Genetic analysis of speciation by means of introgression into *Drosophila melanogaster*

Kyoichi Sawamura*†, Andrew W. Davis‡, and Chung-I Wu*

*Department of Ecology and Evolution, University of Chicago, 1101 East 57th Street, Chicago, IL 60637; and ‡Science Department, Webster University, 470 East Lockwood Avenue, St. Louis, MO 63119

Communicated by Jiazhen Tan, Fudan University, Shanghai, People's Republic of China, December 20, 1999 (received for review July 1, 1999)

In the last decade, the genetic basis of reproductive isolation has been shown to be surprisingly polygenic, and yet even the most efficient system currently in use could lend itself to molecular analysis only in highly selected cases. By extending the recent discovery of fertility rescue between *Drosophila melanogaster* and *Drosophila simulans*, we show that this hybridization can permit systematic and precise delineation of the genetic and molecular basis of speciation. In a region of 5% of the *D. simulans* genome introgressed into *D. melanogaster*, we discover at least six genes of hybrid male sterility and none for female sterility by deficiency mapping. A single case of hybrid inviability has been tracked down to a 3-Kb element that was inserted into the *Cyclin E* locus during species hybridization. The extent of interspecific genetic divergence underlying hybrid male sterility, especially in contrast with the low degree of inviability and female sterility, is far greater than expected from previous studies.

One of the important subjects in speciation study is the genetic mechanism of reproductive isolation (1, 2). *Drosophila melanogaster*, with its wealth of resources, could allow the delineation of genes of reproductive isolation by genetic, cytological, and molecular means. Unfortunately, this species does not produce progeny beyond the F1 generation when crossed to any of its known sibling species, including *Drosophila simulans* (3–5). Since the time A. H. Sturtevant observed this sterility 80 years ago (3), evolutionary geneticists have relied on other species pairs. Th. Dobzhansky (6) was the first to do so, and many have since followed. Nevertheless, the progress has been slow (7), and not until recently has a gene of reproductive isolation been cloned from *Drosophila* (8). To move the field at a faster pace, a more efficient system involving *D. melanogaster* is highly desirable.

Various attempts at bringing this species into speciation research have been made, including mating triploid *D. melanogaster* females to heavily irradiated *D. simulans* males to obtain the equivalent of F2s (9, 10), searching for sibling species that might be crosshybridizable (4, 5), examining far-flung populations that might be reproductively isolated (11, 12), and surveying for hybrid inviability rescue mutations (13–16). None has succeeded in making this "model organism" usable material for systematic studies of the genetics of speciation. The recent discovery of combinations of lines from *D. melanogaster* and *D. simulans* that yield fertile hybrid females has perhaps given us the best chance to tap into the resources of *D. melanogaster* (17).

Materials and Methods

Nomenclature. We shall briefly explain the conventional *Drosophila* nomenclature used in this study. For example, *Df(2L) J39* (31C-D; 32D-E) denotes a deletion on the left arm of the second chromosome with the two breakpoints determined to be at the locations of 31C (or 31D) and 32D (or 32E) of the polytene chromosome, respectively. This deletion is named J39. Following this convention, we use *Int* to denote introgressions. We name the two introgressed segments shown in the present paper *Int(2L)D* (21A1; 22D1–23A2) and *Int(2L)S* (30F1–31E7; 35D7–36A14) with the general format *Int(chromosome arm)name* (left breakpoint; right breakpoint). *D* and *S* are named after Dobzhansky and Sturtevant, respectively. *Int(2L)D+S* denotes a chromosome that carries the introgressions of both regions. Second chromosome balancers (multiply marked and serially inverted chromosomes that prevent recombination) have also been indispensable for this study. A most common balancer chromosome called *CyO* and several others (referred to simply as *Bal*) were used.

Construction of Introgressions. We obtained fertile F1 hybrid females between females from the C167.4 strain of *D. simulans* and males from the *C(1)M4, y^2/In(1)AB, f* strain of *D. melanogaster* (ref. 17; the exact meaning of this nomenclature is immaterial here) and backcrossed them to the Oregon-R strain of *D. melanogaster*. The resulting F2 females were again backcrossed to Oregon-R males individually. Although more than 500 F2 females tested were sterile, 21 females did produce progeny. A founder F2 female with a notched wing was selected for further crosses. This dominant phenotype is linked to the second chromosome bearing the introgressions. F4 and F5 females were individually crossed to *CyO/bw^{VDe2}* males to establish the introgression over the *CyO* balancer chromosome.

