BRIEF COMMUNICATIONS

Evolution, 54(2), 2000, pp. 686–691

EFFECT OF DELETERIOUS MUTATION-ACCUMULATION ON THE FITNESS OF RNA BACTERIOPHAGE MS2

MARCOS DE LA PEÑA,[1,2] SANTIAGO F. ELENA,[1] AND ANDRÉS MOYA[1,3]
[1]*Institut Cavanilles de Biodiversitat i Biología Evolutiva and Departament de Genètica, Universitat de València, Apartado 22085, 46071 València, Spain*

Abstract.—RNA viruses show the highest mutation rate in nautre. It has been extensively demonstrated that, in the absence of purifying selection, RNA viruses accumulate deleterious mutations at a high rate. However, the parameters describing this accumulation are, in general, poorly understood. The present study reports evidences for fitness declines by the accumulation of deleterious mutations in the bacteriophage MS2. We estimated the rate of fitness decline to be as high as 16% per bottleneck transfer. In addition, our results agree with an additive model of fitness effects.

Key words.—Bacteriophage evolution, deleterious mutations, experimental evolution, fitness, RNA viruses.

Received April 9, 1999. Accepted August 20, 1999.

Due to the lack of proofreading by their replicates, RNA viruses show the highest mutation rates in nature (Drake et al. 1998). For this reason, RNA virus populations usually consist of highly heterogeneous mixtures. In a defined environment, the mutant spectrum of a viral population is generally centered around one or several sequences that are more fit, but nonetheless often represent only a minor proportion of the mutant distribution (Domingo and Holland 1988). Thus, during some infectious processes, when a single genome of the heterogeneous population is randomly chosen to replicate and to generate a new population, there is a high probability that the new population carries a slightly deleterious mutation relative to the most fit members of the ancestral population. After further replication, the debilitating mutation may revert or be compensated for by mutations at other sites. However, when the effective population size is small, compensatory mutations might not arise. Continuous sampling or genetic bottleneck events followed by low-fidelity genome replication may lead to substantial fitness losses. Muller (1964) predicted that when mutation rates are high and populations sizes are small, this process occurs in an irreversible, ratchetlike manner that leads to the gradual decrease of fitness in a population, particularly in asexual organisms.

In recent years, increasing evidence of this pernicious effect of deleterious mutations combined with genetic drift has been reported. Chao (1990) found that under demographic conditions in which genetic drift is the major force, mean fitness of the segmented RNA bacteriophage $\phi 6$ systematically decreased. Later, this observation was extended to mammalian RNA viruses. Duarte et al. (1992, 1993, 1994), Clarke et al. (1993), Novella et al. (1995, 1996), and Elena et al. (1996) extended this observation to the nonsegmented RNA virus vesicular stomatitis virus (VSV). These authors made detailed analyses of the role that population size and initial fitness have in the final outcome of the deleterious mutation-accumulation process. Escarmís et al. (1996) made similar observations with another nonsegmented RNA virus, the foot-and-mouth disease virus (FMDV). In addition, they characterized, at the molecular level, which mutations were responsible for the phenotypic fitness changes observed. Most recently, Yuste et al. (1999) also described the negative effect that strong genetic drift and high mutation rates have in the fitness of the retrovirus human immunodeficiency virus type 1 (HIV-1).

Here, we extend the study of Muller's ratchet to populations of the nonsegmented RNA bacteriophage MS2. We also propose a new technique, based on northern blot analyses, that allows competition between two variants of a virus when easily distinguishable phenotypic markers are difficult or impossible to find. In addition, because our method distinguishes the competitors based on their genotypes and not on their phenotypes, the method avoids the problem of phenotypic masking during mixed infections. (Phenotypic masking appears when coat proteins of one genotype encapsulate RNA of the competitor genotype. This problem has been described for the monoclonal antibody detection of MAR mutants in the VSV system; Holland et al. 1989).

MATERIALS AND METHODS

Bacteriophages and Bacterial Host

The bacterial host employed in these experiments was *Escherichia coli* strain GM-1 (*F′lacI*$^{Q}\Delta$*pro-lac ara*$^{-}$ *thi*; Miller et al. 1977). The bacteriophage employed was the coliphage MS2. MS2 belongs to the Leviviridae family and is able to infect male-specific (F^{+}) *E. coli.* Its genome is a positive sense, single-stranded RNA of 3569 nucleotides that encodes four proteins. Two different genotypes of this phage were employed, a wild-type clone and mutant A3. Mutant A3 of MS2 contains a 13-base insertion between positions 1305 and 1306 and a point mutation (C to U) at position 1304 of the wild type genome (Olsthoorn and van Duin 1996). The insertion is located at the maturation-gene terminator helix, a conserved region in all RNA coliphages. This insertion has been demonstrated to be highly stable (Olsthoorn and van Duin 1996). The fitness of the wild-type clone relative to the A3 mutant was estimated to be 3.1732 ± 0.0071

[2] Present address: Institut de Biología Cel·lular i Molecular de Plantes, CSIC-UPV, Camí de Vera 14, 46022 València, Spain.
[3] Corresponding author; E-mail: andres.moya@uv.es.

(mean ± SEM), indicating that the insertion clearly has a negative effect on the life cycle of the A3 mutant (see below for a detailed description of the protocol for fitness estimation).

Stocks of both phages were prepared as described elsewhere (Sambrook et al. 1989) and kept at −80°C with 15% glycerol (v/v). These two stocks constituted the entire viral stock employed in our experiments.

Media and Culture Conditions

All phages and bacteria were grown in LC broth (10 g/L bactotryptone, 5 g/L yeast extract, 8 g/L NaCl, 2 g/L $MgSO_4$, 140 mg/L thymine, and 1 mL/L 1 M Tris-HCl, pH = 7.6). For plating, LC broth was solidified by adding agar at the following concentrations: 1.5% for hard agar (bottom) and 0.7% for soft agar (top), respectively.

Detection of the A3 Mutant by Means of Northern Blot

To determine the ratio of wild type to A3 in a mixed infection, a specific oligonucleotide hybridization in situ assay was performed (Sambrook et al. 1989; Beekwilder et al. 1996). The designed oligonucleotide was 5′-CTATCTG-GATGTTCTATCTGAA-3′, carrying a digoxygenin-uracil group covalently linked to the 3′ end that allowed us to visually detect mutant A3 plaques. In brief, the detection protocol is as follows: first, a nylon membrane (Hybond N+, Amersham, Uppsala, Sweden) was carefully put in contact with the surface of the soft agar plates for 3 min, followed by RNA fixation to the membrane by alkaline denaturation (1 min with 0.1 M NaOH). Then, prehybridization (30 min, room temperature), hybridization (1.5 h, at 37°C with 10 pmol/mL of labeled probe), and stringency washes were performed as described by Beekwilder et al. (1996). Finally, detection of the hybridization signals was performed following the manufacturer's instructions (Boehringer, Mannheim, Germany).

