with quantitative inheritance. *Evol. Ecol.* **8**, 74–94.

Fritz, R.S. & Simms, E.L. 1992 *Plant resistance to herbivores and pathogens*. Chicago: University of Chicago Press.

Gabriel, D.W. & Rolfe, B.G. 1990 Working models of specific recognition in plant-microbe interactions. *A. Rev. Phytopathol.* **28**, 365–391.

Gaston, M.A., Strickland, M.A., Ayling-Smith, B.A. & Pitt, T.L. 1989 Epidemiological typing of *Enterobacter aerogenes*. *J. clin. Microbiol.* **27**, 564–565.

Gouyon, P.-H. & Couvet, D. 1985 Selfish cytoplasm and adaptation: variations in the reproductive system of thyme. In *Structure and functioning of plant populations* (ed. J. Haeck & J. W. Woldendorp), pp. 299–319. New York: North-Holland.

Gouyon, P.-H., Vichot, F. & Van Damme, J.M.M. 1991 Nuclear-cytoplasmic male sterility: single-point equilibria versus limit cycles. *Am. Nat.* **137**, 498–514.

Hamilton, W.D., Axelrod, R. & Tanese, R. 1990 Sexual reproduction as an adaptation to resist parasites (a review). *Proc. natn. Acad. Sci. U.S.A.* **87**, 3566–3573.

Hanson, M.R. 1991 Plant mitochondrial mutations and male sterility. *A. Rev. Genet.* **25**, 461–486.

Hardy, K.G. 1975 Colicinogeny and related phenomena. *Bacteriol. Rev.* **39**, 464–515.

Harvell, C.D. 1990*a* The evolution of inducible defence. *Parasitology* **100**, S53–S61.

Harvell, C.D. 1990*b* The ecology and evolution of inducible defenses. *Q. Rev. Biol.* **65**, 323–340.

Kessler, C. & Manta, V. 1990 Specificity of restriction endonucleases and DNA modification methyltransferases – a review. *Gene* **92**, 1–248.

Koelewijn, H.P. 1993 On the genetics and ecology of sexual reproduction in *Plantago coronopus*. Ph.D. thesis, University of Groningen, The Netherlands.

Korona, R., Korona, B. & Levin, B.R. 1993 Sensitivity of naturally occurring coliphages to Type I and Type II restriction and modification. *J. gen. Microbiol.* **139**, 1283–1290.

Korona, R. & Levin, B.R. 1993 Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution* **47**, 556–575.

Kruger, D.H. & Bickle, T.A. 1983 Bacteriophage survival: multiple mechanisms for avoiding deoxyribonucleic acid restriction systems of their hosts. *Microbiol. Rev.* **47**, 345–360.

Laser, K.D. & Lersten, N.R. 1972 Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. *Bot. Rev.* **38**, 425–454.

Lawrence, G.J., Mayo, G.M.E. & Shepherd, K.W. 1981 Interactions between genes controlling pathogenicity in the flax rust fungus. *Phytopathology* **71**, 12–19.

Lenski, R.E. 1984 Coevolution of bacteria and phage: are there endless cycles of bacterial defenses and phage counterdefenses? *J. theor. Biol.* **108**, 319–325.

Lenski, R.E. 1988 Dynamics of interactions between bacteria and virulent bacteriophage. *Adv. microb. Ecol.* **10**, 1–44.

Lenski, R.E. & Levin, B.R. 1985 Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *Am. Nat.* **125**, 585–602.

Leonard, K.J. & Czochor, R.J. 1980 Theory of genetic interactions among populations of plants and their pathogens. *A. Rev. Phytopathol.* **18**, 237–258.

Levin, B.R. 1986 Restriction-modification immunity and the maintenance of genetic diversity in bacterial populations. In *Evolutionary processes and evolutionary theory* (ed. S. Karlin and E. Nevo), pp. 669–688. New York: Academic Press.