Deficiency Mapping. To test the viability and fertility of flies heterozygous for the *Int(2L)D+S* introgression and the *Df(2L)* deficiency, *Int(2L)D+S/CyO* females were crossed to *Df(2L)/Bal* males (where *Bal* denotes a second chromosome balancer). Five *Int(2L)D+S/Df(2L)* flies were mated to five flies of the opposite sex of the Oregon-R line (at least four replicates). If no offspring appeared, they were regarded as sterile. If fewer than one offspring per tested parent appeared, they were regarded as semisterile. The names of the *Df(2L)* deficiencies used and their breakpoints (18) are given below: 1, *net-PMF* (21A1; 21B7–8); 2, *al* (21B8-C1; 21C8-D1); 3, *ast2* (21D1–2; 22B2–3); 4, *dp-79b* (22A2–3; 22D5-E1); 5, *Mdh* (30D-F; 31F); 6, *J39* (31C-D; 32D-E); 7, *Prl* (32F1–3; 33F1–2); 8, *prd1.7* (33B2–3; 34A1–2); 9, *b87e25* (34B12-C1; 35B10-C1); 10, *osp29* (35B1–3; 35E6); 11: *r10* (35E1–2; 36A6–7); 12, *cact-255rv64* (35F-36A; 36D); 13, *H20* (36A8–9; 36E1–2); 14, *TW137* (36C2–4; 37B9-C1). All deficiencies, except no. 5 and no. 6, are normal as heterozygotes in *D. melanogaster*. These two deficiencies cause female sterility and reduced viability, respectively, as heterozygotes; in other words, they are haplo insufficient for either trait. Because of the low viability, only five females and seven males were tested for deficiency no. 6. For control, we have used 17 deficiencies that map to between *Int(2L)D* and *Int(2L)S* but do not overlap with either introgression. None of them uncovered any hybrid inviability or sterility, as expected.

†To whom reprint requests should be addressed at present address: Drosophila Genetic Resource Center, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, 606-8585 Japan. E-mail: sawamura@drochan.bio.kit.ac.jp.

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.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.050558597. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.050558597

DNA Marker-Assisted Mapping. We use 14 DNA markers for genotyping to determine the extent of introgression. Eleven of these markers were amplified from genomic DNA by PCR and detected as restriction fragment length polymorphism (five markers), single-strand conformation polymorphism (two markers), or microsatellite-associated length difference (four markers). Genomic Southern hybridization was performed for three markers. The primer sequences and the conditions of PCR, gel electrophoresis, or Southern hybridization are available from the corresponding author on request.

Results and Discussion

In this study, we first obtained fertile F1 hybrid females between the appropriate strains of these two species (see *Materials and Methods*) and backcrossed them to males of the Oregon-R strain of *D. melanogaster*. The resultant F2 females were again backcrossed to Oregon-R males individually. This step has been a difficult hurdle in several unpublished attempts at introgression because of the classical "F2 breakdown" (1). The F2 genotypes are highly heterogeneous and most of them are inviable or sterile, presumably because of complex genic interactions. A substantial fraction of the F2 genotypes are either the reconstituted pure species genotype or something very close to it. Because these unwanted pure types are the most viable and fertile kinds, the task is to filter them out to retain the true F2 introgression types that are viable and fertile.

We noticed that many F1 and F2 hybrids are associated with various phenotypic aberrations in wing and eye morphology, bristle number, cuticular pattern, and so on. For example, a notched-wing phenotype could be observed in the F1 hybrid males from the cross between *D. melanogaster* females and *D. simulans* males (Fig. 1 *e* and *f*). [Because these males are generally inviable, the phenotype can be revealed only when inviability rescue mutations are also present in the crosses (13, 14)]. The molecular basis of morphological variations within and between species has attracted great attention recently (19–21). In this study, we use these morphological anomalies also as indicator of introgressions. A single F2 female with a notched-wing phenotype would eventually become the founder of subsequent introgression lines (see *Materials and Methods*). Because this notched-wing phenotype is retained by these introgression lines as shown in Fig. 1, it is a natural marker as well as good material for probing morphological divergence in the future.