Relative Fitness Assays

The determination of the relative fitness of the wild-type phages relative to the genetically marked A3 ancestral mutant was performed by an adaptation of the method described in Holland et al. (1991). Previously titered populations of both phages were mixed at a 1:1 ratio. Approximately 10^7 viruses from this mixture were used to infect a growing culture of *E. coli* (2 mL, $\sim 10^6$ cells). The infection was kept at 37°C and shaken at 300 rpm overnight. Infections were always done at a multiplicity of infection of 0.1 phages per cell. The resulting lysate was centrifuged (twice, 5 min. at 3000 rpm) and an aliquot of 10 μL of supernatant (containing $\sim 10^7$ viral particles) was used to continue the serial competition passages. After each competition passage, appropriate dilutions of the resulting mixed virus yield were plated on an *E. coli* lawn in triplicate. After overnight incubation, the total number of plaques was scored for each plate and the proportion of A3 mutant plaques was determined by means of the northern blot method described above. These two determinations gave the proportion of wild type to A3 mutant at time t (R_t). Fitness was defined as $W^t = R_t/R_0$ (Chao 1990) and obtained by fitting $\ln W$ to the time series data by the least-squares method (Sokal and Rohlf 1995).

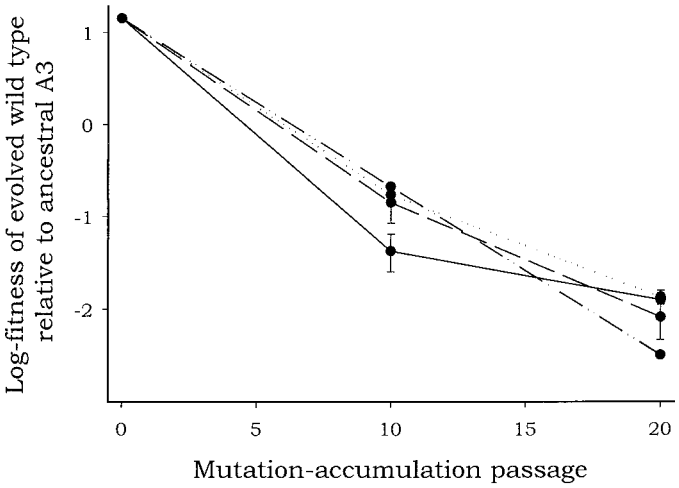

FIG. 1. Change in the natural logarithm of relative fitness observed after 20 consecutive passages through bottlenecks of size one. The bars represent standard errors.

Mutation-Accumulation Process

All plaque-to-plaque transfers were done with the wild-type bacteriophage MS2. From the initial stock, four clearly isolated plaques were randomly picked and separately dispersed in 0.5 mL of LC broth. Each of these four samples was employed to initiate independent passage series. For each series, viruses from a plaque were picked and transferred daily in a plaque-to-plaque regime for a total of 20 days. Each day, the population was generated from a single viral particle and, after a day of growth, the size expanded to $(1.7446 \pm 0.0354) \times 10^9$ viruses/plaque. The relative fitness of virus isolated after 10 and 20 passages was determined as described above.

RESULTS

Figure 1 shows the change in fitness with the number of passages of mutation accumulation for the evolved lines. All four accumulation lines showed significant declines in fitness relative to the ancestral wild-type clone.

Fitness Decline Due to Accumulation of Deleterious Mutations

After 10 passages of mutation-accumulation, the average fitness across lines decreased from the initial value of 3.1732 ± 0.0639 to 0.4158 ± 0.0424. This difference, which was highly significant ($t_{15} = 35.9713$, $P < 0.0001$), represents a loss of fitness of approximately 87%. After 10 more passages of mutation-accumulation (i.e., a total of 20 passages), the fitness of the lines declined even more, reaching a value of 0.1365 ± 0.0110. Obviously, this value was significantly different from that obtained for the ancestral wild-type MS2 ($t_{15} = 46.8478$, $P < 0.0001$). At the end of the experiment, the loss of fitness suffered by the mutation-accumulation lines

TABLE 1. Results of the linear regression analysis performed for each of the four accumulation lines (data from Fig. 1). Under the column $U\bar{s}$, the values of the rate of fitness decline per transfer (± standard error) are reported.

Accumulation line	R^2	F	df	P	$U\bar{s}$
1	0.9229	131.6673	1, 11	< 0.0001	0.1680 ± 0.0146
2	0.9979	4851.8920	1, 10	< 0.0001	0.1824 ± 0.0026
3	0.9854	810.7074	1, 12	< 0.0001	0.1551 ± 0.0054
4	0.9792	517.9010	1, 11	< 0.0001	0.1689 ± 0.0074

was about 96%, which represents an enormous competitive disadvantage.

It is evident from Figure 1 and from the comparison of the mean values across lines obtained after 10 and 20 passages of mutation-accumulation (both values are significantly different: $t_{14} = 6.3794$, 1-tail $P < 0.0001$) that the decline in fitness was constant during the experiment, in a ratchetlike way, as a consequence of the fixation of several deleterious mutations of equal effect, as opposed to a hatchetlike fashion (Kondrashov 1993), which would be compatible with the appearance one rare, extremely deleterious mutation. However, the decline suffered by one of the lines (solid line in Fig. 1) may be compatible with the second possibility: This line shows a large fitness decline in the first 10 passages and almost no change in the last 10 passages.

The Rate of Fitness Decline

According to Lynch and Gabriel (1990) and Kibota and Lynch (1996), and assuming a multiplicative fitness model, the slope of the regression of log-fitness on passage number can be taken as an estimate of the product $U\bar{s}$, where U is the deleterious genomic mutation rate per unit time, in this case, daily transfer, and $\bar{s}$ is the expected reduction in fitness due to any mutation. In other words, this product is the expected rate of fitness decline per transfer due to the introduction of deleterious mutations in the genome of MS2. Independent linear regressions were fitted to each of the accumulation lines shown in Figure 1. Table 1 shows the corresponding regression parameters. All four linear regressions were highly significant, even after applying Bonferroni's correction for multiple comparisons (Rice 1989). An analysis of variance showed that these four estimates of $U\bar{s}$ were homogeneous ($F_{3,21} = 0.6447$, $P = 0.5949$), thus allowing us to pool all of the estimates in a single value. After doing so, the average rate of fitness decline per transfer was estimated to be 0.1595 ± 0.0082. In other words, after each daily transfer, fitness declined, on average, by about 16%.

TABLE 2. Analysis of variance comparing the fit of multiplicative and synergistic models to the data from each independent mutation-accumulation line shown in Figure 1. A partial-F-test was used to assess the significance of adding an additional term to the model. The addition of a synergistic term (β) uses one degree of freedom. SSR, residual sum of squares; df, residual degrees of freedom; P, significance level.

