



Differences in (G + C) Content between Species: a Commentary on Forsdyke's "Chromosomal Viewpoint" of Speciation*

RICHARD M. KLIMAN^{†‡}, BRYAN T. ROGERS[§] AND MOHAMED A. F. NOOR[§]

[†]*Department of Biological Sciences, Kean University, 1000 Morris Ave., Union, NJ 07083, U.S.A. and*
[§]*Department of Biological Sciences, Louisiana State University, LA, U.S.A.*

(Received on 8 May 2000, Accepted in revised form on 29 November 2000)

Forsdyke (1999) has recently argued that differences in (G + C)%, or G + C content, may trigger new species formation. He further argues that the genic model has shortcomings that can be overcome by his "chromosomal" (hereafter, "G + C") model. We disagree on several counts. First, we do not accept that the genic model has the shortcomings suggested by Forsdyke. There is an abundance of empirical support for the contribution of individual genes, as well as of mapped chromosomal regions, to post-zygotic reproductive isolation (and Haldane's rule). Further, we argue that the G + C model suffers from the same theoretical difficulties as other speciation models based on underdominance. We also question the evidence Forsdyke uses to support his model. Finally, we describe analyses of G + C content in a well-studied model system of speciation (the *Drosophila melanogaster* species complex), the results of which are incompatible with the G + C model. Thus, while Forsdyke's G + C model cannot be explicitly ruled out, it is not directly supported by empirical data. In contrast, the genic model is well supported by empirical data, holds up on theoretical grounds, and does not require any assistance from the G + C model.

© 2001 Academic Press

1. Introduction

Reproductive isolating mechanisms that produce hybrid sterility can involve both genic and chromosomal differences between species. Several lines of evidence have led researchers to believe that genic differences play a greater role in the *initiation* of the isolation process that leads to

species formation (Coyne & Orr, 1998). In recent papers, Forsdyke (1996, 1999) argues that differences in genomic base composition (G + C content) can explain the initiation of new species formation. He further argues that alleged shortcomings in the genic model (that post-zygotic reproductive isolation between species reflects genetic incompatibilities in the hybrid offspring) can be addressed by incorporating the effect of G + C content into the "chromosomal" model. Specifically, Forsdyke proposes that impaired gametogenesis in compositional heterozygotes is responsible for hybrid sterility in incipient species. While sequence divergence among taxa could naturally cause G + C content divergence, Forsdyke's suggestion that the latter contributes

*This manuscript is being submitted as a comment on: Forsdyke, D. R. (1999). Two levels of information in DNA: Relationship of Romanes' "intrinsic" variability of the reproductive system, and Bateson's "residue" to the species-dependent component of the base composition, (C + G)%. *J. theor. Biol.* **201**, 47–61.

[‡] Author to whom correspondence should be addressed.
E-mail: rkliman@turbo.kean.edu

directly to speciation is, in our view, not well supported.

We begin our discussion of Forsdyke's model by reminding the reader of the wealth of evidence in support of genic models of speciation. We then focus on the problems with Forsdyke's model, which we will henceforth call the G + C model (to distinguish it from the more traditional chromosomal model that deals with chromosomal macromutations). First, we explain the major theoretical problem associated with the G + C model, that it requires deleterious mutations to rise in frequency in at least one of the incipient species. Second, we explain why Forsdyke's *post hoc* evaluations of published data do not clearly support his model. Finally, we present an analysis of base composition variation within and among closely related species (the data for which has been available since 1993) which suggests that, at least in the case of *Drosophila simulans* and its closest relatives, the G + C model is unsupported on empirical grounds.

2. Support for the Traditional Genic Model

When discussing speciation, it is customary to distinguish the factors that lead to initial divergence from the secondary effects that either decrease the fitness of hybrids (post-zygotic isolation) or prevent mating between incipient species (pre-zygotic isolation). Although pre-zygotic isolation may be directly favored by natural selection (e.g. Dobzhansky, 1940; Noor, 1999), post-zygotic isolation can only be directly favored by natural selection under very limited conditions (Coyne, 1974; Johnson & Wade, 1995). Instead, divergence can result simply from independent evolution, either by drift or indirect natural selection, of geographically isolated populations. These incipient species may then produce hybrids with reduced fitness (i.e. the initiation of post-zygotic isolation).

The general concept of underdominance can be applied to any unit of heredity in a diploid system. Chromosomal macromutations, such as translocations or multiple inversions, can, in principle, lead to reduced fitness of chromosomal heterozygotes. Similarly, heterozygotes for individual genes can also have reduced fitness. The difficulty with using underdominance to explain

hybrid fitness reduction has been addressed at great length by population geneticists (see the discussion by Coyne & Orr, 1998). Basically, fixation of a mutation in one incipient species, either by drift or positive selection, is difficult if the original heterozygote is less fit than the homozygotes that represent the rest of the population—unless intermediate states are involved or selection is less effective due to inbreeding or otherwise low effective population size.