Through a series of further crosses (see *Materials and Methods*), we obtained six sublines carrying introgressions on *2L* (left arm of the second chromosome). Each subline is homogeneous for the introgression and is maintained over the second chromosome balancer, *CyO*, to prevent further recombination. One of the sublines is homozygote inviable, and the others are homozygote viable but female and male sterile. This male sterility is because of the failure to produce mature sperm, but females with introgressions do lay many normal-looking eggs, unlike F1 females, which have atrophied ovaries. The developmental details of the female and male sterility will be described elsewhere (K.S., M.-T. Yamamoto, C.-I.W., and T. L. Karr, unpublished work).

We first determined the extent of the introgression by both molecular and cytological means. Given the history of unconfirmed reports of introgressions between *D. melanogaster* and its sibling species (22), we present the verification of introgressions in Fig. 2. The complete mapping results are summarized in Fig. 3*a*. The introgression in fact has two components on *2L*: one between cytological locations 21A1 and 22D1–23A2 and the other between 30F1–31E7 and 35D7–36A14. We propose to name the two introgressed segments *Int(2L)D* and *Int(2L)S*, respectively, where *Int* stands for introgression (see *Materials and Methods*). Four sublines, including the one that is homozygote inviable, have both the *D* and *S* components, whereas two others have only the proximal *S* introgression. Further molecular analysis showed that all three crossovers demarcating the introgressed segments must have occurred in F1, in the germline of the mother of the founder F2 female.

Because the introgression sublines are homozygote sterile in both males and females (hence the sterility alleles behaving as recessives), we could map the sterility loci by using a series of

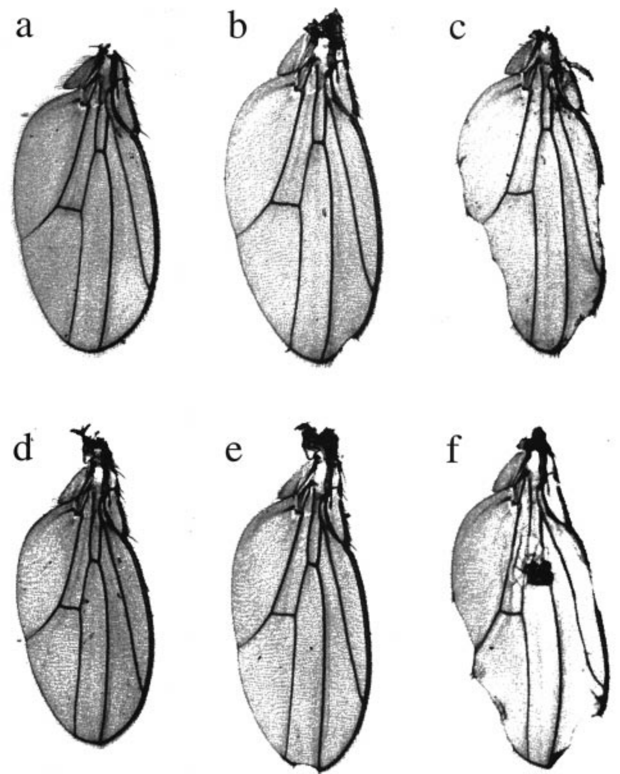

Fig. 1. Wing morphology associated with introgressions and F1 hybrids. Wings were taken from males of: (*a*) *D. melanogaster* Oregon-R; (*b*) introgression heterozygotes; (*c*) introgression homozygotes; (*d*) *D. simulans* C167.4; (*e* and *f*) F1 hybrids rescued by the lethal hybrid rescue mutations, *Lhr* and *Hmr* (13, 14), respectively.