Accumulation line	Model	SSR	df	F	P
1	multiplicative	12.3951	12		
	synergistic	12.0832	11	0.2840	0.6050
2	multiplicative	12.4396	11		
	synergistic	11.9989	10	0.3672	0.5580
3	multiplicative	12.0922	13		
	synergistic	12.0394	12	0.0526	0.8225
4	multiplicative	12.2356	12		
	synergistic	12.1773	11	0.0527	0.8226

Apparent Lack of Epistasis among Deleterious Mutations

In the previous section, we assumed a multiplicative fitness function, which is equivalent to saying that each individual mutation contributes independently to the final decline in fitness. However, it is very important to ascertain whether this assumption is true. For instance, in the mutational deterministic hypothesis for the origin of sexual reproduction (Kimura and Maruyama 1966; Kondrashov 1988), it is theoretically shown that sexual reproduction provides a selective advantage over asexual reproduction only if deleterious mutations interact synergistically, that is, the reduction in fitness suffered by an individual carrying two mutations is larger than that obtained by merely multiplying the effects observed for each mutation independently. Mathematically, this can be described by the equation $\ln W_k = -\alpha k - \beta k^2$, where W_k is the average fitness of a genotype with k mutations, $\alpha > 0$ for deleterious mutations, and β defines the interaction among mutations. If mutations interact synergistically ($\beta > 0$), then the frequency at equilibrium of the less mutated class is higher and the mutation load smaller in a sexual than in an asexual population.

Table 2 shows the results of the partial-F-test (Kleinbaum and Kupper 1978) comparing the goodness of fit of a synergistic versus a multiplicative model. None of the four independent mutation-accumulation lines showed a significant reduction in the residual sum of squares with the addition of a quadratic term to the regression equation. This indicates that it is not necessary to call for a more complex model (the synergistic one) to explain the observed data.

DISCUSSION

The study of viral fitness has an important role in understanding the evolution of viruses. The emergence of antiviral resistant mutants, the screening of new antiviral drugs/vaccines, and the analysis of the evolution of viral populations in different cell types and/or demographic situations can be monitored via studies of fitness variation (Lee and Yin 1996). With our system, we have been able to determine fitness differences between two viral populations easily and after only a maximum of five competition passages. To our knowledge, this work represents the first time that viral fitness measurements have been performed by quantifying the ge-

notypes of both competitors instead of their phenotypes. This helps in avoiding problems like phenotypic mixing and masking (pseudotypes) during mixed infections, where surface proteins of one genotype can enclose the RNA of another genome, as described for the monoclonal antibody detection of MARM virus in the VSV system (Holland et al. 1989). Our system also allows us to evaluate the biological fitness of predesigned mutants (taking into account its evolutionary stability) to determine accurately the replicative effect of any change in the wild-type genome of the phage. Moreover, there is even a chance for biological fitness studies between the viruses belonging to the ssRNA coliphages family (Qβ, GA, SP, etc.). This could help to clarify the phylogenetic relationships between them.

The results reported here with MS2 extend previous observations of the deleterious effects of repeated bottlenecks on the fitness of RNA viruses. Despite important biological differences among the five viral systems studied (a segmented phage, ϕ6; a nonsegmented phage, MS2; two animal viruses, VSV and FMDV; and a retrovirus, HIV-1), in all cases significant decreases in mean fitness have been observed, providing support for Muller's ratchet hypothesis.

The type of data obtained in these experiments allows, in theory, an estimate of $U\bar{s}$, the product of the deleterious genomic mutation rate and the average effect of deleterious mutations. Several different methods have been proposed for independently estimating U and $\bar{s}$ (for a review of the different methods, see Keightley 1998). However, because this was not the main objective of our design (our original objective was to demonstrate the effect of Muller's ratchet in MS2), we did not include enough replicates in our experiments to employ any of these methods. For this reason, we decided to present data on the change in fitness rather than erroneous estimates of its components. In fact, the study of $U\bar{s}$ by itself is of great importance in evolutionary biology. A heated debate surrounds the idea that many mutations of small effect produce a fitness decline of about $U\bar{s} = 1\%$ per generation (Mukai et al. 1972; Ohmishi 1977; Kibota and Lynch 1996; Shabalina et al. 1997) versus the conflicting idea that fewer mutations of larger effect produce smaller declines in fitness of $U\bar{s} \simeq 0.1\%$ per generation (Fernández and López-Fanjul 1996; Keightley 1996; García-Dorado 1997, Keightley and Caballero 1997). The decline in fitness that we estimated for MS2 is 16% per day; however, we must caution against straight comparisons for two reasons. First, there are enormous differences among MS2 and DNA-based organisms in terms of mutation rates and, second, our estimate is on the scale of *per day,* not *per generation.* It is difficult to determine how many generations of viral replication occur each day. However, we can base our study in two different models of viral replication. The first replication model assumes a pure doubling process; the number of generations can be estimated then as $\log_2(1.7446 \times 10^9/1) = 30.7$ generations per day, which gives a per generation estimate of $U\bar{s} \simeq 0.52\%$. A more realistic replication model based on mixed lineal-exponential growth through negative-sense intermediate strains (Schuster and Stadler 1999) gives a much smaller figure of only eight generations per day, therefore $U\bar{s} \simeq 2\%$. These two replication models can be considered as possible extremes. Thus, even at the upper boundary, our estimate is closer to the one supported by the first group of authors than to the second, although the lower limit is somewhat intermediate between both positions. In addition, our lower boundary of a 0.52% fitness decline per generation is comparable with the value of 0.39% recently found by Elena and Moya (1999) in VSV. This similarity between the values obtained for two different RNA viruses gives extra support to the generality of the findings here presented.

Recombination has been reported to occur during MS2 replication (Olsthoorn and van Duin 1996) at a rate of $\sim 10^{-6}$ events per replication round. The possible beneficial effect of recombination in the elimination of deleterious mutations (Muller 1964) may have a limited effect during plaque-to-plaque transfers of MS2 because of the low probability that the same cell could be infected by two debilitated viruses during the development of a plaque. Furthermore, the beneficial effect of recombination on fitness may not be important under certain biological scenarios because even experiments specifically designed to test recombination's effect on fitness recovery found only a minimal advantage (Chao et al. 1997).

