

REPLY

## Genes and speciation

CHUNG-I WU

*Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA*

### Reproductive isolation: separating the pools of genomes

Reproductive isolation (RI) is a population genetic concept denoting that two groups do not share a common gene pool. When the F1's are completely inviable or sterile in both sexes, it is obvious that the groups are reproductively isolated. Ambiguity arises when RI is not complete (as in most cases germane to the understanding of the process of speciation). Two groups that produce hybrids that have reduced (but nonzero) viability, mating success, fertility, etc. could still be sharing a common gene pool. Clearly, RI should not be defined by hybrid fitness reduction, as is often done. It is a concept of separation of gene pools.

Let us now consider a hypothetical example which illustrates the conceptual and operational difficulties with RI being the primary criterion of speciation. Imagine two groups that are divergent in morphology, behaviour and reproductive biology. Hybrids are formed frequently where the two groups overlap. The divergent characters are determined by genes scattered over five (but only five) of their 10 chromosomes. These underlying genes are so strongly coadapted that any combination, other than the parental constitutions, suffers substantial pre- and post-mating fitness reduction. These two groups will thus remain permanently distinct and their divergence will become progressively greater.

Given that more than half of the two groups' genomes still form a common pool, whereas the rest is separated into two distinct gene pools, we ask this question: are these two groups reproductively isolated? Indeed, whether the two groups are reproductively isolated may be more difficult to answer than whether they are different species. It is intriguing that E. Mayr (this issue) and some others (e.g. J. Britton-Davidian, this issue; H.A. Orr, this issue) appear to give an answer of yes to the above question. Would it then not be possible to claim two populations to be reproductively isolated even when 90% of their genomes form a common gene pool? Under such a relaxed standard of RI, most populations can probably be considered reproductively isolated. After all, even human populations fail to share a small portion of their genomes (e.g. Broadberry & Lin, 1996; Hamblin & Di Rienzo, 2000).

*Correspondence:* Chung-I Wu, Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA. Tel.: +1 773 702 2565; fax: +1 773 702 9740; e-mail: ciwu@uchicago.edu

RI, as applied to the whole genome, is not without sound theoretical justification. With appropriate levels of selection, recombination, epistasis and migration, whole-genome RI can be achieved with only a modest number of 'speciation genes'. (Speciation genes are loci where the allelic form of one population cannot be integrated into the genomes of another population because of fitness reduction in the foreign genetic and/or ecological background; in other words, such genes are differentially adapted). For example, it takes only 10 speciation genes, each affecting the 10 adjacent centimorgans (Eqn 1 of C-I. Wu, this issue), to congeal a genome of 100 cM. The defence by the proponents of RI, however, is to allow the genome to be 'porous', presumably because everyone knows that 'the isolating mechanism comprises only a very small portion of the genome' (E. Mayr, this issue). Such a defence is not only unnecessary but represents a retreat down a slippery road. Because 10% porosity is no more defensible than 90%, the concept of RI with nontrivial gene flow becomes vacuous.

To be logically tight, RI has to be a whole-genome concept (hence, the above example is *not* reproductively isolated). Whether it applies to the real world, especially between nascent species, has to be answered by empirical observations. It is time to consider whole-genome RI as a genetic hypothesis to be tested, rather than a fact established in the absence of rigorous genetic data. This is not a radical proposal as the concept of genome-wide RI has not been well accepted by plant scientists (L.H. Rieseberg and J.M. Burke, this issue) and animal data have cast further doubt on its generality (C-I Wu, this issue).