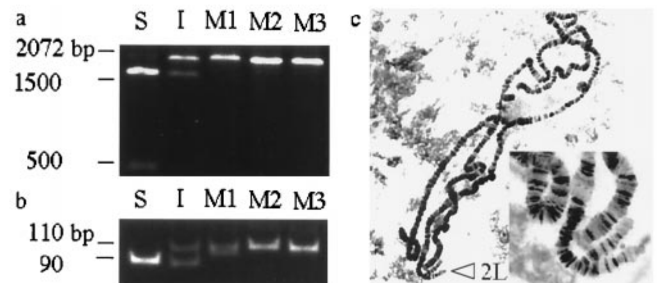

Fig. 2. Physical evidence for *D. simulans* introgressions on *2L*. (*a*) The *Pst*I-digested PCR products at the *Adh* locus at the cytological position 35B2: Lane S, the C167.4 strain of *D. simulans*; lane I, an introgression subline (*Int(2L)D+S/CyO*); lanes M1-M3, *D. melanogaster* lines used in the crosses (M1, *CyO/bw*^{VDe2}; M2, Oregon-R; M3, *C(1)M4, y*^{2}*/In(1)AB, f*). Lane I clearly exhibits the heterospecific pattern. All other five introgression sublines show a pattern identical with that of lane I. (*b*) The PCR products at the *ex* locus at the tip of *2L* (21C3–4). DNA sources for each lane are identical with *a*. In total, four sublines carry this distal introgression. (*c*) Polytene chromosomes of the Oregon-R/*Int(2L)D+S* heterozygote. Note the asynapsis at the tip of *2L*, which is always associated with heterospecific chromatids between *D. melanogaster* and *D. simulans* (35).