Our results indicate that the fitness effects of deleterious mutations in RNA virus MS2 interact in a multiplicative way. This is not the first time that experimental results from RNA viruses have indicated that deleterious mutations interact in a multiplicative way. Elena (1999) also observed that deleterious mutations interact in such way in FMDV. In this case, the number of mutations was directly inferred from sequencing the genome of FMDV (Escarmís et al. 1996). This lack of synergistic interactions among deleterious mutations could be real or the apparent result of two opposing effects: mutations with synergistic effects ($\beta > 0$) and mutations with antagonistic effects on fitness ($\beta < 0$), producing in the former a log-concave-downward and in the latter a log-concave-upward function. This mixed effect has been observed in *E. coli* (Elena and Lenski 1997), in which the number of mutations with synergistic and antagonistic effects is approximately the same, generating, on average, a linear relationship among log-fitness and mutation load.

It is questionable how much of the results about epistasis obtained from experiments with microbes can be extrapolated to more complex organisms. Obviously, the more complex the genome, the more room for interaction among genes. However, three arguments support these kinds of studies. First, these studies intend to evaluate whether sex would be advantageous for the microbe under study. The goal of these studies is ''to determine if there is a general tendency for genetic architectures to exhibit synergistic epistasis among deleterious mutations'' (Elena and Lenski 1997, p. 396). Second, this lack of epistasis has been observed in microorganisms of dramatically different complexity, from simple RNA viruses with only few genes encoded (this study; Elena 1999) to *E. coli* (Elena and Lenski 1997), which represents an enormous jump in complexity when compared with RNA viruses, and, finally, to the fungus *Aspergillus niger* (de Visser et al. 1997a) or the algae *Chlamydomonas moewusii* (de Visser et al. 1996, 1997b), which also represent a tremendous increase in complexity relative to *E. coli.* If we strive for a general theory for the origin of sexual reproduction, we may want to base it on some general and basic principle. Recombination

is common at all levels of complexity, not only in higher organisms, but also in certain microbes. Thus, any basic principle to be used as the foundation of the theory must be common at all organismic levels (i.e., including microbes). Otherwise, we will have only organism-specific theories. Third, how much of the results (if available) from higher organisms can be extrapolated to microbes? Are we assuming that higher organisms are more important than microbes? The number, diversity, and clinical or economic importance of microbes is large enough to justify studies of their evolutionary mechanisms, even at the risk of being unable to extrapolate conclusions to higher organisms, which, from our point of view, is not the case.

Acknowledgments

We thank J. van Duin for kindly providing the phages and host cells and F. González-Candelas and P. E. Turner for critical reading and English correction. This research has been supported by grant PM97-0060-C02-02 from the Spanish Dirección General de Enseñanza Superior e Investigación Científica.

Literature Cited

Beekwilder, J., R. Nieuwenhuizen, A. H. Havelaar, and J. van Duin. 1996. An oligonucleotide hybridization assay for the identification and enumeration of F-specific RNA phages in surface water. J. Appl. Bacteriol. 80:179–186.

Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. Nature 348:454–455.

Chao, L., T. T. Tran, and T. T. Tran. 1997. The advantage of sex in the RNA virus ϕ6. Genetics 147:953–959.

Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland. 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. J. Virol. 67:222–228.

de Visser, J. A. G. M., R. F. Hoekstra, and H. van den Ende. 1996. The effect of sex and deleterious mutations on fitness in *Chlamydomonas.* Proc. R. Soc. Lond. B. 263:193–200.

———. 1997a. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. Evolution 51:1499–1505.

———. 1997b. An experimental test for synergistic epistasis and its application in *Chlamydomonas.* Genetics 145:815–819.

Domingo, E., and J. J. Holland. 1988. High error rates, population equilibrium, and evolution of RNA replication systems. Pp. 3–36 *in* E. Domingo, J. J. Holland, and P. Ahlquist, eds. RNA genetics. Vol. 3. CRC Press, Boca Raton, FL.

Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow. 1998. Rates of spontaneous mutation. Genetics 148:1667–1686.

Duarte, E. A., D. K. Clarke, A. Moya, E. Domingo, and J. J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc. Natl. Acad. Sci. USA 89:6015–6019.

Duarte, E. A., D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland. 1993. Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. J. Virol. 67:3620–3623.

Duarte, E. A., I. S. Novella, S. Ledesma, D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland. 1994. Subclonal components of consensus fitness in an RNA virus clone. J. Virol. 68:4295–4301.

Elena, S. F. 1999. Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus. J. Mol. Evol. 49: 703–707.

Elena, S. F., and R. E. Lenski. 1997. Test of synergistic interactions among deleterious mutations in bacteria. Nature 390:395–398.

Elena, S. F., and A. Moya. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J. Evol. Biol. 12:1078–1088.

Elena, S. F., F. González-Candelas, I. S. Novella, E. A. Duarte, D. K. Clarke, E. Domingo, J. J. Holland, and A. Moya. 1996. Evolution of fitness in experimental populations of vesicular stomatitis virus. Genetics 142:673–679.

Escarmís, C., M. Dávila, N. Charpentier, A. Bracho, A. Moya, and E. Domingo. 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. J. Mol. Biol. 264:255–267.

Fernández, J., and C. López-Fanjul. 1996. Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster.* Genetics 143:829–837.

García-Dorado, A. 1997. The rate and effects distribution of viability mutations in *Drosophila*: minimum distance estimation. Evolution 54:1130–1139.

Holland, J. J., J. C. de la Torre, D. A. Steinhauer, D. K. Clarke, E. A. Duarte, and E. Domingo. 1989. Virus mutation frequencies can be greatly underestimating by monoclonal antibody neutralization of virions. J. Virol. 63:5030–5036.

Holland, J. J., J. C. de la Torre, D. K. Clarke, and E. A. Duarte. 1991. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. J. Virol. 65:2960–2967.

Keightley, P. D. 1996. Nature of deleterious mutation load in *Drosophila.* Genetics 144:1993–1999.

———. 1998. Inference of genome-wide mutation rates and distribution of mutation effects for fitness traits: a simulation study. Genetics 150:1283–1293.

Keightley, P. D., and A. Caballero. 1997. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans.* Proc. Natl. Acad. Sci. USA 94:3823–3827.

Kibota, T. T., and M. Lynch. 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli.* Nature 381: 694–696.

Kimura, M., and T. Maruyama. 1966. The mutation load with epistatic gene interactions in fitness. Genetics 54:1337–1351.

Kleinbaum, D. G., and L. L. Kupper. 1978. Applied regression analysis and other multivariable methods. Duxbury Press, North Scitute, MA.

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336:435–440.

———. 1993. Classification of hypothesis on the advantage of amphimixis. J. Heredity 84:372–387.

Lee, Y., and J. Yin. 1996. Detection of evolving viruses. Nature Biotechnology 14:491–493.

Lynch, M., and W. Gabriel. 1990. Mutation load and the survival of small populations. Evolution 44:1725–1737.

Miller, J. H., D. Ganem, P. Lu, and A. Schmitz. 1977. Genetic studies of the *lac* repressor. I. Correlation of mutational sites with specific amino acid residues: construction of a colinear gene-protein map. J. Mol. Biol. 109:275–301.