One solution is provided by the “Dobzhansky–Muller model”, elegantly described by Coyne & Orr (1998). Essentially, one mutation rises to fixation in one incipient species, while a mutation at a different gene rises to fixation in the second incipient species. The ancestral $A_1A_1B_1B_1$ genotype has typical fitness, as do the $A_2A_2B_1B_1$ and $A_1A_1B_2B_2$ genotypes of the respective incipient species. However, the novel hybrid genotype, $A_1A_2B_1B_2$, has reduced fitness due to epistasis (Dobzhansky, 1936; Muller, 1942; Orr, 1995). As pointed out by Coyne and Orr, the Dobzhansky–Muller model makes specific, *testable* predictions. First, there should be an exponential decline in hybrid fitness as the number of divergent interacting loci increases. Second, effects of reciprocal introgressions should be asymmetric, since the ancestral allele that remains in one species should not have much effect when recombined into the other species. Third, complex incompatibilities, those involving epistatic effects of three or more genes, may also be common. As Coyne & Orr (1998 and references therein) point out, the predictions of this genic model are borne out by experimental data.

In some species pairs, reproductive isolation is incomplete, and hybrid offspring can be produced. “Haldane's (1922) rule” is the observation that when hybrids of one sex show a greater reduction in fitness than the other (either by reduced viability or reduced fertility), the less-fit sex is generally the heterogametic sex. Citing Coyne (1992), Forsdyke (1999) argues that convincing genic explanations for Haldane's rule have “eluded everyone”, and that existing genic explanations must cope with “adaptive valleys” and other problems”. This view, however, requires that one ignores work from the past 7 years that has led to broad agreement on two major genic explanations contributing to Haldane's rule.

First, according to the “dominance theory” (Orr, 1993; Turelli & Orr, 1995), recessive X-linked incompatibilities should more seriously affect heterogametic hybrids than homogametic hybrids. Second, according to the “faster male theory” (e.g. Wu & Davis, 1993), diverging taxa tend to accumulate male-sterilizing incompatibilities more quickly than female-sterilizing incompatibilities. Both theories have gained broad support from empirical studies of hybrid sterility (see Wu *et al.*, 1996; Laurie, 1997; Orr, 1997; Presgraves & Orr, 1998; Turelli & Begun, 1997; Turelli & Orr, 2000). Interestingly, analyses by Presgraves & Orr (1998) demonstrate that Haldane’s rule does not require hemizyosity (i.e. a true heterogametic sex). They show that Haldane’s rule for hybrid infertility is obeyed in mosquitoes of the genus *Aedes* that have single-locus sex determination, ruling out any type of chromosomal basis for Haldane’s rule in these insects. Further, implicating simple allelic intermediate solves the problem of “adaptive valleys”.

Forsdyke’s (1999) explanation for Haldane’s rule relies on disproportionate G+C divergence between X and Y chromosomes, due to degeneration of loci on the Y chromosome (e.g. see Rice, 1994). Meiotic disruptions in hybrid offspring of the heterogametic sex would hypothetically cause sterility. However, interactions besides those involving the X and Y chromosomes cause hybrid sterility (e.g. Johnson *et al.*, 1992; Pantazidis *et al.*, 1993), and the Y chromosome sometimes has no effect on hybrid fertility (Coyne & Orr, 1989).

One of the most important model systems for studying the genetics of Haldane’s rule has been the species triad of *Drosophila simulans*, *Drosophila mauritiana* and *Drosophila sechellia*. In matings between these species, fertile hybrid female offspring are produced, while hybrid males are sterile or inviable. Introgression studies (e.g. True *et al.*, 1996) have identified at least a dozen chromosomal regions (including regions of the two large autosomal chromosomes) associated with reproductive isolation between *D. simulans* and *D. mauritiana*. The ability to map these factors to specific regions in the genome strongly suggests that factors represent genes. One chromosomal region that contributes directly to hybrid male sterility in *D. simulans* ×

D. mauritiana matings has been cloned and sequenced, and it contains the gene *Odysseus* (Perez *et al.*, 1993; Perez & Wu, 1995). This gene has a rapidly evolving homeodomain, the divergence being most easily explained by positive selection (i.e. adaptive evolution) on amino acid sequence (Ting *et al.*, 1998). Direct comparison of this gene’s sequence in representatives of each species demonstrates that G+C content is exactly the same (44%, C.-T. Ting, pers. comm.). Hence, the coding potential, and not local differences in base composition, at this locus contributes to its effect on hybrid male sterility. One could argue that G+C differences in, as of yet unsequenced, flanking DNA could produce the reproductive isolating effect via Forsdyke’s compositional model, but this would be special pleading. In fact, the repeated backcross method employed by Wu and colleagues to isolate the *Odysseus* locus would be expected, under Forsdyke’s “compositional” model, to rescue hybrid fertility, since such backcrossing should homogenize base composition. As Coyne & Charlesworth (1986) showed, however, this procedure did not decrease hybrid sterility associated with a region near the *forked* locus of *Drosophila simulans* and *D. mauritiana*. Obviously, given Wu and colleagues’ success in locating the gene, the same goes for *Odysseus*.

