

No.91

Reprint Series
20 November 1998, Volume 282

SCIENCE

Molecular Origin of Species

Masatoshi Nei and Jianzhi Zhang

Molecular Origin of Species

Masatoshi Nei and Jianzhi Zhang

In sexually reproducing organisms, new species originate when a reproductive barrier is established between different groups of organisms. This event, called speciation, is one of the most important biological processes, yet its mechanism remains largely unknown. Reproductive isolation between species can occur either before or after the formation of hybrid zygotes (1). In prezygotic isolation, mating or fertilization is prevented, whereas postzygotic isolation occurs by the sterility or inviability of the hybrid offspring.

The first step in understanding the mechanism of speciation is to identify the genes involved in speciation and to study their effects on molecular aspects of mating or development. Two research groups, one led by Vacquier (2) and the other by Wu (3) (page 1501 of this issue), did just this for prezygotic and postzygotic isolation, respectively. Previously, Vacquier and colleagues had shown that in abalone the fertilization of eggs by sperm is mediated by lysin, a protein that species specifically creates a hole in the egg envelope, and that in evolution this protein undergoes rapid amino acid substitution—apparently as a result of positive Darwinian selection (4). Swanson and Vacquier (2) have now cloned the egg receptor gene for lysin and shown that the receptor protein consists of about 28 repeats of a 153-amino acid sequence motif. The repeat sequences vary among species but are similar within species, apparently because of concerted evolution, which homogenizes the repeat sequences through unequal crossing over or gene conversion.

For many decades, fruitflies (*Drosophila* species) have been favorite organisms for studying speciation, and Wu and his colleagues have now used this organism to clone and characterize a gene involved in speciation. They used a genetic technique to narrow down the chromosomal location of a gene that causes the sterility of male hybrids between *Drosophila mauritiana* and *Drosophila simulans* and cloned it. They found a gene (*OdsH*) containing a 60-codon homeobox that is homologous to the paired-type subfamily of homeobox genes from mice, rats, and nematodes. The real function of this gene is unknown, but

it is likely to control the transcription of a target gene that participates in spermiogenesis.

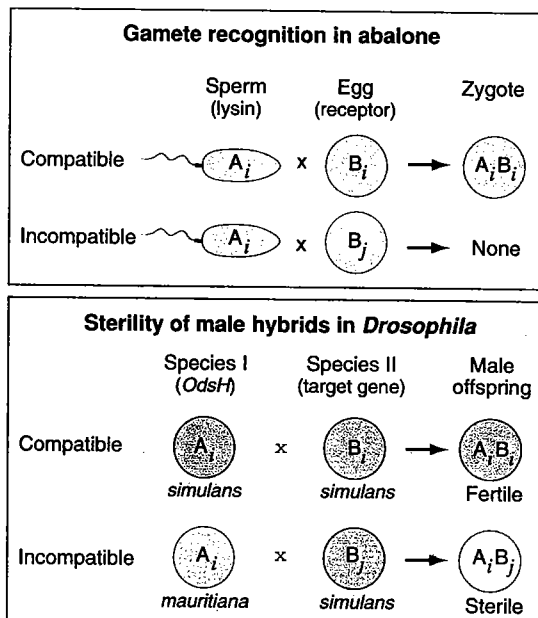
This homeobox gene has evolved surprisingly rapidly among the four sibling species of *Drosophila* (*mauritiana*, *simulans*, *sechellia*, and *melanogaster*), but it is highly conserved in other evolutionary lineages. The extent of sequence divergence among the homeoboxes of the four sibling species is much greater than that between

top panel). Similarly, the male offspring from a mating in *Drosophila* are fertile only when the sequences of *OdsH* and its target genes are compatible (see the figure, lower panel). In general, reproductive isolation between different species appears to be caused by the incompatibility of alleles at two or more loci that control mating, spermiogenesis, and development. The new findings in abalone and *Drosophila*, as well as a similar finding about the genes controlling the binding of sperm to eggs in sea urchins (5, 6), are consistent with this view.

