Evolutionary Ecology, 1996, **10**, 307–317

Statistical properties of polymorphism in host–parasite genetics

STEVEN A. FRANK
Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA

Summary

Host–parasite interactions often have complex dynamics. At the level of individual allele frequencies, the dynamics are difficult to predict and difficult to measure. However, aggregate properties of polymorphism, such as allelic diversity or the frequency of resistance, may be relatively easy to work with. I study this problem with computer simulations of a host–parasite model. In one example, the simulations show that the allelic diversity at a locus is similar in a host–parasite model and a neutral model in which drift is the only evolutionary process. Allelic diversity is similar in the two models, even though the temporal dynamics of individual allele frequencies are very different. In a second example, the genetic system that would be inferred from analysing samples of hosts and parasites is quite different from the actual specificity that determines the dynamics of the system. Thus, general conclusions about the specificity of host–parasite genetics must be analysed in the context of the expected statistical distributions of polymorphism. The final example shows that the frequency of resistance provides an interesting aggregate measure of host–parasite polymorphism. If the ratio of parasite generation time to the time between the reproductive seasons of the hosts is small, then no regular periodicity in the frequency of resistance occurs. However, if parasites have many generations per reproductive season of the host, then resistance fluctuates with a period equal to the seasonality of the host. The important role of seasonality shown here differs from the emphasis in previous theories on the relative generation times of host and parasite.

Keywords: population genetics; disease; resistance; statistics; non-linear dynamics

Introduction

Simple host–parasite interactions can have complex or chaotic dynamics (May, 1986). This important discovery has changed the way the pattern is interpreted. For example, the idea that polymorphic host resistance alleles can be stably maintained by a delicate balance of opposing forces has given way to a list of temporal and spatial possibilities: limit cycles, strange attractors, fractal dimensions, colonization–extinction dynamics and metapopulation theory.

These new theories about dynamical complexity are fascinating, but the possibilities of testing alternatives for host–parasite genetics are limited by the difficulty of collecting sufficient data. The first point of this paper is to illustrate that a statistical approach to polymorphism can help to bridge the gap between hopelessly complex dynamics and measurable patterns with meaning. I show, with computer simulations of a host–parasite model, that certain aggregate properties of polymorphism and host–parasite dynamics follow simple patterns. One surprising conclusion is that allelic diversity is similar in a model of host–parasite co-evolution and a neutral model, in spite of the fact that the dynamics of allele frequencies are very different between the two models.

The second point of this paper is to show that the pattern of host–parasite specificity cannot be inferred directly from reported observations in the literature. Parker (1994) has recently argued that the data from plant–pathogen interactions support a particular, asymmetric form of specificity. In Parker's (1994) summary, pathogens carry some alleles that can be matched by

0269–7653

host alleles, leading to host resistance, but there are also some alternative pathogen alleles that are unmatched and confer a universal host range. Hosts carry some resistance alleles that match pathogen alleles, but there are also alternative, unmatched host alleles that lead to universal susceptibility.

Parker (1994) contrasted this observed, asymmetric pattern of specificity with the assumptions of models that have been used to study whether parasites can favour the evolution of sexual reproduction. Those models typically assume that each host allele has a matching parasite allele (Hamilton *et al.*, 1990). Parker (1994) showed that symmetric, matching specificity favours sex much more strongly than the asymmetric specificity observed in plant–pathogen interactions. He concluded that models based on the observed genetics of specificity do not support the idea that pathogens are the crucial selective pressure favouring sex. I will show that the statistical properties of polymorphism in host–parasite systems make it very difficult to infer whether genetic specificity is truly symmetric or asymmetric. Thus, although Parker (1994) has raised an interesting and important problem, interpretation requires a deeper understanding of host–parasite genetics than is presently available.

The third point of this paper is that seasonality of host reproduction combined with a short generation time for parasites can lead to seasonal fluctuations in the frequency of host resistance. This point arose while I was searching for properties of host–parasite polymorphism that showed statistical regularity. The idea that the relative generation time of hosts and parasites is important for the intensity and fluctuation of parasite pressure has long been a part of the parasite theory of sex (Hamilton *et al.*, 1990). However, I found that the relative generation time alone had little effect on the aggregate statistical properties of host–parasite polymorphism and dynamics. By contrast, the scaling of host seasonality and parasite generation time leads to simple and potentially testable predictions about host–parasite dynamics.

Assumptions and methods

I chose to study a very simple host–parasite interaction in which the number of loci and alleles per locus could be increased easily. The purpose is to focus on the aggregate properties of polymorphism when each allele is only a small part of a larger system.