Essentially, while no evolutionary geneticist would suggest that evidence is available to explain *every* case of speciation by a genic model, it is far from a theory in crisis.

3. The Problem of Simple Underdominance in Forsdyke’s Model

Forsdyke argues that sequence differences between maternal and paternal chromosomes could interfere with synapsis during meiosis, triggering checkpoint control and, ultimately, impairing gametogenesis. Specifically, he argues that even small differences in G+C content alter the formation of stem-loop structures necessary for synapsis (see also Forsdyke, 1998). Thus, if two individuals have sufficiently different G+C, their offspring might suffer sterility as a consequence of this impaired gametogenesis.

As we indicated earlier, this type of model requires underdominance. The problem of underdominance is solved by, among other things, the

Dobzhansky–Muller model; that is, there are non-additive effects of allelic differences at individual loci on the quantitative character *hybrid fertility*. The G + C model, however, cannot easily get around the problem of underdominance. For there to be G + C divergence between species, there must first be deleterious G + C variation within species.

This is not to say that the problem of underdominance is irresolvable. The effects of compositional differences could be additive or synergistic, such that heterozygosity within species could be nearly or effectively neutral as long as an arbitrary threshold of G + C divergence—which would be exceeded in interspecies heterozygotes—was not reached. Inbreeding could also allow for rapid fixation of A/T ↔ G/C substitutions. G + C heterozygosity could be maintained by linkage disequilibrium with functional polymorphisms maintained by balancing selection, though one would have to then consider how many functional balanced polymorphisms are required for the G + C model to work. Finally, some mutations may be effectively neutral if effective population size is sufficiently low; this depends, of course, on the effect on fitness of these mutations, something that needs to be explicitly stated in the model. Where that threshold is placed in a quantitative model, and what type of epistatic model is used, will determine if sufficient divergence at G + C is likely to arise at a rate comparable to that of functional genic divergence. In fact, this is something that one could conceivably model and evaluate by computer simulation. Of course, the specific parameters and epistatic models employed would be open to debate, but at least the G + C model would be presented in a way that one could more objectively judge it.

Essentially, we do not argue that the G + C model is impossible. In the absence of quantitative analysis, however, the model may suffer from the same problems inherent in other models that require underdominance.

4. The Nature of the Supporting Evidence for Forsdyke's Model

Forsdyke's (1999) support for the G + C model is based on *post hoc* examination of published

research. He cites several papers that report base composition differences among species. However, it should be obvious that some compositional variation among species should be a byproduct of genetic divergence. Distinguishing cause and effect—specifically, showing that compositional divergence has *initiated* post-zygotic reproductive isolation in specific instances—is crucial to his case. Here, we challenge Forsdyke's suggestion that work on viruses, plants and yeasts supports the G + C model.

Forsdyke (1999) discusses an article by Yin & Hu (1997) on retroviruses, writing,

“In certain virus species, however, signs of the initial isolation process may still exist. This, indeed, may be the case with two virus species whose members have the potential to coexist in the same host cell (Yin & Hu, 1997); here they would exert a selection pressure on each other to sustain and expand any differences which might prevent recombination (otherwise they would recombine and destroy each other as individual species)”.

First, we take issue with the teleological argument suggested by the last clause. What is the source of this selection pressure against recombination? Why is it assumed that maintenance of viral diversity is important, from a fitness standpoint, to either species?

Regardless, do the findings of Yin and Hu support Forsdyke's contention that “signs of the initial isolation process might still exist?” The authors showed that two retroviruses from different subgenera, murine sarcoma virus (MSV) and spleen necrosis virus (SNV), were capable of copackaging genomic RNA. However, they do not propose any base composition-derived barrier to recombination between these two species. In fact, they note the public health implications of retroviral recombination.

Schachtel *et al.* (1991) compare the base compositions of two human herpes viruses. They propose that the compositional differences between herpes simplex virus 1 and varicella zoster virus are the result of independent evolution in different organisms or cell types. It does not seem reasonable that these very different viruses, with different host cells, should serve as a model

system for reproductive isolation driven by base composition differences.