GENETICS

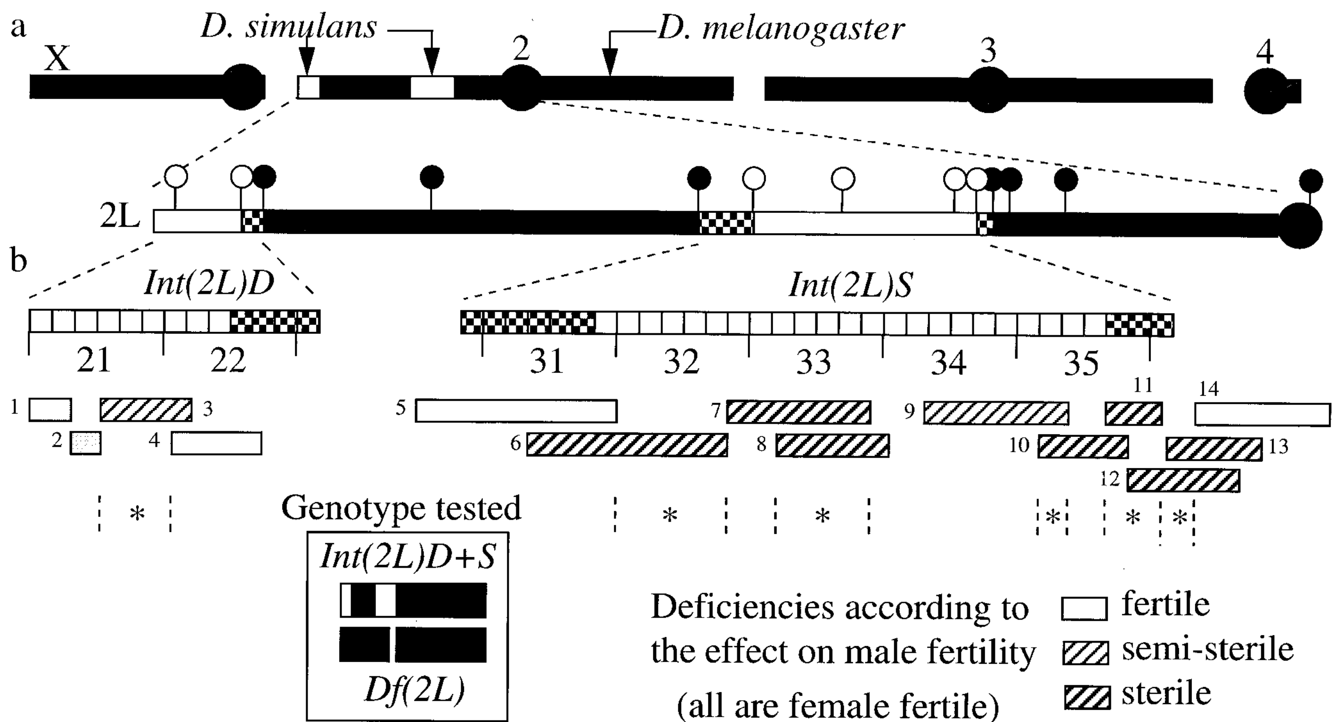

Fig. 3. (*a*) Mapping of *D. simulans* introgressions, *Int(2L)D+S*, by molecular markers. Dark and light bars denote *D. melanogaster* and *D. simulans* chromatids, respectively. Checkered bars denote regions of uncertain species origin. The markers used [and their cytological locations (18)] are, from *Left* to *Right*: [*ex* (21C3–4), *aop* (22D1–2)], *Pgk* (23A1–2), *Acp26Aa* (26A1–5), *bib* (30F1–6), [*da* (31E1–7), *prd* (33C1), *Adh* (35B2), *CycE* (35D7)], *grp* (36A10–14), *dl* (36C2), *Ddc* (37C1), *Rsp* (h39), where [ ] corresponds to the boundaries of introgressions. (*b*) Mapping genes of reproductive isolation by deficiencies. Each deficiency is tested over *Int(2L)D+S* as shown in *Inset*, where blank regions indicate introgressions, and the gap denotes a deletion. The numbers below the chromosome indicate the cytological intervals; further divisions of each interval into six letter regions (*A–F*) are also given. Fourteen deficiencies were used, and the fertility of the tested genotype is indicated. Loci of hybrid male sterility are indicated by stars; no female sterility is observed over any of these deficiencies. Deficiencies no. 10 and no. 11 result in inviability over the introgressions in the one subline that is homozygote inviable.

chromosomal deletions in *D. melanogaster*. In these experiments, the introgression sublines bearing both *D* and *S* components were used. All mappings are based on the fertility of flies heterozygous for the introgression and the deficiency, as shown in Fig. 3*b Inset*. The deficiencies shown in Fig. 3*b* generally overlap with the introgressions and, unless stated otherwise, do not cause inviability or sterility in the pure *D. melanogaster* background (i. e., are not haplo insufficient). The most salient observation of this study is that the majority of deficiencies uncover hybrid male sterility, whereas none gives rise to female sterility. If we postulate a factor in each overlapping region between sterility-causing deficiencies, the minimum number of loci for hybrid male sterility is six. Each putative locus is indicated by an asterisk flanked by dashed lines in Fig. 3*b*.

In comparison with previous analyses, which usually examined a relatively large segment of the introgressed chromosome for the existence of hybrid sterility or inviability genes (23, 24), this present study analyzes a small region uncoverable by a single deficiency at a time. Hence, that so many deficiencies overlapping the proximal *Int(2L)S* causes male sterility is surprising. It is equally unexpected that no female sterility is observed with any of the deficiencies, which collectively cover nearly the entire introgressions. If there are two or more loci on the introgressions that are jointly needed for female sterility but are too far apart to be uncovered by any single deficiency, none would have been detected. This may help to explain the observed female sterility when the entire *Int(2L)S* or *Int(2L)D+S* are made homozygous. (Recombination analysis has confirmed that genes within *Int(2L)S* are responsible for the female sterility.)

It is thus natural to wonder why there is so much hybrid male sterility (25). An increasingly popular view is that genes pertaining to male reproduction evolve faster than genes of other functions (26, 27), an idea originating from the observations of rapid changes in males' genitalia morphology (28). Our observations of the very high density of genes of hybrid male sterility, but little female sterility, offer further support for this view of sexual selection driving species differentiation, especially in their male reproductive systems (25).