The host and parasite are haploid, with n alleles labelled $1, \ldots, n$ at each of L loci. When sexual reproduction occurs, pairs of haploid individuals fuse, recombine and segregate immediately into haploid offspring. The recombination rate is 0.5 between all loci. Each of the L host and parasite loci are paired. If, at matching loci, the host and parasite have alleles with the same numeric label, then the host can resist that parasite. One match at any of the L loci is sufficient for resistance. The parasite can attack the host only if all L loci are unmatched. During reproduction each locus mutates at rate μ, with an allele changing with equal probability to any of the other n-1 possibilities. The mutation process here encompasses actual mutational changes and the immigration rate of alleles, thus I will refer to the process as 'mutation–immigration'. The contribution of immigration is measured as the uncorrelated rate of introduction of alleles relative to the current distribution of allelic variation.

The simulation is stochastic with fixed population sizes for the host and parasite. I used a population size of 500 unless noted otherwise. The life cycles and interactions occur on a discrete time scale. In each step, each of the hosts is tested for resistance or susceptibility against each of the parasites. A host's fitness is proportional to the number of parasites that it resists. A parasite's fitness is proportional to the number of hosts that it can attack. The parasites reproduce in each time step, with reproduction set to either sexual or asexual for a particular run of the model. For sexual reproduction, pairs of parents are chosen stochastically in proportion to the fitness. The

parents fuse, recombine and segregate. One of the progeny is added to the next generation and the process is repeated until the population for the next generation is filled. For asexual reproduction, a parent is chosen stochastically in proportion to its fitness and a replica is added to the population for the next time step. Mutation–immigration occurs after reproduction but before any interactions in the next time step.

Two parameters control the timing of host reproduction. First, the parameter g determines seasonality. Hosts can reproduce only during every gth time step. Second, the parameter h determines the proportion of hosts added to the $g + 1$ time step that were created by reproduction in the gth time step. In some time steps, some of the hosts are carried forward without reproduction, for example, if it is not the season for reproduction or if hosts have overlapping generations with only a fraction h reproducing in each reproductive season. The hosts that are carried forward to the next time step without reproduction are sampled stochastically from the current period in proportion to their fitness (resistance to parasites). The sampling is done with replacement, so that carrying forward is essentially asexual reproduction without mutation or immigration. Hosts have a mean generation time of g/h, with overlapping generations if h is less than 1. Hosts always reproduce sexually, with changes caused by mutation and immigration occurring before the start of the next time step.

To summarize the parameters, each host and parasite is haploid with L loci and n alleles per locus. The population sizes of the host and parasite are set at 500 for most runs. The parasite population is either entirely sexual or entirely asexual for a given run. Mutation–immigration occurs at rate μ. The hosts reproduce seasonally every gth time step, with only a fraction h of the hosts being newborns in each reproductive period.

The populations were initialized with randomly chosen genotypes and then the model was run for 2000 generations without collecting statistics. In the following 2000 generations, 20 equally spaced samples were taken and a variety of statistics about polymorphism and resistance were measured, as discussed in the following sections. In addition, some time-series data were collected in each of the 2000 generations. Most of the results were taken from a factorial design with the number of loci $L = 2^1, 2^3, 2^5$, the number of alleles $n = 2^1, 2^3, 2^5$, the mutation–immigration rate $\mu = 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}$, the seasonality parameter $g = 2^1, 2^3, 2^5$, the age-structure parameter $h = 1/2^0, 1/2^2, 1/2^4$ and the parasites either entirely sexual or entirely asexual for a given run. This design, with 810 runs, was repeated a second time but with all host and parasite fitnesses assigned constant values in each time step. This second set is a study of polymorphism in a neutral model with population sizes and life cycles the same as in the host–parasite model.

The computer programs used are available upon request.

Statistical properties of polymorphism

Distribution of allele frequencies

A common measure of genetic diversity at a locus is $D = \sum p_i^2$, where p_i is the frequency of the ith allele. This measure is referred to as Simpson's index in the ecological literature (Magurran, 1988; see Weir (1990) for its use in genetics). When $1/D$ is used, the value ranges from 1, when there is a single allele at a locus, to n, where n is the number of equally frequent alleles at a locus. Thus, $1/D$ is diversity when measured in the number of equally frequent allele equivalents.