Levin, B.R. 1988 Frequency-dependent selection in

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

bacterial populations. *Phil. Trans. R. Soc. Lond.* B **319**, 459–472.
Levin, B.R., Stewart, F.M. & Chao, L. 1977 Resource-limited growth, competition, and predation: a model and experimental studies with bacteria and bacteriophage. *Am. Nat.* **111**, 3–24.
Levin, S.A. 1983 Some approaches to the modelling of coevolutionary interactions. In *Coevolution* (ed. M. H. Nitecki), pp. 21–65. Chicago: University of Chicago Press.
Lewin, B. 1977 *Gene expression*, (*Plasmids and phages*). New York: Wiley.
Lewis, D. 1941 Male sterility in natural populations of hermaphrodite plants: the equilibrium between females and hermaphrodites to be expected with different types of inheritance. *New Phytol.* **40**, 56–63.
Luria, S.E. & Human, M.L. 1952 A non-hereditary host-induced variation in bacterial viruses. *J. Bact.* **64**, 557–559.
May, R.M. & Anderson, R.M. 1983 Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. Lond.* B **219**, 281–313.
Person, C. & Mayo, G.M.E. 1974 Genetic limitations on models of specific interactions between a host and its parasite. *Can. J. Bot.* **52**, 1339–1347.
Potts, W.K. & Wakeland, E.K. 1993 Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends Genet.* **9**, 408–412.
Reeves, P. 1972 *The bacteriocins*. New York: Springer-Verlag.
Richards, A.J. 1986 *Plant breeding systems*. London: Unwin Hyman.
Riley, M.A. 1993*a* Positive selection for colicin diversity in bacteria. *Molec. Biol. Evol.* **10**, 1048–1059.
Riley, M.A. 1993*b* Molecular mechanism of colicin evolution. *Molec. Biol. Evol.* **10**, 1380–1395.
Riley, M.A. & Gordon, D.M. 1992 A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col-plasmid lineages. *J. gen. Microbiol.* **138**, 1345–1352.
Roberts, R.J. 1990 Restriction enzymes and their isoschizomers. *Nucl. Acids Res.* **18** (Suppl.), 2331–2365.
Rocha, E.R. & de Uzeda, M. 1990 Antagonism among *Bacteriodes fragilis* group strains isolated from middle ear exudates from patients with chronic suppurative otitis media. *Ear Nose Throat J.* **69**, 614–618.
Saloniemi, I. 1993 A coevolutionary predator-prey model with quantitative characters. *Am. Nat.* **141**, 880–896.
Scarpino, P.V. 1978 Bacteriophage indicators. In *Indicators of viruses in water and food* (ed. G. Berg), pp. 201–227. Ann Arbor, Michigan: Ann Arbor Science.
Seger, J. 1988 Dynamics of some simple host–parasite models with more than two genotypes in each species. *Phil. Trans. R. Soc. Lond.* B **319**, 541–555.
Seger, J. 1992 Evolution of exploiter-victim relationships. In *Natural enemies: the population biology of predators, parasites and diseases* (ed. M. J. Crawley), pp. 3–25. Oxford: Blackwell Scientific.
Senior, B.W. & Vörös, S. 1989 Discovery of new morganocin types of *Morganella morganii* in strains of diverse serotype and the apparent independence of bacteriocin type from serotype of strains. *J. med. Microbiol.* **29**, 89–93.
Sharp, P.M. 1986 Molecular evolution of bacteriophages: evidence of selection against the recognition sites of host restriction enzymes. *Molec. Biol. Evol.* **3**, 75–83.
Thompson, J.N. & Burdon, J.J. 1992 Gene-for-gene coevolution between plants and parasites. *Nature, Lond.* **360**, 121–125.
Traub, W.H. 1991 Bacteriocin typing and biotyping of clinical isolates of *Serratia marcescens*. *Int. J. med. Microbiol.* **275**, 474–486.
Vanderplank, J.E. 1968 *Disease resistance in plants*. New York: Academic Press.
Wakelin, D.M. & Blackwell, J.M. 1988 *Genetics of resistance to bacterial and parasitic infection*. London: Taylor & Francis.
Wilson, G.G. & Murray, N.E. 1991 Restriction and modification systems. *A. Rev. Genet.* **25**, 585–627.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES