Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 3983–3986, April 1995
Evolution

Founder-effect speciation theory: Failure of experimental corroboration

(*Drosophila pseudoobscura*/evolution of sexual isolation/founder principle/genetic drift)

ANDRÉS MOYA*†, AGUSTÍ GALIANA*, AND FRANCISCO J. AYALA†‡

*Departmento de Genètica, Universitat de València, 46100 Burjasot, Spain; and †Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717

Contributed by Francisco J. Ayala, January 18, 1995

ABSTRACT **The theory of founder-effect speciation proposes that colonization by very few individuals of an empty habitat favors rapid genetic changes and the evolution of a new species. We report here the results obtained in a 10-year-long and large-scale experiment with *Drosophila pseudoobscura* designed to test the theory. In our experimental protocol, populations are established with variable numbers of very few individuals and allowed to expand greatly for several generations until conditions of severe competition for resources are reached and the population crashes. A few random survivors are then taken to start a new population expansion and thus initiate a new cycle of founding events, population flushes, and crashes. Our results provide no support for the theories proposing that new species are very likely to appear as by-products of founder events.**

Mayr (1, 2) has formulated a speciation model according to which the probability of speciation is enhanced when a new population is started by a few migrant individuals colonizing an available habitat. According to this model, a genetic restructuring of the population, often yielding a new species, is likely to ensue as the population adapts to the new habitat under the conditions of genetic depauperation caused by the founder event. A number of related models have thereafter been advanced (3–10) that largely share in common certain demographic and genetic features. The demographic components include colonization by (or constriction to) very few individuals, followed by rapid population expansion until severe competition for resources is reached. The genetic components include initial loss of variability, inbreeding, genetic drift, recombination, and eventual epistatic restructuring of the genome.

Carson (4) sketched a protocol for testing his founder–flush–crash version of the theory, which stimulated a number of experimental tests yielding inconsistent results (11–21). In these experiments, populations are established with two or very few individuals and allowed to expand greatly for several generations (the flush phase) until conditions of severe competition for resources are reached (the crash phase). In some experiments, a few random survivors are then taken to start a new population expansion and thus initiate a cycle of founding events, population flushes, and crashes. Several of these cycles may be carried out, as it is thought that every founding event offers an additional chance for the desired founder effect. Experimental populations are monitored for traits relevant to speciation, mainly ethological isolation assayed by mating tests.

We report the results of a large-scale experiment that follows the founder–flush–crash protocol, with variable numbers of founders (n = 1–9 pairs). The experiment involves 59 populations, 13 founder–flush–crash cycles extending over 10 years and 100 generations, and hundreds of mating tests. Our results fail to support the theory in that no persistent reproductive isolation has evolved between any two populations.

MATERIALS AND METHODS

Two ancestral populations of *Drosophila pseudoobscura*—BCA from Bryce Canyon National Park, Utah, and MA from Lake Zirahuén, Mexico—were used; these populations were derived from many individuals collected in the wild (21). In December 1984, 45 experimental populations were established, 27 derived from BCA and 18 from MA, as well as 14 control populations, half from each locality. Each experimental population was founded with n virgin pairs with n = 1, 3, 5, 7, or 9, allowed to grow exponentially for about six generations, followed by one crash generation with drastic competition for limiting food and space resources. A new founder–flush–crash cycle was then started with the same initial number of founders, for a total of 13 cycles and about 100 generations. The control populations were of three types: two ancestral (C.A) that were maintained in large mass cultures, without subjecting them to bottlenecks or flush–crash cycles; six inbred (C.I) subject to inbreeding but not flush–crash cycles; and six partial (C.P) whose cycles consisted of three consecutive generations of one-pair inbreeding followed by a flush two generations shorter than in the experimental populations (see ref. 21 for additional experimental details).

Evolution of reproductive (sexual) isolation in the populations was tested by assortative mating tests: 12 males and 12 females from each of two populations were confined in mating chambers where the types of matings are directly observed. (Several of these individual tests, typically four, were accumulated for each two-population combination in every analyzed cycle.) Assortative mating was measured by χ^2 homogeneity tests and the Y index (15, 21), which is positive when matings occur preferentially between members of the same population and is not affected by differences in mating propensity.

The mating tests were performed at various intervals, starting with the fourth cycle, when at least one population was tested for each n value and geographic origin. In cycle 5, all populations with n = 1 or 3 were tested; in cycle 7 most tests involved populations with n = 5, 7, or 9, although a few combinations (involving populations BC2, M3, BC7, and M7) were retested because of the previous appearance of incipient ethological isolation. Finally, in cycles 11, 12, and 13, we sought primarily to retest all population pairs that had earlier exhibited significant assortative mating, to ascertain whether the isolation had persisted rather than being accidental or ephemeral. Thus, the set of population combinations tested during the three last cycles (and partially also in cycle 7) is not a random sample with respect to the probability of reproductive isolation but biased toward combinations exhibiting assortative mating. The total number of pairwise combinations of popu-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡To whom reprint requests should be addressed.

lations tested (counting separately a given combination when analyzed in different cycles) is 678.

Some features of our experimental design deserve attention. *D. pseudoobscura*, a species endemic through much of western North America, was chosen as a model system, rather than a cosmopolitan species, such as *Drosophila melanogaster* or the common housefly *Musca domestica*, which might be less prone to genetic changes relevant to speciation during colonization (6). Moreover, we sought to meet the postulates of the founder-effect theory of speciation by incorporating certain features in our experimental design. First, we chose single local populations as our ancestral populations, rather than the hybrid populations used in some previous experiments (11, 12). Second, we tested two natural populations, one (MA) was initially polymorphic for abundant rearrangements in the third chromosome, whereas the other (BCA) was monomorphic, a feature that permits testing the effect of recombination on the evolution of reproductive isolation (recombination is thought, in some models, to be essential for the postulated epistatic restructuring of the genome). Third, populations were reared in culture bottles instead of large cages, to control larval competition during the flush phase (10 pairs of adult flies were allowed to lay eggs for 1 week in each of the 10 bottles making up each population). Flies were transferred each generation to new cultures by means of a particular protocol that assures gene flow among all cultures making up each particular population (21). Fourth, the number of experimental populations was large, much larger than in any previous experiment. Finally, we used a range of initial conditions for the founder event, allowing for different levels of genetic variability, inbreeding, and genetic drift.

RESULTS

The results are summarized in Table 1. Out of 486 tests between two experimental populations (E × E), 49 showed significant positive assortative mating and 6 showed significant negative assortative mating at the 5% level, using both standard χ^2 homogeneity tests or the Y index. The difference between the number of significant positive and negative cases is significant ($\chi^2 = 33.6$, with 1 degree of freedom; $P < 0.001$). Assortative mating between the experimental (E) and control populations (C.A) was examined in 142 pairwise tests, of which 10 yielded significantly positive and 6 yielded significantly negative assortative mating, a difference that is not statistically significant. Of 50 pairwise tests between control populations, only 1 was significantly positive and 2 were significantly negative.

We have tested the statistical validity of these results by means of a sequential Bonferroni test (22), which corrects for statistical significance arising by chance when many instances are tested. Three combinations between experimental populations [M19 ($n = 7$) × BC7 ($n = 3$) at cycle 4, M7 ($n = 3$) × BC10 ($n = 3$) at cycle 5, and BC10 ($n = 3$) × M3 ($n = 1$) at cycle 11] plus one between an experimental and a control population [BC18 ($n = 5$) × C.A (BCA) at cycle 13] are statistically significant by this test, all with positive assortative mating.