The symmetric host–parasite model that I studied has frequency dependence that tends to favour rare alleles. The only possible stable equilibrium is all host and parasite alleles equally frequent at each locus. Thus, one reasonable hypothesis about an aggregate measure of

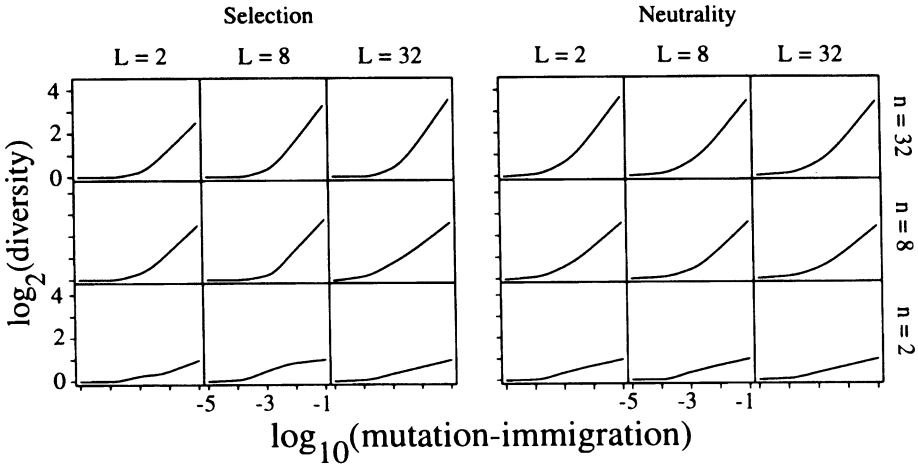

Figure 1. The allelic diversity of hosts when there is co-evolution with parasites (left panels) and when allelic effects are neutral (right panels). Diversity is measured by $1/D$, where D is Simpson's index (see the text). Diversity is measured at only one of the L loci in each sampling period; the same locus was always chosen. Each panel shows a smoothed, locally weighted regression curve fitted to the observations for the design described in the text. For each mutation–immigration rate, number of loci and number of alleles, there are $3 \times 3 \times 2 = 18$ observations corresponding to the three levels of g, the three levels of h and whether the parasites are sexual or asexual. An observation is the median from the 20 samples of each run. The effects of g and h were small (data not shown). An increase in g, the number of parasite generations per host season, decreases diversity in both the selective and neutral models because host mutation occurs only during reproduction. The magnitude of reduced diversity is small and is consistent with the reduced level of mutation. The proportion of hosts breeding in each season, h, and the parasite sexual system had little effect on allelic diversity.

polymorphism is that a locus involved in a host–parasite system will maintain significantly greater allelic diversity than a neutral locus.

Figure 1 shows that there is very little difference in the allelic diversity under host–parasite selection (left panels) and neutrality (right panels). Thus, although the dynamics of genotype frequency are quite different between the selective and neutral models, as shown below, allelic diversity in this host–parasite model is indistinguishable from neutrality. It is difficult to be certain about the processes of co-evolution that lead to this similarity. One possibility is that, in the host–parasite model, the trajectories of individual allele frequencies move near the boundary (frequency of 0) and, with small population sizes, sampling variance causes local extinction of those alleles. Increasing the number of alleles in a host–parasite system tends, on the one hand, to favour many equally frequent but rare alleles because of frequency-dependent selection. On the other hand, the more alleles, the closer the allele frequencies will tend toward zero along their trajectories and the more frequent will be local extinctions of alleles.

One possibility is that very large population sizes could maintain greater diversity in the host–parasite system than in a neutral system. For both systems the effect of drift would decrease. In the host–parasite system, weak frequency dependence would, with sufficiently large populations, be sufficient to protect alleles from extinction even when their frequencies are close to zero. Simulations with large population sizes are very time-consuming because each host individual must be tested against each parasite individual, so that execution time increases with the square of population size. The largest populations that I studied had 2000 host individuals and 2000

parasite individuals, compared with population sizes of 500 for the results shown in Fig. 1. With the larger populations of 2000, both selected and neutral systems maintained more diversity, as expected. However, there was very little difference between the diversity in the host–parasite system and the neutral model. Where differences existed, the neutral model maintained more diversity for larger numbers of alleles per locus ($n = 8,32$) and the host–parasite model maintained more variation for smaller numbers of alleles per locus ($n = 2$).

Although 2000 is a relatively small population, the effective size of large populations in a host–parasite system can be greatly reduced by fluctuating frequencies of resistance. Occasionally, unmatched parasite genotypes that can attack most hosts sweep through the system; hosts counter by a rapid increase in genotypes that match and resist the dominant parasite genotype. Some evidence from the simulations is given below that suggests these kinds of rapid change. Thus, even when population sizes are stable because disease is not a regulating factor, fitness differences caused by differing resistance patterns can create silent, powerful epidemic fluctuations in genotype frequencies that greatly reduce effective population sizes.

Inferred genetic system

Models that study whether host–parasite co-evolution can favour the evolution of sex typically assume a matching recognition system between the host and parasite alleles (Hamilton, 1980, 1993; Hamilton *et al.*, 1990). One assumption is that a parasite allele must match a particular host allele for compatibility and invasion. A second, alternative assumption is that a host allele must match a particular parasite allele for recognition and defence. In the first case, frequency dependence will favour rare host alleles. Rare type advantage diversifies the hosts which, in turn, favours matching diversity in the parasites. In the second case, where hosts must match a parasite for resistance, rare parasite alleles are favoured. This rare type advantage diversifies the parasites which, in turn, favours matching diversity in the hosts.

Models that favour sex because of parasite pressure require that the favoured combination of resistance alleles changes over time (Hamilton *et al.*, 1990). However, the favoured patterns of linkage disequilibrium can change only if there is sufficient allelic diversity at each locus and there is no universal advantage to particular combinations of alleles. Thus, Parker (1994) has argued that models with matching alleles, in which widespread diversity is maintained purely by frequency dependence, are the most conducive to favouring sexual reproduction as a way to escape parasitic attack. However, frequency dependence is not the only mechanism that can maintain allelic diversity. A host allele causing resistance to some parasite genotypes may carry a pleiotropic fitness cost relative to a universally susceptible alternative allele. The balance between the costs and benefits for resistance can maintain a resistance-susceptibility polymorphism. Similarly, a parasite allele conferring a universal host range may carry a pleiotropic fitness cost relative to an allele with a narrow host range. The balance between costs and benefits of a universal host range can maintain polymorphism.

Parker (1994) reviewed observations from plant–pathogen interactions and concluded that the evidence overwhelmingly favoured balanced polymorphisms maintained by costs instead of matching polymorphisms maintained purely by frequency dependence. He then compared the extent to which sex is favoured in host–parasite models with balanced or matching polymorphisms. He concluded that the matching polymorphisms favour sex relatively easily, whereas the balanced polymorphisms are not conducive to sexual reproduction. Because polymorphisms appear to be maintained by balancing selection in plant–pathogen systems, Parker (1994) concluded that parasites are probably not the main factor favouring sex in those systems.

In a previous study, I found that it is very difficult to differentiate between matching allele and cost models for the maintenance of polymorphism (Frank, 1993). Matching allele models often

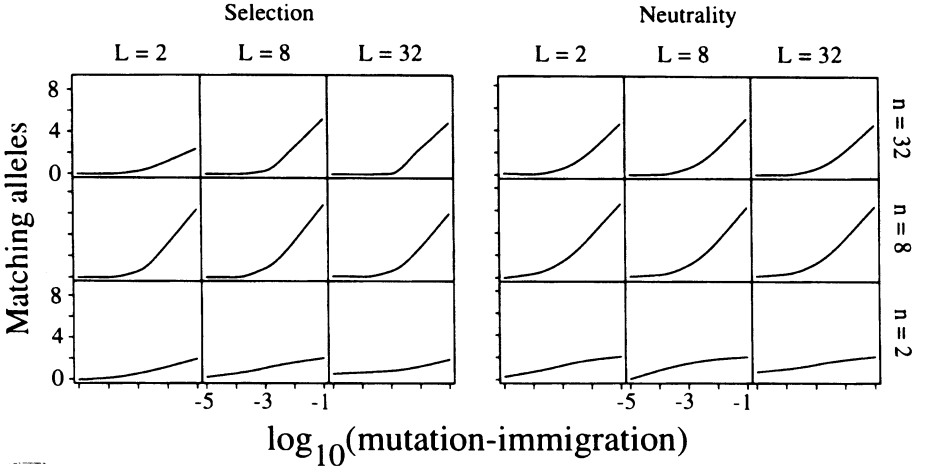

Figure 2. The number of matching alleles inferred from samples of 25 hosts and 25 parasites. The design and methods are the same as in Fig. 1. Also as in Fig. 1, the parameter g influenced the observed pattern through its influence on the mutation–immigration rate; h and the sexuality of the parasite had little effect on the overall trends (data not shown).

led to observations of universally susceptible host alleles or to parasite alleles with a universal host range. Thus, pure frequency dependence is sufficient to explain the observed universal susceptibility or universal host-range alleles; it is not necessary to invoke costs.

I extended my past work by considering the inferred genetic system of host–parasite interactions as an aggregate property of polymorphism. I defined three components of the genetic system that can be inferred from a sample of hosts and parasites. First, the number of matching alleles per locus is the number of host alleles that have different phenotypic effects depending on the sampled parasite alleles at that locus. Second, the frequency of universal-susceptibility alleles at each host locus is the frequency of all alleles at that locus that do not match any parasite alleles. In this model, an unmatched host allele leads to susceptibility. Third, the frequency of universal host-range alleles at each parasite locus is the frequency of all alleles at that locus that do match any host alleles. An unmatched parasite allele leads to successful attack by the parasite.