Bronson & Anderson (1994) note that base composition varies dramatically among retroviruses, possibly because of variation in mutational biases during the process of reverse transcription. They also note that two retrovirus species inhabiting the same cell type often differ in base composition. At first, this might appear consistent with Forsdyke's model, since a (G + C)-based antirecombinogenic effect could prevent homogenization of the two viral species. This, however, presupposes that there would otherwise be sufficient interspecies recombination to homogenize them. But even in the absence of any antirecombinogenic effect, the two species would remain distinct so long as effective population size \times recombination frequency is small. The latter is a function of the *opportunity* for recombination, not just the mechanistic potential. Further, Bronson and Anderson note that the G + C differences in coexisting retroviruses are so great that the two species have different amino acid requirements, suggesting that they coexist because they occupy different ecological niches in the cells. One could reason that species with similar base compositions and amino acid requirements would be unable to coexist, leading to either competitive exclusion or, interestingly, diversifying selection that favors two very different base compositions.

Matassi *et al.* (1991) note that little recombination has occurred over a 6 million year period between the "parental" components of the allopolyploid tobacco genome. They suggest that differences in "nuclear architecture" of the parental genomes suppresses recombination, a point acknowledged by Forsdyke. (The reader should also note that the two components of an allopolyploid genome do not necessarily form tetravalents during meiosis.) Forsdyke (1999) states "in the light of the present work it seems possible that the differences in (C + G)% would themselves be responsible for the isolation"; in fairness, he does not otherwise promote this cause-and-effect relationship. Such an argument does not seem parsimonious when contrasted with a simpler model, whereby two genomes differing in chromosomal architecture, and coincidentally compositionally distinct, maintain compositional

identities because insufficient time has passed for organism-specific compositional biases to homogenize them toward the same compositional equilibrium. In fact, Matassi *et al.* are dealing with two genomes that were quite distinct from a compositional standpoint (albeit, heterogeneous within species due to isochore patterning). The G + C model, however, begins with a single species and, hence, compositional homogeneity between incipient species.

Forsdyke also cites an empirical study of two yeast species that are thought to produce sterile hybrids due to antirecombinational activity (Hunter *et al.*, 1996). However, the two yeast species studied by Hunter *et al.* differ at over 10% of their sites in coding regions of the genome, and about 20% of sites in non-coding regions. This divergence is much greater than would be expected in taxa that have recently speciated. It is, therefore, impossible to argue that this antirecombinational activity evolved more quickly than other barriers to gene exchange. Furthermore, the ability of mutations in the genes of the mismatch repair system to overcome the large *sequence* differences (irrespective of compositional differences) and significantly rescue the infertility between these two species argues for a strong *genic* component of speciation.

It would be ludicrous to suggest that a slime mold and a roundworm with the same overall G + C content ought to be interfertile, and it would be unfair to imply that Forsdyke would suggest this, as the essence of his hypothesis is that subtle differences in G + C content of genetically similar individuals have a role in the hybrid sterility leading to speciation. We can ignore the problem of underdominance for the moment and ask whether patterns of G + C variation in closely related species support Forsdyke's model. If Forsdyke is correct, one should expect significant G + C variation among recently diverged species; that is, there should be more variation than can be explained by variation within species. In the next section, we provide an analysis of G + C content of genes sequenced in the early 1990s in multiple strains of *Drosophila simulans*, *D. mauritiana*, *D. sechellia* and *D. melanogaster*. We show that the G + C model is not supported by the data.

5. G + C Content is not Correlated with Hybrid Infertility in *Drosophila*

The testable question in Forsdyke's hypothesis is whether (i) measurable differences in G+C content precede subsequent genetic divergence that reinforces reproductive isolation, or (ii) coincidental differences in G+C content arise by independent selection and genetic drift following genic reproductive isolation. A prediction from the compositional model is that the variation in G+C content among closely related species should be greater than the variation within a species. However, the analyses that follow show that overall variation in G+C content is due mainly to variation among genes and, more importantly, within species.

There is published data for five genes sequenced in multiple individuals of all four species of the *D. melanogaster* complex (Kliman & Hey, 1993; Hey & Kliman, 1993; Hilton *et al.*, 1994). Two of these, *asense* (*ase*) and *cubitus interruptus* (*ci*), are in regions of the genome where little within-species variation is expected (Begun & Aquadro, 1992). When the DNA sequences are compared within species, we find that *D. simulans* gene copies differ from each other, on average, by 0.0077 substitutions/base pair; this value for *D. mauritiana* is 0.0070. The average pairwise divergence values for the *ase* and *ci* genes are 0.0024 and 0.0058, respectively. The remaining three, *period* (*per*), *yolk protein 2* (*yp2*) and *zeste* (*z*), show an average pairwise divergence of 0.0123 (weighted by gene length) for *D. simulans* vs. *D. mauritiana*. Kliman, Hey and colleagues have recently collected more sequence data, and this average divergence value is essentially unchanged (Kliman *et al.*, in press). The average pairwise within-species variation at the *Esterase-6* gene in *D. simulans* is about 2.5% (Karotam *et al.*, 1995), actually slightly (though not significantly) greater than the divergence between *D. simulans* and *D. mauritiana* (manuscript in preparation). It is clear that, at least for some genes, there is more variation within species than there is between species.