Mukai, T., S. I. Chigusa, L. E. Mettler, and J. F. Crow. 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster.* Genetics 72:335–355.

Muller, H. J. 1964. The relation of recombination to mutational advance. Mut. Res. 1:2–9.

Novella, I. S., S. F. Elena, A. Moya, E. Domingo, and J. J. Holland. 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. J. Virol. 69: 2869–2872.

———. 1996. Repeated transfer of small RNA virus populations leading to balanced fitness with infrequent stochastic drift. Mol. Gen. Genet. 252:733–738.

Ohmishi, O. 1977. Spontaneous and ethyl methanosulfonate-induced mutations controlling viability in *Drosophila melanogaster.* II. Homozygous effect of polygenic mutations. Genetics 87: 529–545.

Olsthoorn, R., and J. van Duin. 1996. Random removal of inserts from and RNA genome: selection against single-stranded RNA. J. Virol. 70:729–736.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Sambrook, J. E., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2d ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Schuster, P., and P. F. Stadler. 1999. Nature and evolution of early replicons. Pp 1–24 *in* E. Domingo, R. Webster and J. Holland, eds. Origin and evolution of viruses. Academic Press, London.
Shabalina, S. A., L. Y. Yampolsky, and A. S. Kondrashov. 1997. Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed natural selection. Proc. Natl. Acad. Sci. USA 94:13034–13039.
Sokal, R. R., and F. J. Rolhf. 1995. Biometry. 3rd ed. W. H. Freeman, New York.
Yuste, E., S. Sánchez-Palomino, C. Casado, E. Domingo, and C. López-Galíndez. 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J. Virol. 73:2745–2751.

Corresponding Editor: M. Riley

Evolution, 54(2), 2000, pp. 691–695

BALANCING SELECTION ON A FLORAL POLYMORPHISM

BANU SUBRAMANIAM[1] AND MARK D. RAUSHER[2]
Department of Zoology, Evolution, Ecology and Organismal Biology Group, Box 90325, Duke University, Durham, North Carolina 27708-0325
[2]*E-mail: mrausher@duke.edu*

Abstract.—The common morning glory, *Ipomoea purpurea,* exhibits a flower color polymorphism at the *W* locus throughout the southeastern North America. The *W* locus controls whether flowers will be darkly pigmented (*WW*), lightly pigmented (*Ww*), or white with pigmented rays (*ww*). In this report, we describe results of a perturbation, or convergence, experiment using five plots designed to determine whether balancing selection operates on the *W* locus. The pattern of gene frequency changes obtained are indicative of balancing selection operating at the *W* locus, providing direct evidence that both the alleles are actively maintained by selection.

Key words.—Balancing selection, frequency-dependent selection, *Ipomoea,* morning glory, polymorphism.

Received November 24, 1998. Accepted August 24, 1999.

A major focus of evolutionary biology is to account for the great genetic diversity exhibited by natural populations of most organisms, both at the molecular level and in ecologically important characters. Four general explanations can account for genetic variation in a particular trait: (1) variation reflects passive accumulation of neutral mutations due to genetic drift; (2) it reflects some form of mutation-selection balance; (3) it reflects a transient phase in which natural selection is causing certain genetic variants to replace others; or (4) it reflects some form of balancing selection. At the molecular level, explanations (1) and (4) seem consistent with much DNA and protein sequence variation (Li 1997). For some ecologically important traits, including sex ratio (Uyenoyama and Bengtsson 1979), self-incompatibility systems (Charlesworth and Charlesworth 1979; Clark 1993; Vekemans and Slatkin 1994; Richman et al. 1995), heterostyly (Eckert et al. 1996), and possibly disease resistance (Ennos 1983; Antonovics 1994), selection is believed to be inherently frequency dependent, resulting in balancing selection. For the vast majority of variable ecological traits, however, inherent frequency dependence is not expected and there is little information regarding which explanation accounts for observed variation. In particular, the general importance of balancing selection in maintaining genetic variation in ecologically important characters is unclear.

In this report, we describe results of a perturbation, or convergence, experiment designed to determine whether balancing selection operates on a locus influencing floral pigmentation in the morning glory, *Ipomoea purpurea.* The *W* locus, which controls whether flowers will be darkly pigmented (*WW*), lightly pigmented (*Ww*), or white with pigmented rays (*ww*), is often polymorphic in populations in southeastern North America (Ennos 1981; Ennos and Clegg 1983; Epperson and Clegg 1986). The frequency of the white (*w*) allele typically ranges from 0.0 to 0.4 (Epperson and Clegg 1986). The primary pollinators, bumblebees, discriminate among the color morphs and undervisit whites when they are in the minority, which results in increased selfing by whites, but no detectable pollen discounting (Brown and Clegg 1984; Epperson and Clegg 1987; Rausher et al. 1993; Fry and Rausher 1997). Flower color is thus ecologically important in this species in that it modulates interactions with symbionts and influences the mating system.

METHODS

We performed a convergence experiment in which we established five experimental populations with initial allele frequencies perturbed from their typical values. We then monitored changes in gene frequency over one generation to determine whether gene frequencies tended to converge toward typical, intermediate values, as would be expected if bal-

[1] Present address: Department of Women's Studies and Department of Ecology and Evolutionary Biology, Communications 108, University of Arizona, Tucson, Arizona 85721; E-mail: banu@u.arizona.edu.

ancing selection were operating. Two of the populations (populations 1 and 2) were begun with low frequencies of the white allele ($p_w = 0.13$). This frequency is approximately equal to the mean frequency of the white allele ($p_w = 0.11$) in 19 polymorphic populations reported in a gene frequency survey by Epperson and Clegg (1986). Although a lower initial frequency in these two experimental populations would have been desirable, the value chosen reflected a necessary compromise between the need to begin at a frequency as low as possible and the need to have a sufficient number of plants of the least common (white) genotype to provide sufficient statistical power for detecting selection and a change in gene frequency. A third population (population 5) was begun with equal frequencies of the two alleles, which represents a higher frequency of the white allele than has been recorded in natural populations. Finally, two experimental populations (populations 3 and 4) were begun with high frequencies of the white allele ($p_w = 0.81$). Our a priori expectation, based on gene frequencies observed in natural populations and the assumption that balancing selection operates, was that p_w in populations 3–5 would decrease from high to intermediate values, whereas p_w populations 1 and 2 would increase from low to intermediate values.

Experimental Procedures

Experimental procedures followed those of previous studies (Rausher and Simms 1989; Simms and Rausher 1989; Rausher and Fry 1993). Each experimental population was established by planting a sample of seeds of known genotypic composition at the *W* locus. The seeds were planted in a rectangular grid with one-foot spacing between rows and columns. Individual seeds were randomly assigned to positions in the grid and were allowed to grow and reproduce naturally. After capsules began to mature, plants were visited twice weekly to collect all seeds produced. For each population, a random sample of offspring seeds were scored for genotype at the *W* locus to estimate the change in gene frequency after one generation. Locations of experimental populations were at open fields within the Duke Forest at Duke University, Durham, North Carolina. These sites were isolated from each other and from all other localities with *I. purpurea* present by at least 3 km, thus ensuring that gene flow into the sites was probably nonexistent. In addition, all nonexperimental *I. purpurea* were eradicated to prevent pollen contamination from native plants.

Experimental population 5 was established in early July 1989 by planting 1968 experimental seeds (656 *WW*, 640 *Ww*, and 672 *ww*). These seeds were generated by a crossing design similar to that reported by Rausher and Fry (1993), which randomizes the genetic background for genes unlinked to the *W* locus. This design involves three generations of crosses. In generation 1, six pairs of dark × white crosses were performed, with each pair constituting a "unit." In generation 2, one light (heterozygous) offspring from each unit was selfed to produce numerous offspring. In generation 3, self-sibs (offspring within the same unit) were paired in three ways (darks with darks, darks with whites, and whites with whites) to produce the experimental seed (dark, light, and white, respectively). Within each unit each cross type was replicated with three or four pairs, and within each pair reciprocal crosses were performed.

Experimental populations 1–4 were established in early July 1992 using seed produced in 1989 by population 5. To ensure that the frequency of the white allele was low in populations 1 and 2, these populations were established by planting 350, 40, and 20 seeds chosen randomly from those produced by dark, light, and white plants respectively in population 5. Similarly, to ensure that the frequency of the white allele was high in populations 3 and 4, these populations were established by planting 350, 40, and 20 seeds chosen randomly from those produced by white, light, and dark plants respectively in population 5. The initial gene frequencies in these samples were determined by planting a "companion" sample of seeds in the greenhouse at the same time and scoring the color of the flowers they produced, as described below.

Statistical Analyses

All statistical analyses of changes in gene frequencies are based on two sets of seed samples. One set, the greenhouse companion sample, consists of approximately 1500 seeds produced by the experimental plants in plot 5. In particular, random samples of 523, 435, and 534 seeds were selected from the seeds produced by darks, lights, and whites, respectively, in the Plot 5 experiment. These seeds were germinated in the greenhouse, allowed to grow to flowering, and their genotype determined by inspection. This sample of seeds was used for two purposes: (1) to estimate the final gene frequency in population 5 by comparing offspring gene frequency with the initial gene frequency in that population; and (2) to estimate the initial gene and genotype frequencies in plots 1–4 for comparison with final (offspring) gene frequencies in these plots. The second set of samples, the "offspring" samples, consisted of 545, 366, 680, and 444 randomly chosen seeds produced by the plants in plots 1–4 respectively. The seeds of these samples were also germinated and grown in the greenhouse to estimate final genotype and gene frequencies in each plot. Germination success of all samples was greater than 95%, thus providing estimates of genotype frequencies that are relatively free of potential bias introduced by natural selection operating in the greenhouse. In particular, this high germination rate makes it unlikely that any observed change in gene frequency in population 5 could be ascribed to differential seed mortality during the almost three years between collection and scoring the offspring seeds of that population.

The statistical significance of the change in gene frequency in each population was assessed using likelihood methods (Edwards 1972; Weir 1990). In each of the four populations with extreme allele frequencies (populations 1–4), the unconstrained joint likelihood of the observed genotype counts in the greenhouse samples and in a random sample of offspring seed produced by that population is given by

$$L = \left(\prod_{i,j} p_{ij}^{N_{ij}}\right)\left(\prod_{i} p_i^{N_i}\right),$$

where i and j can take on the values D, L, and W, corre-

sponding to dark, light and white genotypes; p_{ij} is the probability that a parent plant of genotype i in population 5 produced an offspring seed of genotype j; N_{ij} is the number of offspring of genotype j produced by a maternal parent of genotype i in the greenhouse sample; p_i is the probability that an offspring seed produced by the population is genotype i, and N_i is the number of offspring in a random sample of seeds produced by the population that has genotype i. The null hypothesis that the allele frequencies in the offspring seed produced by one of these populations is equal to the initial allele frequencies, that is, that there is no change in gene frequency from one generation to the next, is represented by the constraint

$$\left(\frac{K_D}{410}\right)\left(p_{DD} + \frac{1}{2}p_{DL}\right) + \left(\frac{K_L}{410}\right)\left(p_{LD} + \frac{1}{2}p_{LL}\right) + \left(\frac{K_W}{410}\right)\left(\frac{1}{2}p_{WL}\right) = p_D + \frac{1}{2}p_L,$$

where K_i is the number of offspring seed produced by genotype i planted (note that $\Sigma_i K_i = 410$, the total number of seeds planted in each experiment). The left side of this constraint represents the initial gene frequency as a weighted average of the gene frequencies in the three greenhouse samples derived from maternal parents of different W-locus genotype, were the weights are the fractions of planted seeds that are offspring from these genotypes. This null hypothesis is tested using the standard likelihood-ratio statistic $\Lambda = -2$ ln where L_0 is the maximum likelihood corresponding to L with the constraint incorporated and L_1 is the maximum likelihood corresponding to the unconstrained version of L. Under the null hypothesis of no change in gene frequency, the statistic Λ has an approximate χ^2 distribution with one degree of freedom.

A similar analysis was performed for population 5. In this case, the initial gene and genotype frequencies were known exactly. The likelihood function for the offspring seeds is

$$L = \left(\prod_{i,j} p_{ij}^{N_{ij}}\right),$$

and the constraint corresponding to no gene frequency change over one generation is given by

$$\left(\frac{S_D}{S_T}\right)\left(p_{DD} + \frac{1}{2}p_{DL}\right) + \left(\frac{S_L}{S_T}\right)\left(p_{LD} + \frac{1}{2}p_{LL}\right) + \left(\frac{S_W}{S_T}\right)\left(\frac{1}{2}p_{WL}\right) = \text{initial frequency of dark allele,}$$

where S_i [$i \in (D, L, W)$] is the total number of seeds produced by plants of genotype i and $S_T = \Sigma_i S_i$.

For all analyses, likelihoods were assessed by an iterative grid-search method. Likelihoods were first calculated for all combinations of the parameters p_{ij} at intervals of 0.1. (For example, for the unconstrained models, there are five free parameters, each with the eleven possible values 0.0, 0.1, . . . 0.9, 1.0. This yielded a grid of 11^5 points. The likelihood was evaluated at each of these points.) Inspection of the resulting likelihood surface indicated that in all cases the surface had a single, global maximum. Using the grid point with the maximum likelihood as an approximation of the global maximum, a second grid search was then performed. For this search, the grid was centered on the maximum of the first search and a new grid with intervals of 0.02 was searched for the grid point with the maximum likelihood. A third search was then performed using a grid with intervals of 0.002 for each parameter. The maximum likeihood found in the third search was taken as a good estimate of the true maximum because the difference in likelihood value between that grid point and adjacent points was always less than 2%.

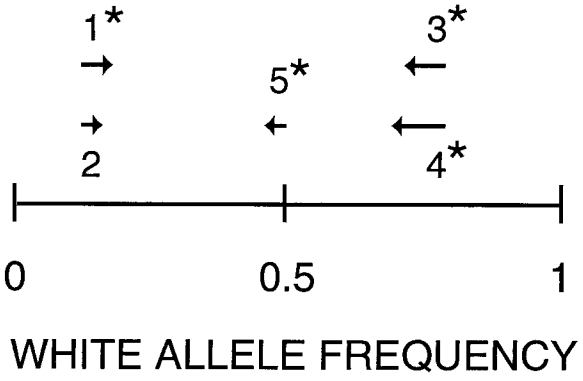

FIG. 1. Change in white allele frequency in experimental populations 1–5. Arrows indicate direction and magnitude of change in frequency. Base of arrow indicates initial allele frequency. Numbers correspond to populations. Asterisks indicate significant changes in gene frequency.

Viability and Seed Production in Population 5

In addition to estimating changes in gene frequency in population 5, we also compared the viability and fecundity of plants of the three genotypes to obtain information about the causes of any observed change in gene frequency. Viability, which was estimated as the proportion of plants surviving to produce at least one offspring seed, was compared among genotypes with a standard G-test (Sokal and Rohlf 1969). Total seed production was determined for each plant by collecting and counting seeds three times a week (Rausher and Fry 1993). Comparison of seed production among genotypes was performed using analysis of variance (ANOVA).

RESULTS

The pattern of gene frequency changes in the five populations are indicative of balancing selection operating at the W locus (Fig. 1). The three populations with initial frequencies of the white allele greater than normally observed (populations 3–5) all exhibited a decrease in the frequency of the white allele (Table 1). In contrast, the two populations (1 and 2) with initially low frequencies of the white allele each exhibited an increase in its frequency. In all populations except population 2, these changes were highly significant (Table. 1). Overall, the direction of gene frequency change in each of the five populations was as expected under the hypothesis that selection is balancing. The probability of this result under the null hypothesis of neutrality (equal likelihood of increase or decrease in frequency in a given experimental population) is $P = 0.034$. This pattern indicates that regardless of initial allele frequencies, selection causes frequencies

TABLE 1. Genotype and gene frequencies among seed samples used to estimate initial and final gene frequencies. Greenhouse samples were obtained from the offspring of population 5, whereas offspring samples were obtained from the offspring of populations 1–4 (see text). p_w is the frequency of the white allele in the sample. Δp_w is the estimated change in gene frequency in the indicated experimental population during the experiment. χ^2_1 and P are the chi-square value and associated probability corresponding to the likelihood-ratio test for change in gene frequency (see text). Genotype frequencies for population 5 were estimated indirectly rather than by direct count, but this does not affect the validity of the likelihood-ratio test.

Sample	Genotype			p_w	Δp_w	χ^2_1	P
	WW	*Ww*	*ww*				
Greenhouse: *WW* parents	0.87	0.13	—				
	456	67	—				
Greenhouse: *Ww* parents	0.37	0.37	0.26				
	160	162	113				
Greenhouse: *ww* parents	—	0.21	0.79				
	—	114	420				
Offspring: population 1	0.70	0.21	0.09	0.19	+0.06	11.37	0.001
	384	114	47				
Offspring: population 2	0.75	0.18	0.08	0.17	+0.04	1.75	0.02
	274	64	28				
Offspring: population 3	0.17	0.19	0.64	0.73	−0.08	20.70	0.001
	114	131	435				
Offspring: population 4	0.13	0.33	0.54	0.70	−0.10	24.09	0.001
	57	146	241				
Offspring: population 5	0.41	0.24	0.35	0.47	−0.04	441.1	0.0001
	523	435	534				

to converge to a region in which the frequency of the white allele is between approximately 0.25 and 0.45, which corresponds quite well to the range of frequencies commonly observed in natural populations.

In population 5, genotypes exhibited no detectably significant differences in either viability or seed production (Table 2). Moreover, the estimates of these fitness components for the two homozygote genotypes were remarkably similar. Whereas the viability estimate was slightly higher for the dark genotype, the opposite was the case for seed number. These differences compensated almost exactly to produce an estimate of expected seed production (viability × seed number) that differed by less than 0.05% (Table 2). Mean seed number exhibited a slight but significant degree of overdominance (Table 2).

DISCUSSION

The principle result of this study is that the *W* locus in *I. purpurea* appears to be subject to balancing selection, at least in the area in which this investigation was conducted. We infer balancing selection from the observation that both alleles seem to have increased in frequency when rare. Previous investigations have suggested the operation of a mechanism that likely causes the white allele to increase in frequency when rare. When this allele is at low frequency, white (*ww*) plants exhibit a higher selfing rate than darks or lights, apparently because bumblebees, the primary pollinator, visit whites less frequently (Brown and Clegg 1984; Fry and Rausher 1997). Coupled with an apparent lack of pollen discounting (Rausher et al. 1993) and minimal inbreeding depression (Pear 1983; S.-M. Chang and M. D. Rausher, unpubl. ms.), this increased selfing is expected to contribute to a net advantage to the white allele (Fisher 1941; Lloyd 1979; Wells 1979; Holsinger et al. 1984). The increase in the gene frequency of the w allele reported here for populations 1 and 2 lends credence to this explanation and provides the first direct evidence indicating that selection acts to prevent the elimination of the white allele.

When the white allele is not at low frequency, pollinators visit all color variants at the same rate, selfing rates are similar for the three genotypes (Epperson and Clegg 1987; Rausher et al. 1993; Fry and Rausher 1997), and there is thus not expected to be an analogous increase in the frequency of the white allele. If no other evolutionary forces were operating on the *W* locus, the white allele would be expected to eventually drift to fixation. Our observation that the frequency of the white allele decreases when initially at high and intermediate frequencies (populations 3–5) indicates that such fixation is not occurring and that therefore some other process must cause the dark allele to increase in frequency when rare.

Previous experiments have provided inconclusive evidence regarding the nature of this additional selective force. No dark-allele advantage was detected by Rausher and Fry (1993) in either survival or seed production. Overdominance in seed size does seem to occur (Rausher and Fry 1993; Mojonnier and Rausher 1997; Table 2), but its effects on

TABLE 2. Fitness components of *W*-locus genotypes in population 5. *P* is probability fitness component differs among genotypes. For viability, *P* was determined by a *G*-test with two degrees of freedom ($G = 1.02$). For ln seed number, *P* was determined by ANOVA ($F_{2,105} = 0.09$), which contained other sources of variation in addition to genotype. Numbers in parentheses are standard errors.