Figure 2 (left panels) shows the number of matching alleles inferred from a random sample of 25 hosts and 25 parasites. In each case I analysed the matching alleles at one of the L loci, always using the same locus. When mutation or immigration of new alleles is $< 10^{-3}$, the number of matching alleles is usually 0 or 1. The right panels of Fig. 2 from a neutral model shows that the inferred number of matching alleles in a sample is similar under strong host–parasite selection or drift. (A host–parasite match occurs when the numeric labels of the alleles at a locus are the same.)

The inferred number of matching alleles depends on sample size. Larger samples include more rare alleles, which increase the inferred number of matching alleles. In the computer methods I used, a single copy of a host allele in the sample matched with a single copy of a parasite allele is sufficient to be counted. In practice, finding every matching pair of alleles in a sample is very difficult. First, one has to backcross both the host and parasite so that the effects of the locus under study are not masked by the epistatic effects of other loci. In the matching allele system, the host resists attack if it matches the parasite at any one of the loci involved in the interaction. After backcrossing to expose the effects of a single locus, there are N^2 pairwise interactions,

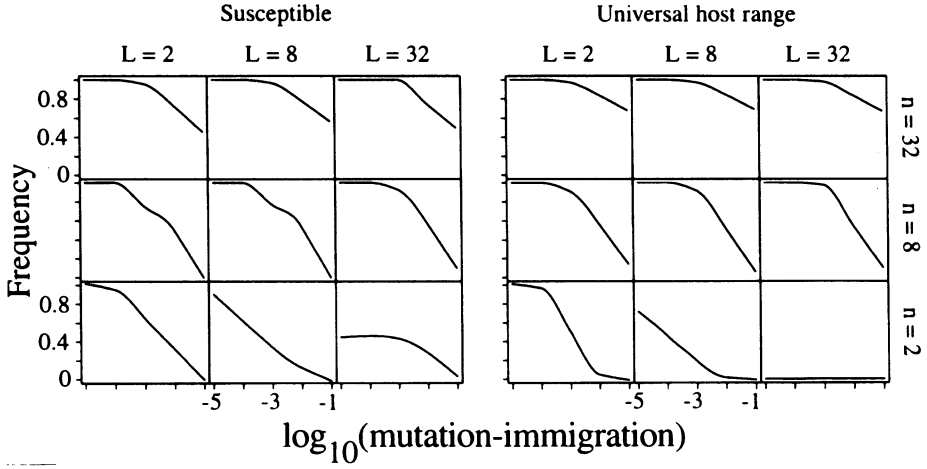

Figure 3. The inferred frequencies of universal-susceptible alleles in hosts and universal host-range alleles in parasites in samples of 25 hosts and 25 parasites. The design and methods are the same as in Fig. 1. Also as in Fig. 1, the parameter g influenced the observed patterns through its influence on the mutation–immigration rate; h and the sexuality of the parasite had little effect on the overall trends. The results for a neutral model were similar (data not shown).

where N is the sample size. Thus, Fig. 2 undoubtedly overestimates the number of matching alleles that would actually be inferred from samples of 25 hosts and 25 parasites.

The largest sampling that I am aware of from natural populations of plants and pathogens are the studies by Bevan *et al.* (1993a,b). They analysed 50 host lines of *Senecio vulgaris* when tested against 24 pathogen isolates of the powdery mildew *Erysiphe fischeri*. The host lines were inbred, undoubtedly causing some loss of genetic variability. That loss, combined with the inherent practical difficulties of detecting rare events in any study, suggests that the information in their sample is probably somewhat less than obtained from my computer analysis of 25 hosts and parasites. More extensive samples have been obtained in agricultural studies. However, whether the pathogens collected from crop epidemics and the host varieties are representative of natural variation is unclear and the detection and analysis of rare events is not discussed.

It is certainly possible to obtain larger samples. The point here is that the polymorphism detected depends on the range of potential variation (the parameters L and n), the intensity of migration and mutation (μ), the dynamics of natural populations and the sampling and experimental methods. Thus, any strong claims about the actual genetic system must be weighed carefully against the difficulty of obtaining all the required information.

When zero matching alleles are observed, the locus would not be recognized as contributing to reciprocal co-evolutionary pressures between the host and parasite. The classical gene for gene model from plant genetics (Flor, 1955, 1971), as described in the current literature (Burdon, 1987), has one matching allele pair at each locus. In addition to the matching allele, the host in a gene for gene model has a universal-susceptible allele and the parasite has a universal host-range allele.