Because only one of the six sublines is homozygote inviable, it appears that there is even less hybrid inviability than sterility. The disparity between hybrid sterility and inviability is in fact even greater because the observed sterility may be multifactorial (29–31). The molecular genetic basis of the one, and perhaps rare, case of inviability is thus of great interest. We mapped the inviability gene(s) by using a series of deficiencies that also include those shown in Fig. 3*b*. Among them, all except no. 10 and no. 11 of Fig. 3*b* could restore the viability. Precise complementation mapping is possible because the overlapping region between the no. 10 and no. 11 deficiencies is saturated with lethal mutations (32). The results of Table 1 clearly show that the inviability gene is allelic to *Cyclin E* (*CycE*). To our surprise, further molecular analysis of the *CycE* gene reveals a 3-Kb insertion to be uniquely associated with the inviable introgression but absent in both pure species as well as all other sublines (data not shown). Because this hybrid inviability is apparently induced during species hybridization, there is in fact no hybrid inviability at all in this region. Hybridizations between different populations or different species have been shown to activate transposition, presumably because the mechanisms of transpo-

Table 1. Deficiency and complementation mapping of an inviability gene on the introgression subline

	Genotype of tester strain						No. of flies		
Tester	*sna*	*lace*	*CycE*	*[l(2)35Df*	*l(2)35Di*	*Gli]*	Cy	Non-Cy	χ^2 test
Df(2L)b88c75	X	X	X	X	X	X	186	0	
Df(2L)TE35D-GW1	X	X	X				311	0	
$Df(2L)Sco^{rv25}$	X	X					335	148	NS
$Df(2L)Sco^{rv19}$	X						346	193	NS
sna^{HG31}	X						367	215	NS
$lace^{HG34}$		X					374	159	NS
$CycE^{TE35D}$			X				269	0	
$CycE^{PZ05206}$			X				135	0	
$l(2)35Df^{P15}$				X			464	222	NS
$l(2)35Di^{RAR8}$					X		489	270	NS
Gli^{P29}						X	556	314	NS

Int(2L)D+S/CyO females were crossed to the tester strains, which are *Df/CyO* (the top four) or *m/CyO* (the bottom seven), where *Df* and *m* are a deficiency and a mutation, respectively. The numbers of Cy and non-Cy flies from each cross are presented. Because *CyO* homozygotes are inviable the ratio of Cy/non-Cy is expected to be 2:1 if the tester does not carry a mutation allelic to the inviability gene in the introgression. The genotypes of the tester strains at each of the six loci are shown, with X denoting a null or mutant allele at that location. For example, the first deficiency is null at all six loci, the second one at three loci, and so on. The order of the last three loci (delineated with square brackets) is not known. The breakpoints of deficiencies are based on ref. 32 (J. Roote, personal communication). NS, not significant.

sition suppression break down in the hybrids (33, 34). We suspect that transposition in our hybrids may be quite active because the observed insertion is screened from a very small number of lines and in a small region coextensive with the introgression only. It is, however, premature to speculate whether our observation is purely fortuitous or is part of a general pattern.

The advantage of using *D. melanogaster* to study reproductive isolation seems evident. In this study of introgression, we have achieved a resolution that has not been attainable in most previous studies. Further elaborations that have been difficult in other systems are now feasible. By observing a very high density of hybrid sterility genes, we also extend the recent results that the genetic divergence between very closely related species can be extensive. We confirm that there is a discrepancy between the rate of divergence for male fertility loci and that for female fertility (or viability) loci. Such observations have been reported between much more closely related species pairs (23–25). Apparently, this contrast is not restricted to the early stage of species divergence, and the same forces that drive nascent species apart must continue at work at a higher level of species divergence. As divergence progresses, each genetic element causing incompatibility is expected to exert an increasingly stronger effect. The present introgression system thus presents a unique opportunity to analyze these genes and their effects on the emergence of hybrid incompatibility.

We thank Bloomington Drosophila Stock Center, T. L. Orr-Weaver (Massachusetts Institute of Technology, Boston), and J. Roote (Cambridge University, Cambridge, U.K.) for materials; and M. Ashburner, I. Boussy, T. L. Karr, M. F. Palopoli, C.-T. Ting, S.-C. Tsaur, M.-T. Yamamoto, and L.-W. Zeng for discussions and comments. We thank M.-L. Wu for technical assistance. K.S. was supported by a postdoctoral fellowship from the Human Frontier Science Program Organization, Strasbourg, France.