Fitness component	Genotype			P
	WW	*Ww*	*ww*	
Viability	0.83 (0.04)	0.73 (0.05)	0.79 (0.04)	> 0.90
ln seed number	3.34 (0.99)	3.35 (1.03)	3.39 (0.96)	> 0.87
Viability × seed number	23.42	20.81	23.43	
ln mean seed size	2.89 (0.21)	2.93 (0.19)	2.91 (0.18)	< 0.008

offspring fitness appear not to be of sufficient magnitude to maintain high frequencies of the dark allele (Mojonnier and Rausher 1997). In an experimental population, Fry and Rausher (1997) detected biased inheritance favoring the dark allele in pollen produced by heterozygotes, which could act to prevent the dark allele from being eliminated when rare. However, subsequent experiments (S. Paulsen and M. D. Rausher, in review) indicate that this mechanism does not operate in all, or even most, polymorphic *I. purpurea* populations.

Regardless of mechanism, our results provide direct evidence that the dark allele is actively maintained by selection. In addition, our estimates of viability and female fitness components in population 5 provide additional support for some of the conclusions derived from previous experiments. In particular, the lack of differences among genotypes in these fitness components confirms Rausher and Fry's (1993) results and means that a dark-allele advantage in these fitness components cannot explain the decrease in frequency of the white allele observed in this population. The increase in frequency of the dark allele in this population must therefore have been due to a fitness advantage of that allele either through male fitness or during the gametic stage, a conclusion that is consistent with Fry and Rausher's (1997) observation of non-Mendelian inheritance in pollen from heterozygotes. Our results thus reinforce the notion that in some polymorphic populations, but not all, some sort of biased inheritance may contribute to maintenance of the dark allele. As Fry and Rausher (1997) have shown, an interaction of this type of force with the frequency-dependent transmission advantage of whites associated with increased selfing is sufficient to maintain gene frequencies by balancing selection at roughly the values typically seen in nature.

Acknowledgments

This research was supported in part by NSF grant DEB 9318919 to MDR.

Literature Cited

Antonovics, J. 1994. The interplay of numerical and gene-frequency dynamics in host-pathogen systems. Pp. 129–145 *in* L. A. Real, ed. Ecological genetics. Princeton Univ. Press, Princeton, NJ.

Brown, B. A., and M. T. Clegg. 1984. The influence of flower color polymorphism on genetic transmission in a natural population of the common morning glory, *Ipomoea purpurea.* Evolution 38: 58–67.

Chang, S.-M., and M. D. Rausher. 2000. The role of inbreeding depression in maintaining the mixed mating system of the common morning glory, *Ipomoea purpurea.* Evolution 54:*in press.*

Charlesworth, D., and B. Charlesworth. 1979. The evolution and breakdown of S-allele systems. Heredity 43:41–55.

Clark, A. G. 1993. Evolutionary inferences from molecular characterization of self-incompatibility alleles. Pp. 79–108 *in* N. Takahata and A. G. Clark, eds. Mechanisms of molecular evolution. Sinauer, Sunderland, MA.

Eckert, C. G., D. Manicacci, and S. C. H. Barrett. 1996. Frequency-dependent selection on morph ratios in trystylous *Lythrum salicaria* (Lythraceae). Heredity 77:581–588.

Edwards, A. W. F. 1972. Likelihood. Cambridge Univ. Press, Cambridge.

Ennos, R. A. 1981. Quantitative studies of the mating system in two sympatric species of *Ipomoea.* Genetica 57:93–98.

———. 1983. Maintenance of genetic variation in plant populations. Pp. 129–155 *in* M. K. Hecht, B. Wallace, and G. T. Prance, eds. Evolutionary biology. Vol. 16. Plenum Press, New York.

Ennos, R. A., and M. T. Clegg. 1983. Flower color variation in the morning glory, *I. purpurea.* J. Hered. 74:247–250.

Epperson, B. K., and M. T. Clegg. 1986. Spatial auto-correlation analysis of flower color polymoprhisms within substructured populations of morning glory (*Ipomoea purpurea*). Am. Nat. 128: 840–858.

———. 1987. Frequency-dependent variation for outcrossing rate among flower-color morphs of *Ipomoea purpurea.* Evolution 41: 1302–1311.

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53–63.

Fry, J. D., and M. D. Rausher. 1997. Selection on a floral color polymorphism in the tall morning glory (*Ipomoea purpurea*): transmission success of the alleles through pollen. Evolution 51: 66–78.

Holsinger, K. E., M. W. Feldman, and F. B. Christiansen. 1984. The evolution of self-fertilization in plants: a population genetic model. Am. Nat. 124:446–453.

Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.

Lloyd, D. L. 1979. Some reproductive factors affecting the selection of self-fertilization in plants. Am. Nat. 113:67–79.

Mojonnier, L. E., and M. D. Rausher. 1997. Selection on a floral color polymorphism in the common morning glory (*Ipomoea purpurea*): the effects of overdominance in seed size. Evolution 51:608–614.

Pear, J. R. 1983. Viability and fecundity consequences of breeding systems in *Ipomoea purpurea* (L.). MSc. thesis, University of Georgia, Athens, GA.

Rausher, M. D., and J. D. Fry. 1993. Effects of a locus affecting floral pigmentation in *Ipomoea purpurea* on female fitness components. Genetics 134:1237–1247.

Rausher, M. D., and E. L. Simms. 1989. The evolution of resistance to herbivory in *Ipomoea purpurea.* I. Attempts to detect selection. Evolution 43:563–572.

Rausher, M. D., D. Augustine, and A. VanderKooi. 1993. Absence of pollen discounting in genotypes of *Ipomoea purpurea* exhibiting increased selfing. Evolution 47:1688–1695.

Richman, A. D., T.-H. Kao, S. W. Schaffer, and M. K. Uyenoyama. 1995. S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). Heredity 75:405–415.

Simms, E. L., and M. D. Rausher. 1989. The evolution of resistance to herbivory in *Ipomoea purpurea.* II. Natural selection by insects and costs of resistance. Evolution 43:573–585.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman, San Francisco, CA.

Uyenoyama, M. K., and B. O. Bengtsson. 1979. Towards a genetic theory for the evolution of the sex ratio. Genetics 93:721–736.

Vekemans, X., and M. Slatkin. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics 137: 1157–1165.

Weir, B. S. 1990. Genetic data analysis. Sinauer, Sunderland, MA.

Wells, H. 1979. Self-fertilization: advantageous or deleterious. Evolution 33:252–255.

Corresponding Editor: D. Piñero