Parker (1994) noted that alleles with universal effects are typical of plant–pathogen systems, whereas they appear to contradict the assumptions of a purely matching system of host and parasite alleles. Figure 3 shows the inferred frequency of universal-susceptible alleles in the host

and universal host-range alleles in the parasite. The methods were the same as in Fig. 2. A single locus was analysed in a random sample of 25 hosts and 25 parasites. All host alleles that did not match any parasite alleles in the sample were classified as a single, universally-susceptible allele because they could not be distinguished by observation. In the samples, most host alleles would be classified in this category of universal susceptibility because the matching parasite alleles were absent. Similarly, parasite alleles were classified as a single, universal host-range allele if they were not matched by any host alleles. In most cases, the majority of parasite alleles were classified as having a universal host range.

In summary, the data from natural populations of plant–pathogen interactions do not allow one to distinguish between the matching and balanced models of polymorphism. Matching specificity would lead to a high observed frequency of universal-susceptible alleles and universal host-range alleles. Alternatively, models that assumed some host alleles never matched parasites and some parasite alleles are never matched by hosts would also be consistent with the observations (Frank, 1993, 1994). Thus, although Parker (1994) raised an interesting problem by showing that balanced polymorphisms are less conducive to sex than purely frequency-dependent polymorphisms, the data do not allow one to determine which model best describes plant–pathogen genetics.

Rhythms of sex and disease

In the previous sections I have shown that several statistical properties of polymorphism are very similar in neutral models and in models with intense co-evolutionary interactions between host and parasite. I was surprised by these results, because I had intended to show that simple aggregate properties of polymorphism differed in predictable ways between neutral and host–parasite models. Although those aggregate properties have turned out to be similar, there must be strong differences in the dynamics. However, my goal was to show that one does not have to resort to the complex trajectories of individual allele and genotype frequencies to study interesting properties of host–parasite polymorphism. Put another way, I wanted to find an appropriate scale of analysis at which there is a simple and meaningful pattern. I found one particularly interesting example in the seasonality of host reproduction, which I report in this section.

The frequency of host resistance when each host individual is tested against each parasite individual is an important aggregate measure of polymorphism. The frequency of resistance fluctuates according to the complex dynamics of matching pairs of host–parasite alleles and the changing statistical associations of alleles at different loci (linkage disequilibrium). Models in which sex can be favoured by parasite pressure depend on sex and recombination to provide the hosts with a way of increasing their resistance through altering linkage associations among loci (Hamilton *et al.*, 1990).

I examined the temporal fluctuations in the frequency of resistance for different scaling relations between host seasonality and parasite generation time. Recall that parameter g is the number of parasite generations between the hosts' reproductive seasons. Parameter h is the proportion of hosts that reproduce in each season. Figure 4 shows the temporal autocorrelations for the frequency of resistance. The rows of panels, from top to bottom, have seasonality ratios of $g = 2$, 8, and 32, respectively. The left column is for host–parasite co-evolution and the right column is for a neutral model. (Autocorrelation is the correlation between successive observations in a time series. The time lag is the time between pairs of observations used to form the correlation. For example, if the lag is 3, then pairs of observations at times 1 and 4, 2 and 5, 3 and 6 and so on, are used to form the bivariate data from which the correlation coefficient is calculated.)