### Differential adaptation and RI

The genic view of speciation also redirects attention from RI to differential adaptation. An expressed concern is that the re-focusing is too restrictive, ignoring nonadaptive phenomena like chromosome rearrangements (J. Britton-Davidian, this issue), polyploidization and cytoplasmic incompatibilities (J. van Alphen, this issue; H.A. Orr, this issue), which are only briefly discussed in the perspective. The issue is how common these mechanisms are. How often does polyploidization or cytoplasmic symbiont play a role in the RI between nascent species? On a grander scale, these two mechanisms probably should be considered special cases, although interesting they may be. Chromosomal rearrangement is a different matter. Although rearrangements sometimes cause fertility reduction, the question is how these rearrangements ever become fixed (and cause fertility reduction as a byproduct). Despite the extensive documentation of selection on chromosomal rearrangements (Dobzhansky, 1970), the arguments for the role of rearrangements in RI seem to ignore whether and how selection drives them to fixation. In this respect, chromosomal rearrangements and genic substitutions causing RI are not different conceptually.

Another related concern is that the genic substitutions underlying RI could be neutral, rather than adaptive (H.A. Orr, this issue; K.L. Shaw, this issue). This can only be ascertained with the genes in hand. So far, every known gene that pertains to species or race differentiation bears the signature of positive selection (e.g. Lee *et al.*, 1995; Takahashi *et al.*, 2001; see C-I Wu, this issue, for others). As genes contributing to RI, by definition, render large fitness reduction in some genetic backgrounds, the nonadaptive argument must assume that these genes are effectively neutral in their native background. This is far more stringent an assumption than those behind the conventional neutrality argument that certain genetic changes are too subtle to make any fitness difference.

The relative importance of selection and drift has also played an important role in the debate on sympatric speciation (J. Mallet, this issue; J. van Alphen, this issue). In general, sympatric speciation may be viewed as a most extreme form of the model of Fig. 1 (C-I Wu, this issue) with unimpeded gene flow and strong diversifying selection. If sympatric speciation should prove to be common, the genic view of speciation will be even more relevant (Dieckmann & Doebeli, 1999; Kondrashov & Kondrashov, 1999). Nevertheless, it is premature to suggest that Darwinian selection driving speciation should be considered mainly in the sympatric mode (Schilthuizen, 2000; H.A. Orr, this issue). Allopatric and parapatric speciation driven by selection have been well established since the publication of 'Origin of Species', regardless of whether sympatric speciation will be proven to be common or not.

The issue that certain isolating mechanisms may not be a byproduct of differential adaptation but may themselves be the direct consequence of selection is raised by H.D. Rundle *et al.* (this issue). The most obvious example is the reinforcement mechanism which could lead to pre-mating isolation if post-mating isolation is already present. Although this is true, one still needs to explain how the post-mating isolation evolved in the first place. Moreover, genetic analysis has shown that pre-mating isolation can evolve extensively on its own (e.g. Hollocher *et al.*, 1997). The interest in reinforcement is understandable as it is a unique RI mechanism that is adaptive.

### The process of speciation vs. the definition of species

The process of speciation depicted in Fig. 1 (C-I Wu, this issue) is by necessity an oversimplified description. As there are many processes (behavioural, morphological, physiological, reproductive and so on) evolving concurrently, their relative rates would vary greatly in different taxa. Even within the genus of *Drosophila* (Doi *et al.*, 2001; Ting *et al.*, 2001), the dynamics can vary greatly. It should be emphasized that, in Fig. 1, RI is not narrowly defined as post-mating incompatibility. It can also be

behavioural (as emphasized by J.R. Bridle and M.G. Ritchie, this issue) or ecological misfit that results in reduced fitness in nature, even as the hybrids remain fairly viable and fertile. In *Drosophila*, species often evolve to stage IV with great reproductive and behavioural divergence but with little change in morphology whereas, in the cichlids, incipient speciation may be associated mainly with ecological, behavioural and morphological differentiation (E. Mayr, this issue; J. van Alphen, this issue).

Such disparity in rates of divergence may cause disagreement in defining species. It is quite possible to agree on the process of speciation without being able to agree on what species are, or at which stage species can be considered formed. Therefore, it seems most reasonable to define species in the context of the process of speciation. For example, one may choose to define species only when they reach stage IV, as an RI-based species definition would demand. By that, at least the basis of disagreement in defining species would be clear.