We first performed one-way analysis of variance (ANOVA) on overall G+C content (following arcsine-square-root transformation) at the five loci to determine if there were significant

differences among species. The results are given in Table 1. The analyses indicate some significant differences between species pairs using Tukey's test for honestly significant differences. However, there is no consistent pattern. *D. simulans* and *D. sechellia* are identical at *ase*, while *D. mauritiana* and *D. sechellia* are identical at *ci*. The difference between *D. simulans* and *D. mauritiana* at the *yp2* locus is non-significant. There are no significant differences at *per* or *z*. On average, only 36.9% of the observed variation in G+C content stems from differences between species rather than from differences among individuals within species. Further, a two-way ANOVA (species \times genes) indicates that most of the variance in G+C content is among genes, with no significant direct effect of species (genes: $F_{4,75} = 56639.95$, $p < 0.001$, species: $F_{2,75} = 0.06$, $p = 0.941$, genes \times species: $F_{8,75} = 4.99$, $p < 0.001$). Put another way, the among-species variation in G+C content is no greater than would be expected by random sampling of within-species variation. The reader should note that the *D. mauritiana* strains used in the DNA sequence analyses are completely interfertile. The strains, along with others, were crossed in all pairwise combinations to generate a genetically variable population as part of a quantitative genetics study on abdominal bristle number. Subsequent generations retained normal fertility (C. Seay and R. Kliman, unpublished results).

One can always argue that there is something unique about this group of species that should disqualify it from a test of the G+C model. It does not seem that the partial fertility of these species pairs should disqualify them from consideration; to the contrary, this is evidence of relatively recent divergence. On the other hand, it would be improper to call them incipient, or partial, species. There is considerable divergence among these species in reproductive anatomy [e.g. male genitalia (Coyne, 1983; Coyne & Kreitman, 1986)], physiology [e.g. hydrocarbon pheromones (Coyne *et al.*, 1994)] and behavior [e.g. courtship songs (Cowling & Burnet, 1981; Cobb *et al.*, 1989)]. In short, they can easily distinguish among themselves. One could, however, point out that the power to detect significant among-species variation in G+C content is limited by the size of the data set. This may be

TABLE 1
G+C content at five loci in the D. simulans, D. mauritiana and D. sechellia

	<i>ase</i>	<i>ci</i>	<i>per</i>	<i>yp2</i>	<i>z</i>
N <i>sim</i> sequences*	6	9	6	6	6
<i>sim</i> range†	0.502–0.502	0.389–0.390	0.622–0.636	0.588–0.589	0.530–0.538
<i>sim</i> mean	0.502	0.389	0.631	0.588	0.534
N <i>mau</i> sequences*	5	6	6	6	6
<i>mau</i> range†	0.501–0.505	0.391–0.391	0.626–0.628	0.588–0.589	0.531–0.534
<i>mau</i> mean	0.503	0.391	0.627	0.589	0.533
N <i>sec</i> sequences*	6	4	6	6	6
<i>sec</i> range†	0.502–0.502	0.391–0.391	0.628–0.629	0.586–0.587	0.535–0.536
<i>sec</i> mean	0.502	0.391	0.629	0.587	0.536
<i>F</i> (d.f)‡	5.65 (2,14)	152.0 (2,16)	2.08 (2,15)	19.49 (2,15)	4.38 (2,15)
<i>P</i> §	0.016	<0.001	0.159	<0.001	0.032
Proportion of total variance from among-species differences	34.2%	58.6%	13.8%	49.2%	28.7%
<i>sim</i> vs. <i>mau</i> (HSD)	0.027	<0.001	0.140	0.309	0.607
<i>sim</i> vs. <i>sec</i> (HSD)	1.000	<0.001	0.461	0.001	0.163
<i>sim</i> vs. <i>sec</i> (HSD)	0.027	1.000	0.703	<0.001	0.028

Note: Sequences were originally reported as follows: all *per*, Kliman & Hey (1993); all *yp2* and all *z*, Hey & Kliman (1993); all *ase* and *ci* (*D. mauritiana* and *D. sechellia* only), Hilton *et al.* (1994); *ci* (*D. melanogaster* and *D. simulans* only), Berry *et al.* (1991).