To ascertain whether instances of positive assortative mating reflect stable reproductive isolation between populations, we tested in successive cycles all population combinations that had yielded significant results at least once. The results are summarized in Table 2. The number of repeated tests for a particular population combination is a function of when positive assortative mating was first observed. Thus, combinations that yielded significant assortative mating in the first set of tests (i.e., after the fourth population cycle) could be tested again in the ensuing five sets of tests, whereas other combinations could only be retested four or fewer times. In particular, sexual isolation tests were performed after each of the last three cycles (cycles 11–13) for the purpose of including among the combinations tested all those that had yielded assortative mating in some previous test. Of 51 population combinations tested more than once, only 11 exhibited assortative mating the second time, 1 the third time, and none four or more times. It is clear that reproductive isolation is not stable, and, of course, there is no trend toward increasing isolation as time proceeds.

Table 3 summarizes the results of the mating tests for 8 experimental populations (out of the 45 total) that are highly represented among the pairs yielding significant assortative

Table 1. Number of significant mating tests between populations of *D. pseudoobscura* after various founder–flush–crash cycles

Population cycles completed	Test dates	Population combination	Combinations tested	Significant ($P < 0.05$)	
				Positive	Negative
4	Oct.–Dec. 1986	E × E	45	7	0
		E × C.A	20	0	1
5	Jan.–Sept. 1988	E × E	171	19	3
		E × C.A	38	0	2
7	Jan. 1989–May 1990	E × E	154	10	1
		E × C.A	60	8	2
11	Oct. 1990–Mar. 1991	E × E	66	7	1
		E × C.A	13	0	0
12	June–Nov. 1991	E × E	38	6	0
		E × C.A	8	1	1
13	Dec. 1991–May 1992	E × E	12	0	1
		E × C.A	3	1	0
	Oct. 1986–May 1992	C.A × C.A	6	0	0
	July 1988–Nov. 1989	C.A × C.P	12	0	2
	Mar.–June 1990	C.A × C.I	6	0	0
	June–Nov. 1991	C.I × C.I	5	0	0
	June–Nov. 1991	C.P × C.P	21	1	0

E represents the 45 experimental populations; C.A are two control populations not subject to any experimental treatment and maintained by mass culture; C.I are six control populations that were each subject to inbreeding for five generations, by one-pair brother × sister mating, and maintained thereafter by mass culture; C.P are six populations subject in each cycle to three generations of one-pair brother × sister mating and then to flush–crash for two fewer generations than the E populations. The large majority (595) of test combinations were replicated four times; others (83) were replicated between 1 and 9 times (4 times on average). Deviations from random mating are appraised by the Y statistic (15, 21).

Table 2. Number of two-population combinations yielding significant prezygotic isolation in more than one test

Number of combinations retested	Number of tests					
	1	2	3	4	5	6
25	22	3				
19	13	6	0			
3	3	0	0	0		
2	0	2	0	0	0	
2	1	0	1	0	0	0

The first row indicates that 25 combinations were tested two times, of which 22 exhibited significant assortative mating only in one test and 3 in two tests; other rows should be interpreted similarly.

mating. Indeed, these 8 populations are involved in 46 of the 60 combinations yielding significant assortative mating (out of 678 combinations tested); in 12 of the significant combinations, both populations come from this subset of 8 populations. This table gives a Y value calculated by accumulating all mating tests that involve the particular population for a given cycle. An estimate of the coefficient of inbreeding (F) has also been calculated on the assumption that the effective size of the populations during the flush phase is $N = 100$ (for $n = 1$) or 150 (for $n > 1$). There is no evidence that positive assortative mating is increasing as time proceeds. The proportion of significant Y values among the eight populations is 100%, 75%, 45%, 60%, 38%, and 29% for cycles 4, 5, 7, 11, 12, and 13, respectively. The lack of increase in assortative mating is not attributable to the increase in inbreeding that takes place as time proceeds: the correlation between Y and F is not statistically significant ($r^2 = 0.155$; $P > 0.10$; 32 degrees of freedom). More generally, there is no correlation between level of inbreeding and degree of assortative mating over the whole data set.

Table 3. Populations that yield most of the positive significant tests

Population	n	Cycle	Combinations tested		F	Y
			N	Significant		
M3	1	4	11	4	0.63	0.162***
		5	20	5	0.71	0.122***
		7	5	1	0.82	0.113*
		11	12	2	0.93	0.089***
		12	10	0	0.95	0.042
		13	2	0	0.96	0.112
M5	1	5	20	3	0.71	0.069**
		12	2	1	0.95	0.209***
		13	2†	—	0.96	—
M7	3	4	11	3	0.29	0.120***
		5	20	2	0.35	0.047
		7	5	1	0.45	0.070
		11	12	2	0.61	0.035
		12	7	0	0.64	0.024
		13	3	0	0.67	0.147
M14	5	7	17	4	0.31	0.107***
		11	12	4	0.44	0.081***
		12	8	1	0.47	0.085
		13	1	0	0.49	0.139
BC7	3	4	11	2	0.29	0.076*
		5	20	5	0.35	0.085***
		7	5	1	0.45	0.032
		11	12	2	0.61	0.073**
		12	8	5	0.64	0.170***
		13	5	0	0.67	0.140**
BC18	5	7	9	1	0.31	0.004
		12	1	0	0.47	−0.072
		13	1	1	0.49	0.361***
BC21	7	7	10	2	0.23	0.052
		12	3	2	0.37	0.239**
		13	2	0	0.39	0.059
BC22	7	7	10	4	0.23	0.173***
		11	13	0	0.34	0.042
		12	4	0	0.37	0.039
		13	1	0	0.39	0.000

*, $P < 0.1$; **, $P < 0.05$; ***, $P < 0.001$.
†Two tests involving M5 were attempted in cycle 13, but not enough virgins could be collected for the tests owing to the unusual lack of vigor of this population.

We now raise the question of whether the high incidence of assortative mating among the 8 populations in Table 3 is due to their behavioral differentiation or simply reflects a bias introduced by the sampling procedure. These populations were tested more often than the rest because they exhibited significant assortative mating when they were first tested. The 8 populations are represented 58 times (1 population in each of 34 tests, 2 in each of 12 tests) in significant tests out of the 248 times they were tested, an incidence of 23.4%. The other 51 populations (37 experimental and 14 controls) are represented 62 times in significant tests out of 1108 times tested, an incidence of 5.6%. If we now remove the first time a given population pair exhibited significant assortative mating (since the pair was selected for additional testing precisely because it had already exhibited assortative mating), the 8 populations are represented 16 (7.8%) out of 206 times in significant tests. The incidence of significant assortative mating is not significantly different between the 8 populations in Table 3 and the other 51 ($P > 0.10$, using a G test with Williams correction; ref. 23, page 704). If we remove from consideration tests that involve control populations, the 8 populations are represented 13 out of 187 times (7.0%) in significant tests, whereas the other 37 experimental populations are represented 51 out of 885 times (5.8%) in significant tests. This difference is not statistically significant either. Thus, there is no evidence that the 8 populations in Table 3 have evolved any more reproductive isolation than the others.

DISCUSSION

Founder-effect speciation theory predicts that the narrowest bottlenecks ($n = 1$ or $n = 3$) are more likely to lead to speciation. If we examine pairwise combinations between experimental populations of the same n value, the fraction of positive significant tests is 10.7% (3 of $N = 28$ tests) for $n = 1$, 13.8% for $n = 3$ ($N = 87$), 14.3% for $n = 5$ ($N = 35$), 7.1% for $n = 7$ ($N = 42$), and 0% for $n = 9$ ($N = 38$). In tests between experimental and control (C.A) populations, no combination is significant for $n = 1$ ($N = 24$), and the largest value is for $n = 7$ (18.7%, $N = 32$). In general, there is no statistical association between the bottleneck size parameter and the probability of positive assortative mating. The regression equation is $Y = 0.043 + 0.003n$ (standard error for the slope is 0.004; $P > 0.10$; 118 degrees of freedom) for the complete data set.