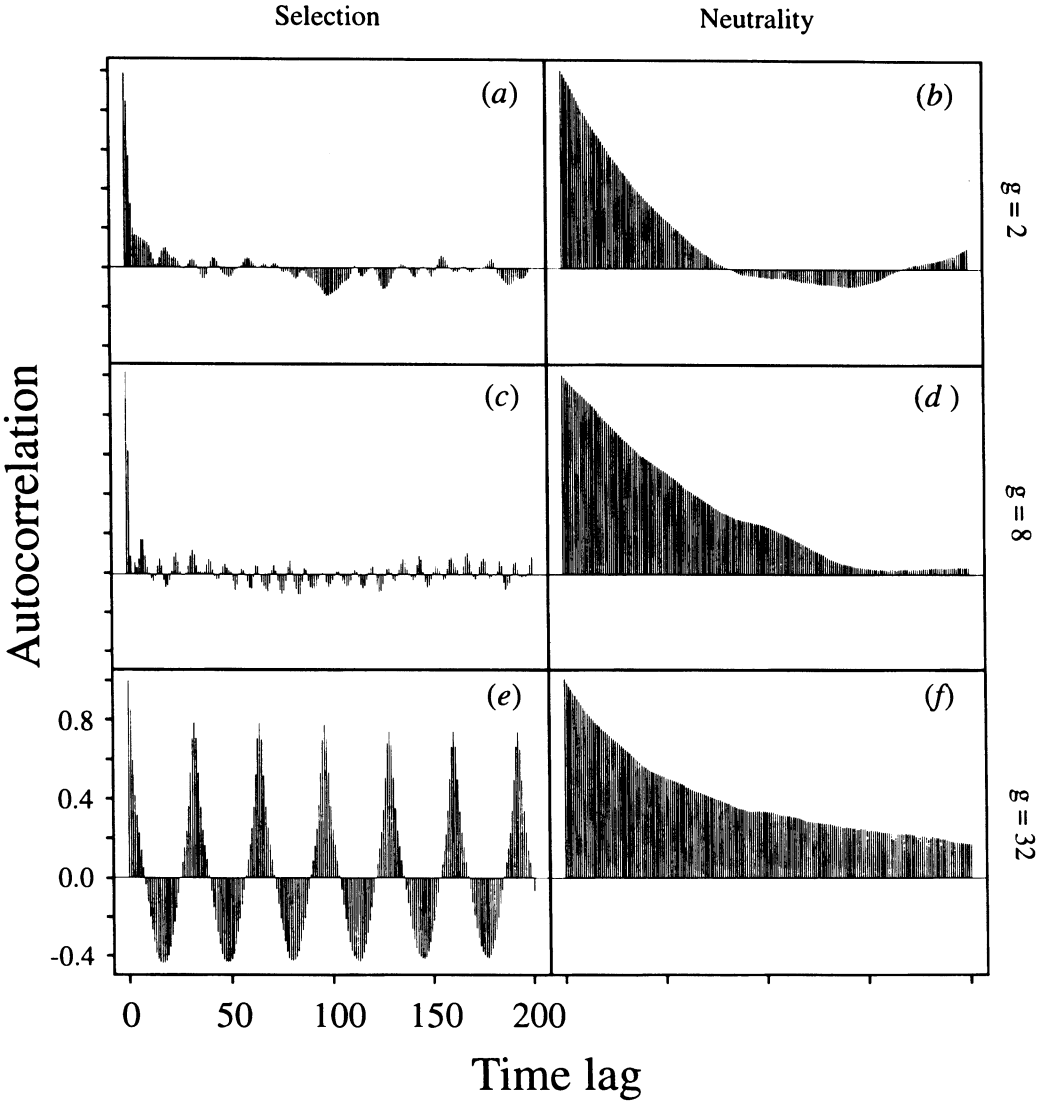

Figure 4. Temporal autocorrelation for the frequency of resistance. For the 2000 generations in which data were collected in a run, all host individuals were tested against all parasite individuals. The left column is a host–parasite model, whereas the right column is a neutral model. The seasonality parameters, from the top row to the bottom row, are $g = 2$, 8 and 32, respectively. The other parameters are $n = L = 8$, $\mu = 10^{-2}$, $h = 0.25$ and sexual parasites.

Looking first at the right column for the neutral model, the rate of decay in the autocorrelation decreases as g increases. That decrease occurs because, with longer time periods between host reproduction, the mutation–immigration rate declines. Less extrinsic perturbation from mutation–immigration maintains higher correlations over time. In addition, statistical associations between the host and parasite genotypic distributions can arise by sampling. These associations decay during reproduction, so the longer the time between reproductive periods, the slower the decay.

In the left column, seasonality ratios of 2 or 8 lead to a negligible periodic effect on the frequency of resistance. When the ratio increases to 32 parasite generations per host season (Fig. 4e), there is a powerful seasonal effect on the frequency of resistance. Apparently, the genotypic distribution of the parasites narrows to those linkage patterns that avoid host resistance. When the hosts have a round of sexual reproduction, some offspring are produced by recombination that resist the majority of parasites. These new host genotypes are strongly favoured and increase in frequency, increasing the frequency of resistance. In response, the parasite genotypes that avoid these new hosts increase in frequency. Thus, resistance declines until the next round of sexuality.

The periodicity of resistance as a function of seasonal scaling suggests a simple prediction. For a sample of parasites taken just before host reproduction, resistance should be low for hosts sampled at the same time and considerably higher among the offspring of the next season. The difference in resistance between pre- and post-reproduction hosts against the sample of parasites should increase with increasing seasonal scaling.

Host generation time, by itself, may have little influence. For example, if each host reproduces with probability 1/32 in each parasite generation, the relative generation time of host to parasite is 32. However, large periodic changes in the frequency of resistance are not expected in this case. By contrast, seasonality appears to build up the pressure of linkage associations between host and parasite populations that is released when some hosts recombine and are favoured after seasonal reproduction.

Discussion

I have emphasized a statistical approach to the analysis of host–parasite polymorphism. I used three examples to illustrate the meaning of a 'statistical approach'. In the first example, I compared the allelic diversity at a locus when the locus is part of a host–parasite interaction system with the diversity when the dynamics of allele frequencies are governed entirely by drift in a neutral model. Surprisingly, the amount of diversity is similar in host–parasite and neutral models (Fig. 1), in spite of the fact that the dynamics of individual alleles are very different between the two models. This example shows that aggregate properties of polymorphism may have statistical distributions that are not easily predicted from the dynamics of the components. Because these aggregate properties are much easier to study and compare among systems than temporal dynamics, a theoretical understanding of these aggregate properties is a prerequisite for the analysis of polymorphism in host–parasite systems.