If reproductive isolation is caused by the incompatibility of multiple alleles at different loci, what is the evolutionary mechanism of reproductive isolation? This

is a difficult question to answer. Suppose that an abalone species has alleles A_i and B_i at the lysin and the receptor loci, respectively, and a new lysin mutation, A_{i+1} , appears in a population. If this mutant allele is incompatible with the receptor allele B_i , it will never be fixed in the population. Theoretically, it is possible that B_i also mutates to B_{i+1} in the same population so that A_{i+1} and B_{i+1} become compatible. However, the chance of the A_{i+1} sperm meeting with the B_{i+1} egg would be vanishingly small in a large population. One way to avoid this problem is to assume that there are intermediate alleles between A_i and A_{i+1} (or B_i and B_{i+1}) and that these alleles are compatible either with B_i or B_{i+1} (or A_i and A_{i+1}). Mathematical models based on this idea indicate that the evolution of reproductive isolation can be developed by the incompatibility of multi-allelic loci (7). Some of the empirical observations (5) are compatible with this model.

The results from abalone, however, are not easy to explain with this model. In this organism the rate of amino acid substitution in lysin is about four times higher than that in the egg receptor. Swanson and Vacquier (2) suggest that the accelerated substitution in lysin is driven by concerted evolution. But concerted evolution is merely a process of homogenizing repeat sequences and cannot enhance the evolutionary rate of lysin. Theoretically, it is possible to modify the model proposed in (7) to enhance the evolutionary rate of locus A (lysin) by considering sperm competition, but this will also enhance the evolutionary rate of locus B (egg receptor). Because the egg receptor has evolved much slower than lysin, additional factors must be



Making new species, two ways. (Top) In this example of prezygotic isolation, fertilization is determined by the compatibility of allelic sequences at the lysin and the egg receptor loci (2). (Bottom) In postzygotic isolation, the sterility of male hybrids is determined by the compatibility of allelic sequences at the *OdsH* and its target gene loci (3). A_i and B_i are the i th alleles at loci A and B, respectively. The target gene in *Drosophila* has not been identified.

the rodent and nematode homeoboxes. Because the rodent and nematode homeobox homologs are expressed in neural tissue whereas *OdsH* is expressed in the testis, *OdsH* seems to have acquired a new function in spermiogenesis. The rapid evolution of *OdsH* appears to be due to this acquisition of a new function.

At first glance, these two studies of the molecular mechanisms of speciation might look unrelated, but in fact they show that the two different phases of speciation apparently have a similar genetic basis. In abalone the fertilization of eggs by sperm occurs only when the lysin and the receptor sequences are compatible (see the figure,

The authors are in the Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park, PA 16802-5301, USA. E-mail: nxm2@psu.edu

involved. One possibility is that to attain optimal binding between lysin and the receptor several amino acid changes are required in lysin when one change occurs in the receptor. This problem will probably be resolved if the three-dimensional structure of the egg receptor is clarified.

The molecular study of speciation has just begun. Although it has given us a

glimpse into the complex nature of gene interactions in the evolution of reproductive isolation, most problems remain untouched (8). As molecular biologists foray into this area, more surprising properties of gene interactions may be revealed.

References

1. T. Dobzhansky, *Genetics and the Origin of Species* (Columbia Univ. Press, New York, 1937).

2. W. J. Swanson and V. D. Vacquier, *Science* **281**, 710 (1998).
3. C.-T. Ting, S.-C. Tsaur, M.-L. Wu, C.-I. Wu, *ibid.* **282**, 1501 (1998).
4. Y.-H. Lee and V. D. Vacquier, *Biol. Bull.* **182**, 97 (1992).
5. E. C. Metz and S. R. Palumbi, *Mol. Biol. Evol.* **13**, 397 (1996).
6. R. A. Cameron, T. S. Walkup, K. Rood, J. G. Moore, E. H. Davidson, *Dev. Biol.* **180**, 348 (1996).
7. M. Nei, T. Maruyama, C.-I. Wu, *Genetics* **103**, 557 (1983).
8. V. D. Vacquier, *Science* **281**, 1995 (1998).