If we consider only the $n = 1$ populations, which is the founder size used in other experiments, our results are consistent with those of Ringo *et al.* (13–15), who obtained 5% ($N = 168$) and 8.3% ($N = 48$) significant cases of assortative mating with *Drosophila simulans*, for tests between experimental populations and between them and the controls, respectively, but not with those of Powell (11) and Dodd and Powell (12), who obtained 35% ($N = 54$) and 31% ($N = 26$) cases of assortative mating for similar tests with *D. pseudoobscura*. An experiment with the common house fly *M. domestica* with $n = 1$ failed to detect the evolution of any significant reproductive isolation (17).

Lack of consistency with the founder-effect theory also occurs with respect to the opportunity of genetic recombina-

tion associated with the geographic origin of the populations. The theory (in general, but particularly in the version of refs. 6 and 9) predicates speciation on the opportunity for ample genetic recombination. Thus, the BCA populations should be favored, since the MA populations are, initially at least, extremely polymorphic for chromosomal inversions. Our results do not favor the theory in this respect. Of the eight populations that represent most of the cases of significant assortative mating (Table 3), half are BCA and half are MA, even though we had 50% more BCA than MA populations (27 vs. 18); and the cumulative Y data shown in Table 3 yield nine significant instances for MA and eight for BCA populations. If we calculate the cumulative Y value for each of the 45 experimental populations in each cycle in which they were tested, we find for BCA populations 20 instances in which positive assortative mating occurs and 50 in which it does not; for MA populations, the numbers are 16 and 32, respectively. The difference is not significant ($P > 0.10$; G test with Williams correction), but there is in any case a greater incidence of positive assortative mating in the MA than in the BCA populations.

We may want to consider only those populations with n = 1 or 3, the ones most likely to yield reproductive isolation (6). The last set of tests (after cycle 13) included 9 MA populations, 3 of which yielded significantly positive assortative mating, but 10 BCA populations, only 1 of which exhibited significantly positive assortative mating. The data, once again, fail to support the proposal that open recombination facilitates ethological divergence. If we examine all those populations that yielded positive assortative mating in any cycle, we find that, among the populations with n = 1 or 3, this occurred in 6 of the 9 MA populations but in 8 of the 10 BCA populations, a difference that is not statistically significant ($G = 0.390$ with Williams correction; 1 degree of freedom; $P > 0.10$).

We conclude that our experiment provides no empirical support for the proposition that genetic and demographic processes associated with the founding of a population by very few individuals are very likely to cause speciation. Moreover, forceful theoretical arguments have been raised against the founder-effect theory of speciation (22, 24–30). It does not, of course, follow that new species never arise from populations established by one or very few founders. Indeed, we conjecture that, on occasion, particularly in remote islands or other isolated habitats, new species have evolved from one or few founders since remote habitats are likely to have been colonized rarely and by few individuals. But in the present case, as in all other evolutionary questions involving natural history, what is at issue is the frequency of a process or event, not whether it ever occurs at all. The theory of founder-effect speciation proposes that founder events, followed by generations of population expansion under relaxed competition until eventual saturation of resources and concomitant genetic recombination and natural selection, are likely to result in speciation. The domain of that likelihood is severely delimited by the magnitude of our 10-year experiment, which has involved 59 populations (51 of which have experienced 13 founder–flush–crash cycles) that have been tested in 678 two-population combinations by means of about 2780 separate assortative mating experiments (in each of which participated 48 individuals collected as virgins, for a total of 53,348 matings directly observed).

We thank Ananias Escalante, Steven A. Frank, and Laurence D. Mueller for reading the manuscript and offering suggestions for improvement. This work was supported by a fellowship from Generalitat Valenciana (Spain) to A.G., Grant PB93-0690 from the Direccion General de Ciencia y Tecnologia (Spain) to A.M., and Grant GM42397 from the National Institutes of Health to F.J.A.