* Number of gene copies sequenced for each species (*sim* = *D. simulans*, *mau* = *D. mauritiana*, *sec* = *D. sechellia*).

† Minimum-maximum G+C content in each species.

‡ $F = MS_{species}/MS_{within}$; d.f. = degrees of freedom.

§ *p*-value for *F*.

|| Statistical significance of *post hoc* two-species comparison using Tukey's test for honestly significant differences.

true, but the fact remains that there is still considerable within-species variation in G+C content, and Forsdyke's model does not yet explain how members of the same species remain interfertile in light of such variation.

We can also compare the G+C content of the outgroup species *D. melanogaster* to that of the species of the *D. simulans* complex. In general, *D. melanogaster* has lower overall G+C content than the others for the five genes analysed above. However, careful examination shows that this is due mainly to G+C content at codon third positions. In *D. melanogaster* and its relatives, all of the "preferred" synonymous codons end in either C or G (mainly C), so codon bias is strongly correlated to G+C content at codon third positions. There is good evidence that selection on codon usage became less effective in the *D. melanogaster* lineage after its split from the *D. simulans* lineage (Akashi, 1995; Kliman, 1999), the result being lower codon bias and, therefore, lower third codon position G+C content in the

D. melanogaster. When one looks, instead, at intron G+C content of these species, no clear pattern emerges, and there is considerable overlap (see Table 2).

6. Final Comments

The final section of Forsdyke's article is entitled "Criticisms Met". Essentially, he predicts the criticisms of his model that would be offered by Coyne & Orr (1998). A careful reading of their review would indicate that Coyne and Orr do not dismiss the possible role of chromosomal macro-mutations in reproductive isolation. They simply point out that genic divergence can do the job. Nevertheless, Forsdyke lists the following potential criticisms of his "chromosomal" model:

1. Many species lack chromosomal differences, and chromosomal differences between others may have accumulated after reproductive isolation was complete.

TABLE 2
*Intron G+C content in the D. melanogaster species complex**

Species	<i>ci</i> †	<i>per</i>	<i>yp2</i>	<i>z</i>
<i>D. melanogaster</i>	0.171 (0.171–0.171)	0.517 (0.511–0.522)	0.384 (0.382–0.388)	0.383 (0.383–0.383)
<i>D. simulans</i>	0.169 (0.169–0.169)	0.546 (0.534–0.550)	0.399 (0.394–0.409)	0.380 (0.377–0.388)
<i>D. mauritiana</i>	0.178 (0.178–0.178)	0.532 (0.529–0.534)	0.394 (0.394–0.394)	0.372 (0.365–0.377)
<i>D. sechellia</i>	0.178 (0.178–0.178)	0.522 (0.518–0.524)	0.379 (0.379–0.379)	0.388 (0.388–0.388)

* The *ase* gene lacks intron sequence.

† The species mean intron G+C is listed first; the minimum–maximum range is given below in parentheses.

2. Heterozygosity for chromosomal arrangements does not automatically decrease fitness.

3. It is difficult to rule out a genic effect even when chromosomal differences are correlated to impaired meiosis.

4. Recent work has demonstrated genic influences on hybrid fertility.

5. Haldane's rule is easily explained by a genic model; however, sterility due to impaired chromosome pairing should affect both sexes.

Forsdyke's response to the first two criticisms is that there is no conflict, because there is no single route to speciation. We agree with him on this point. However, the absence of criticism of the G+C model at the time of publication is not the same as support.

His response to the third point is that (i) Maside & Naveira (1996) have questioned whether hybrid sterility gene-mapping experiments (prior to 1996) had the resolving power to map genes of major effect; and (ii) sometimes hybrid sterility seems to be associated with chromosomal rearrangements. Both statements are correct, but neither relates to the validity of the G+C model. Essentially, Forsdyke states that both genic and macrochromosomal variation affect hybrid fitness, so we ask why he will not extrapolate genic influences to the *initiation* of post-zygotic isolation.

His response to the fourth criticism is based on the paper by Maside & Naveira (1996). That is, while Forsdyke concedes that hybrid sterility genes might be found, he notes the authors' asser-

tion that prior studies on *Drosophila* were over-interpreted. However, much of the *Odysseus* research, as well as other high-resolution introgression studies which showed genomic regions contributing to hybrid sterility (e.g. True *et al.*, 1996), was published after Maside and Naveira's paper was submitted.