GENETICS

1. Dobzhansky, T. (1937) *Genetics and the Origin of Species* (Columbia Univ. Press, New York).
2. Mayr, E. (1942) *Systematics and the Origin of Species* (Columbia Univ. Press, New York).
3. Sturtevant, A. H. (1920) *Genetics* **5,** 488–500.
4. David, J., Lemeunier, F., Tsacas, L. & Bocquet, C. (1974) *Ann. Genet.* **17,** 235–241.
5. Lachaise, D., David, J. R., Lemeunier, F., Tsacas, L. & Ashburner, M. (1986) *Evolution* **40,** 262–271.
6. Dobzhansky, T. (1936) *Genetics* **21,** 113–135.
7. Wu, C.-I. & Palopoli, M. F. (1994) *Annu. Rev. Genet.* **28,** 283–308.
8. Ting, C.-T., Tsaur, S.-C., Wu, M.-L. & Wu, C.-I. (1998) *Science* **282,** 1501–1504.
9. Muller, H. J. & Pontecorvo, G. (1940) *Nature (London)* **146,** 199–200.
10. Muller, H. J. & Pontecorvo, G. (1940) *Science* **92,** 418, 476.
11. Henderson, N. R. & Lambert, D. M. (1982) *Nature (London)* **300,** 437–440.
12. Hollocher, H., Ting, C.-T., Pollack, F. & Wu, C.-I. (1997) *Evolution* **51,** 1175–1181.
13. Watanabe, T. K. (1979) *Jpn. J. Genet.* **54,** 325–331.
14. Hutter, P. & Ashburner, M. (1987) *Nature (London)* **327,** 331–333.
15. Sawamura, K., Taira, T. & Watanabe, T. K. (1993) *Genetics* **133,** 299–305.
16. Sawamura, K., Yamamoto, M.-T. & Watanabe, T. K. (1993) *Genetics* **133,** 307–313.
17. Davis, A. W., Roote, J., Morley, T., Sawamura, K., Herrmann, S. & Ashburner, M. (1996) *Nature (London)* **380,** 157–159.
18. FlyBase (1999) *Nucleic Acids Res.* **27,** 85–88.
19. Rutherford, S. L. & Lindquist, S. (1998) *Nature (London)* **396,** 336–342.
20. Stern, D. L. (1998) *Nature (London)* **396,** 463–466.
21. Takano, T. S. (1998) *Genetics* **149,** 1435–1450.
22. Goulielmos, G. N. & Alahiotis, S. N. (1989) *Genome* **32,** 146–154.
23. True, J. R., Weir, B. S. & Laurie, C. C. (1996) *Genetics* **142,** 819–837.
24. Hollocher, H. & Wu, C.-I. (1996) *Genetics* **143,** 1243–1255.
25. Wu, C.-I., Johnson, N. A. & Palopoli, M. F. (1996) *Trends Ecol. Evol.* **11,** 281–284.
26. Coulthart, M. B. & Singh, R. S. (1988) *Mol. Biol. Evol.* **5,** 182–191.
27. Tsaur, S.-C. & Wu, C.-I. (1997) *Mol. Biol. Evol.* **14,** 544–549.
28. Eberhard, W. G. (1985) *Sexual Selection and Animal Genitalia* (Harvard Univ. Press, Cambridge, MA).
29. Cabot, E. L., Davis, A. W., Johnson, N. A. & Wu, C.-I. (1994) *Genetics* **137,** 175–189.
30. Palopoli, M. F. & Wu, C.-I. (1994) *Genetics* **138,** 329–341.
31. Perez, D. P. & Wu, C.-I. (1995) *Genetics* **140,** 201–206.
32. Ashburner, M., Thompson, P., Roote, J., Lasko, P. F., Grau, Y., el Messal, M., Roth, S. & Simpson, P. (1990) *Genetics* **126,** 679–694.
33. Petrov, D. A., Schutzman, J. L., Hartl, D. L. & Lozovskaya, E. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8050–8054.
34. O'Neill, R. J. W., O'Neill, M. J. & Graves, J. A. M. (1998) *Nature (London)* **393,** 68–72.
35. Horton, I. H. (1939) *Genetics* **24,** 234–243.