1. Mayr, E. (1954) in *Evolution as a Process*, eds. Huxley, J., Hardy, A. C. & Ford, E. B. (Allen & Unwin, London), pp. 157–180.
2. Mayr, E. (1963) *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, MA).
3. Carson, H. L. (1968) in *Population Biology and Evolution*, ed. Lewontin, R. C. (Syracuse Univ. Press, Syracuse, NY), pp. 123–137.
4. Carson, H. L. (1971) *Stadler Genet. Symp. Ser.* **3,** 51–70.
5. Carson, H. L. (1975) *Am. Nat.* **109,** 83–92.
6. Templeton, A. R. (1980) *Genetics* **94,** 1011–1038.
7. Carson, H. L. & Templeton, A. R. (1984) *Annu. Rev. Ecol. Syst.* **15,** 97–131.
8. Carson, H. L. (1986) in *Evolutionary Processes and Theory*, eds. Karlin, S. & Nevo, E. (Academic, New York), pp. 391–409.
9. Templeton, A. R. (1989) in *Genetics, Speciation and the Founder Principle*, eds. Giddings, L. V., Kaneshiro, K. Y. & Anderson, W. W. (Oxford Univ. Press, Oxford), pp. 329–344.
10. Carson, H. L. (1989) in *Genetics, Speciation and the Founder Principle*, eds. Giddings, L. V., Kaneshiro, K. Y. & Anderson, W. W. (Oxford Univ. Press, Oxford), pp. 345–362.
11. Powell, J. R. (1978) *Evolution* **32,** 465–474.
12. Dodd, D. M. B. & Powell, J. R. (1985) *Evolution* **39,** 1388–1392.
13. Ringo, J. M., Wood, D., Rockwell, R. & Dowse, H. (1985) *Am. Nat.* **126,** 642–661.
14. Ringo, J. M., Barton, K. M. & Dowse, H. B. (1986) *Behavior* **97,** 226–233.
15. Ringo, J. M. (1987) in *Evolutionary Genetics of Invertebrate Behavior*, ed. Huettel, M. (Plenum, New York), pp. 79–88.
16. Ringo, J. M., Dowse, H. B. & Lagasse, S. (1987) *J. Hered.* **78,** 271–272.
17. Meffert, L. M. & Bryant, E. H. (1991) *Evolution* **45,** 293–306.
18. Meffert, L. M. & Bryant, E. H. (1992) *Evolution* **46,** 1399–1407.
19. Bryant, E. H. & Meffert, L. M. (1993) *Heredity* **70,** 122–129.
20. Galiana, A., Ayala, F. J. & Moya, A. (1989) in *Evolutionary Biology of Transient Unstable Populations*, ed. Fontdevila, A. (Springer, Berlin), pp. 58–73.
21. Galiana, A., Moya, A. & Ayala, F. J. (1993) *Evolution* **47,** 432–444.
22. Rice, W. R. & Hostert, E. E. (1993) *Evolution* **47,** 1637–1653.
23. Sokal, R. R. & Rohlf, F. J. (1981) *Biometry* (Freeman, San Francisco), 2nd Ed, p. 704.
24. Barton, N. H. & Charlesworth, B. (1984) *Annu. Rev. Ecol. Syst.* **15,** 133–164.
25. Rouhani, S. & Barton, N. H. (1987) *J. Theor. Biol.* **126,** 51–62.
26. Charlesworth, B. & Rouhani, S. (1988) *Evolution* **42,** 1129–1145.
27. Barton, N. H. (1989) in *Speciation and Its Consequences*, eds. Otte, D. & Endler, J. (Sinauer, Sunderland, MA), pp. 229–256.
28. Coyne, J. A. (1990) *Nature (London)* **344,** 30.
29. Charlesworth, B. (1991) *J. Evol. Biol.* **4,** 162–163.
30. Coyne, J. A. (1994) *Evolution* **48,** 19–30.