Forsdyke's response to the fifth criticism is that his model would give the heterogametic sex a "head-start" toward speciation. That is, gametogenesis would be impaired due to compositional divergence between the X and Y chromosomes. There is a fundamental problem with this argument, at least as it relates to *Drosophila*: meiotic crossing-over does not occur in male fruit flies. On a more general level, the response to the criticism is similar to his responses to the others: nothing he has stated in his model is inconsistent with observations. Again, however, this is not evidence in favor of the G+C model. In fact, Coyne & Orr (1989) note that the Y chromosome often has no effect on hybrid sterility, a point difficult to reconcile with Forsdyke's suggestion that degeneration of the Y chromosome accelerates compositional divergence.

Researchers perpetually disagree on the content of papers. In this case, we feel that Forsdyke's (1999) article requires a response, because the genetics of speciation is of central importance in evolutionary biology. We believe that Forsdyke's article provides a misleading account of the state of the field. Regardless of whether differences in gene function or chromosome

architecture can fully explain hybrid sterility, our understanding of the genetics of post-zygotic isolation are not furthered by conjecture. For the model proposed by Forsdyke to have any impact on the understanding of the genetics of hybrid sterility, a detailed explanation of the types of data that would support or reject this model are needed.

We thank Jerry Coyne and four anonymous reviewers for helpful comments on the manuscript.

REFERENCES

- AKASHI, H. (1995). Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. *Genetics* **139**, 1067–1076.
- BEGUN, D. J. & AQUADRO, C. F. (1992). Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**, 519–520.
- BERRY, A. J., AJIOKA, J. W. & KREITMAN, M. (1991). Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. *Genetics* **129**, 1111–1117.
- BRONSON, E. C. & ANDERSON, J. N. (1994). Nucleotide composition as a driving force in the evolution of retroviruses. *J. Mol. Evol.* **38**, 506–532.
- COBB, M., BURNET, B., BLIZARD, R. & JALLON, J.-M. (1989). Courtship in *Drosophila sechellia*: its structure, functional aspects, and relationship to those of other members of the *Drosophila melanogaster* species subgroup. *J. Insect Behav.* **2**, 63–89.
- COWLING, D. E. & BURNET, B. (1981). Courtship songs and genetic control of their acoustic characteristics in sibling species of the *Drosophila melanogaster* subgroup. *Anim. Behav.* **29**, 924–935.
- COYNE, J. A. (1974). The evolutionary origin of hybrid inviability. *Evolution* **28**, 505–506.
- COYNE, J. A. (1983). Genetic basis of differences in genital morphology among three sibling species of *Drosophila*. *Evolution* **37**, 1101–1118.
- COYNE, J. A. (1992). Genetics and speciation. *Nature* **355**, 511–515.
- COYNE, J. A. & CHARLESWORTH, B. (1986). Location of an X-linked factor causing male sterility in hybrids of *Drosophila simulans* and *D. mauritiana*. *Heredity* **57**, 243–246.
- COYNE, J. A., CRITTENDEN, A. & MAH, K. (1994). Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science* **265**, 1461–1464.
- COYNE, J. A. & KREITMAN, M. (1986). Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. *Evolution* **40**, 673–691.
- COYNE, J. A. & ORR, H. A. (1989). Two rules of speciation. In: *Speciation and its Consequences* (Otte, D. & Endler, J. A., eds), pp. 180–207. Sunderland, MA: Sinauer Associates.
- COYNE, J. A. & ORR, H. A. (1998). The evolutionary genetics of speciation. *Philos. Trans. Roy. Soc. Lond. B* **356**, 287–305.
- DOBZHANSKY, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**, 113–135.
- DOBZHANSKY, T. (1940). Speciation as a stage in evolutionary divergence. *Am. Nat.* **74**, 312–321.
- FORSDYKE, D. R. (1996). Different biological species "broadcast" their DNAs at different (G+C)% "wavelengths". *J. theor. Biol.* **178**, 405–417.
- FORSYDKE, D. R. (1998). An alternative way of thinking about stem-loops in DNA. A case study of the human *GOS2* gene. *J. theor. Biol.* **192**, 489–504.
- FORSYDKE, D. R. (1999). Two levels of information in DNA: relationship of Romanes' "intrinsic" variability of the reproductive system, and Bateson's "residue" to the species-dependent component of the base composition, (C+G)%. *J. theor. Biol.* **201**, 47–61.
- HALDANE, J. B. S. (1922). Sex-ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101–109.
- HEY, J. & KLIMAN, R. M. (1993). Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. *Mol. Biol. Evol.* **10**, 804–822.
- HILTON, H., KLIMAN, R. M. & HEY, J. (1994). Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* species complex. *Evolution* **48**, 1900–1913.
- HUNTER, N., CHAMBERS, S. R., LOUIS, E. J. & BORTS, R. H. (1996). The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* **15**, 1726–1733.
- JOHNSON, N. A., PEREZ, D. E., CABOT, E. L., HOLLOCHER, H. & WU, C.-I. (1992). A test of reciprocal X-Y interactions as a cause of hybrid sterility in *Drosophila*. *Nature* **358**, 751–753.
- JOHNSON, N. A. & WADE, M. J. (1995). Conditions for soft selection favoring the evolution of hybrid inviability. *J. theor. Biol.* **176**, 493–499.
- KAROTAM, J., BOYCE, T. M. & OAKESHOTT, J. G. (1995). Nucleotide variation at the hypervariable esterase 6 isozyme locus of *Drosophila simulans*. *Mol. Biol. Evol.* **12**, 113–122.
- KLIMAN, R. M. (1999). Recent selection on synonymous codon usage in *Drosophila*. *J. Mol. Evol.* **49**, 343–351.
- KLIMAN, R. M., ANDOLFATTO, P., COYNE, J. A., DEPAULIS, F., KREITMAN, M., BERRY, A. J., MCCARTER, J., WAKELEY, J. & HEY, J. (2000). The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **156**, 1913–1931.
- KLIMAN, R. M. & HEY, J. (1993). DNA sequence at the period locus within and among species of the *Drosophila melanogaster* complex. *Genetics* **133**, 375–387.
- LAURIE, C. C. (1997). The weaker sex is heterogametic: 75 years of Haldane's rule. *Genetics* **147**, 937–951.
- MASIDE, X. R. & NAVEIRA, H. F. (1996). A polygenic basis of hybrid sterility may give rise to spurious localizations of major sterility factors. *Heredity* **77**, 488–492.
- MATASSI, G., MELIS, R., MACAYA, G. & BERNARDI, G. (1991). Compositional bimodality of the nuclear genome of tobacco. *Nucl. Acids Res.* **20**, 5561–5567.
- MULLER, H. J. (1942). Isolating mechanisms, evolution and temperature. *Biol. Symp.* **6**, 71–125.
- NOOR, M. A. F. (1999). Reinforcement and other consequences of sympatry. *Heredity* **83**, 503–508.
- ORR, H. A. (1993). A mathematical model of Haldane's rule. *Evolution* **47**, 1606–1611.
- ORR, H. A. (1995). The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**, 1805–1813.