In the second example, I examined Parker's (1994) conclusion about host–parasite specificity. Parker (1994) found that, in all available studies of plant–pathogen specificity, plants have some alleles that do not provide resistance against any known pathogen genotypes. Similarly, pathogens have some alleles that can successfully attack all known host genotypes. Parker (1994) argued that these universal-susceptibility alleles in hosts and universal host-range alleles in pathogens contradict a key assumption in theories about the role of parasites in the evolution of sex. Those theories typically assume that each host allele is matched by a pathogen allele, thus no alleles have a single, universal effect against all genotypes of the opponent (Hamilton *et al.*, 1990; Hamilton, 1993). However, my analysis showed that, if one assumes matching host and parasite alleles, the aggregate statistical properties of polymorphism would lead one to infer a high frequency of universal-susceptibility alleles in hosts and universal host-range alleles in parasites (Fig. 3).

In the third example, I showed that the frequency of resistance is an interesting aggregate property of host–parasite polymorphism (Fig. 4). When the number of parasite generations per reproductive season of the host is small, then the frequency of resistance did not show any clear

periodicity. By contrast, if the parasites reproduce rapidly and have many generations per host season, then strong periodicity in resistance occurs. Parasite pressure builds steadily between seasons and recombination in the hosts releases new resistance genotypes by a change in linkage associations among resistance alleles.

The study of aggregate properties of systems is not new. In physics, the study of aggregate properties is called statistical mechanics. For example, the trajectory of a single molecule in a gas is complex, but pressure and volume follow simple laws. In genetics, each allele contributing to a polygenic, quantitative character fluctuates in frequency in an unpredictable way because of linkage disequilibrium, sampling variance and changes in the environment. However, quantitative genetics have been successful in predicting certain aspects of character evolution as an aggregate property of many loci.

Host–parasite interactions often depend on the dynamics of many polymorphic loci. The dynamics of individual loci are too complex to study directly. However, it may be possible to derive predictions for aggregate properties such as allelic diversity or the frequency of resistance. These statistical properties are easier to understand and to measure than the dynamics of particular alleles. Thus, a framework for analysis requires a model of the host–parasite interaction. From this model, the predicted aggregate properties of polymorphism and resistance can be expressed as statistical distributions. These predicted distributions can then be compared with measured distributions of polymorphism and resistance from natural populations.

Acknowledgements

I thank R. Antia, R.M. Bush and B.R. Levin for comments on the manuscript. My research is supported by NSF grant DEB-9057331 and NIH grant GM42403.

References

Bevan, J.R., Crute, I.R. and Clarke, D.D. (1993a) Variation for virulence in *Erysiphe fischeri* from *Senecio vulgaris*. *Plant Pathol.* **42**, 622–35.

Bevan, J.R., Clarke, D.D. and Crute, I.R. (1993b) Resistance to *Erysiphe fischeri* in two populations of *Senecio vulgaris*. *Plant Pathol.* **42**, 636–46.

Burdon, J.J. (1987) *Diseases and Plant Population Biology*. Cambridge University Press, Cambridge.

Flor, H.H. (1955) Host–Parasite interaction in flax rust–its genetics and other implications. *Phytopathology* **45**, 680–85.

Flor, H.H. (1971) Current status of the gene-for-gene concept. *Ann. Rev. Phytopathol.* **9**, 275–96.

Frank, S.A. (1993) Specificity versus detectable polymorphism in host–parasite genetics. *Proc. R. Soc. Lond. B* **254**, 191–7.

Frank, S.A. (1994) Recognition and polymorphism in host–parasite genetics. *Phil. Trans. R. Soc. Lond. B* **346**, 283–93.

Hamilton, W.D. (1980) Sex versus non-sex versus parasite. *Oikos* **35**, 282–90.

Hamilton, W.D. (1993) Haploid dynamic polymorphism in a host with matching parasites: effects of mutation/subdivision, linkage, and patterns of selection. *J. Heredity* **84**, 328–38.

Hamilton, W.D., Axelrod, R. and Tanese, R. (1990) Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl Acad. Sci. USA* **87**, 3566–73.

Magurran, A.E. (1988) *Ecological Diversity and its Measurement*. Princeton University Press, Princeton.

May, R.M. (1986) When two and two do not make four: nonlinear phenomena in ecology. *Proc. R. Soc. Lond. B* **228**, 241–66.

Parker, M.A. (1994) Pathogens and sex in plants. *Evol. Ecol.* **8**, 560–84.

Weir, B.S. (1990) *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.