### Genes and speciation

This view of speciation is gene-based (strictly speaking, it is based on sets of interacting genes). An understanding of the process of speciation at the genic level is the crux of this view, which may be in contrast with the naturalists' thinking that has inspired most speciation studies. For example, E. Mayr (this issue) states that 'a completely consistent (speciation) scenario can be inferred without any reference to the genic basis'. Some concepts only make sense at the genic level. Random mating is one example where alleles are associated with one another independent of their own state, whereas individuals probably never mate randomly. (Even if individuals were to mate randomly, there would be no way of knowing). In a sense, species are nonrandom associations of alleles and so the genic basis does matter greatly. Because different genes would have different evolutionary dynamics and, hence, genealogical histories, sorting out these genes according to their effects (or lack of) on biological differentiation would be a most challenging task. This challenge will have to be addressed by proponents of the genealogical species concept as well (K.L. Shaw, this issue; A.P. Vogler, this issue).

Another reservation with this genic perspective is the heavy reliance on *Drosophila* data. Although the choices are few, it is a valid concern. Nevertheless, the data from the *D. melanogaster* subgroup are as favourable to the concept of whole-genome RI as one could hope for: genes of hybrid male sterility are astonishingly dense between *D. melanogaster* and *D. simulans* (Sawamura *et al.*, 2000). It is expected that the lower the density of speciation genes, the more porous the genome would be.

There are a number of methods that can lead to the systematic identification and characterization of speci-

ation genes, at least in some chosen systems. Such methods usually demand the collective efforts of a large segment of the field, as was the case for developmental genetics. The paucity of speciation genes identified so far may not indicate the difficulty of the task but rather may be a reflection of the conceptual state of the field. Genes are simply not perceived to be that crucial for a deep understanding of the process of speciation. A refocusing on the genic basis of speciation may lead to a more gene-oriented approach that goes beyond a simple analysis of inheritance pattern. Ultimately, experimentation at the molecular level should complement the hundred years of natural history observations in our quest to understand the origin of species.

### Acknowledgments

This reply has obviously benefited from the preceding commentaries which I appreciate sincerely. Loren Rieseberg, Jim Mallet, Ian Boussy, Howard Rundle and Arne Mooers provided very helpful inputs to this reply.

### References

- Broadberry, R.E. & Lin, M. 1996. The distribution of the MiIII phenotype among the population of Taiwan. *Transfusion Med.* **6**: 145–148.
- Dieckmann, U. & Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature* **400**: 354–357.
- Dobzhansky, T. 1970. *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- Doi, M., Matsuda, M., Tomaru, M., Matsubayashi, H. & Oguma, Y. 2001. A locus for female mate discrimination causing sexual isolation in *Drosophila*. *Proc. Natl. Acad. Sci.* **98**: 6714–6719.
- Hamblin, M.T. & Di Rienzo, A. 2000. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *Am. J. Hum. Genet.* **66**: 1669–1679.
- Hollocher, H., Ting, C.T., Wu, M.L. & Wu, C-I. 1997. Incipient speciation by sexual isolation in *Drosophila melanogaster*: Extensive genetic divergence without reinforcement. *Genetics* **147**: 1191–1201.
- Kondrashov, A.S. & Kondrashov, F.A. 1999. Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**: 351–354.
- Lee, Y.-H., Ota, T. & Vacquier, V.D. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* **12**: 231–238.
- Sawamura, K., Davis, A.W. & Wu, C-I. 2000. Genetic analysis by means of introgression into *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **97**: 2652–2655.
- Schilthuizen, M. 2000. Dualism and conflicts in understanding speciation. *Bioessay* **22**: 1134–1141.
- Takahashi, A., Tsaui, S.C., Coyne, J.A. & Wu, C-I. 2001. The nucleotide changes governing pheromonal variation in *Drosophila* and their evolution. *Proc. Natl. Acad. Sci.* **98**: 3920–3925.
- Ting, C.T., Takahashi, A. & Wu, C-I. 2001. Incipient speciation by sexual selection: concurrent evolution at multiple loci. *Proc. Natl. Acad. Sci.* **98**: 6709–6713.