- ORR, H. A. (1997). Haldane's rule. *Annu. Rev. Ecol. Syst.* **28**, 195–218.
- PANTAZIDIS, A. C., GALANOPOULOS, V. K. & ZOUROS, E. (1993). An autosomal factor from *Drosophila arizonae* restores spermatogenesis in *Drosophila mojavensis* males carrying the *D. arizonae* Y chromosome. *Genetics* **134**, 309–318.
- PEREZ, D. E. & WU, C.-I. (1995). Further characterization of the *Odysseus* locus of hybrid sterility in *Drosophila*: one gene is not enough. *Genetics* **140**, 201–206.
- PEREZ, D. E., WU, C.-I., JOHNSON, N. A. & WU, M. L. (1993). Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA marker-assisted mapping and characterization of a hybrid-male sterility gene, *Odysseus* (*Ods*). *Genetics* **134**, 261–275.
- PRESGRAVES, D. C. & ORR, H. A. (1998). Haldane's rule in taxa lacking a hemizygous X. *Science* **282**, 952–954.
- RICE, W. R. (1994). Degeneration of a nonrecombining chromosome. *Science* **263**, 230–232.
- SCHACHTEL, G. A., BUCHER, P., MOCARSKI, E. S., BLAISDELL, B. E. & KARLIN, S. (1991). Evidence for selective evolution in codon usage in conserved amino acid segments of human alphaherpesvirus proteins. *J. Mol. Evol.* **33**, 483–494.
- TING, C.-T., TSAUR, S.-C., WU, M.-L. & WU, C.-I. (1998). A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**, 1501–1504.
- TRUE, J. R., WEIR, B. S. & LAURIE, C. C. (1996). A genome-wide survey of hybrid incompatibility factors by introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* **142**, 819–837.
- TURELLI, M. & BEGUN, D. J. (1997). Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* **147**, 1799–1815.
- TURELLI, M. & ORR, H. A. (1995). The dominance theory of Haldane's rule. *Genetics* **140**, 389–402.
- TURELLI, M. & ORR, H. A. (2000). Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* **154**, 1663–1679.
- WU, C.-I. & DAVIS, A. W. (1993). Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. *Am. Nat.* **142**, 187–212.
- WU, C.-I., JOHNSON, N. A. & PALOPOLI, M. F. (1996). Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol. Evol.* **11**, 281–284.
- YIN, P. D. & HU, W.-S. (1997). RNAs from genetically distinct retroviruses can copackage and exchange genetic information in vivo. *J. Virol.* **71**, 6237–6242.