Long terminal repeat retrotransposons jump between species

Andrew J. Flavell*

Department of Biochemistry, University of Dundee, Medical Sciences Institute/Wellcome Trust Building Complex, Dow Street, Dundee DD1 5EH, Scotland

Retrotransposons are an evolutionarily ancient class of mobile genetic elements that transpose replicatively within their host genomes via RNA intermediates. There are three major retrotransposon groups, the Ty1-*copia* group and the *gypsy* group long terminal repeat (LTR) retrotransposons and the non-LTR retrotransposons or LINE elements (ref. 1; Fig. 1). All three groups are widespread in the eukaryotes, although none of the LTR retrotransposon groups have been detected in mammals or birds as yet. All LTR retrotransposons share striking sequence similarities with the retroviruses of vertebrates, and at least one LTR retrotransposon of *Drosophila* is in fact an infectious retrovirus (2). It is universally believed that modern day retroviruses, LTR retrotransposons, and non-LTR retrotransposons share a common ancestor, though there is some dispute about which came first (1, 3, 4).

Retrotransposons may look a lot like retroviruses, but they are not by themselves infectious. The basic difference between a retrotransposon and a retrovirus is the retrotransposon's lack of an envelope glycoprotein gene. This crucial difference prohibits the formation of an extracellular infectious virus particle, leaving the retrotransposon virus-like particle marooned inside its host cell (5). LTR retrotransposons cannot normally transfer themselves between adjacent cells; they certainly cannot move readily from one animal to another, and even more certainly, they cannot transfer horizontally from one species into another. In that case, why is there a paper in this issue of PNAS that proves beyond all reasonable doubt that an LTR retrotransposon called *copia* has been transferred from one species of *Drosophila*, *Drosophila willistoni*, into another, *Drosophila melanogaster*, within the last 200 years (6)?

We will address that question a little later in the article, but first, let us quickly review the evidence that supports the author's conclusions. The *copia* element of *D. melanogaster* was among the first LTR retrotransposons to be discovered (7). It is ubiquitous in *D. melanogaster*, and it has a broad distribution in other *Drosophila* species (8). *D. willistoni* is a *copia*-containing species, but *copia* is absent from many *willistoni* strains (6). Jordan *et al*. (6) have cloned and sequenced a fragment of the *D. willistoni copia*. This 1-kilobase fragment is identical to the most famous and well used *copia* clone, the *white-apricot copia* that was isolated from a spontaneous insertion of *copia* into the *white* locus of *D. melanogaster* (9, 10). Total sequence conservation of a retrotransposon across the 50-million-year gap that separates these two species is simply unbelievable; thus, either the transposon has jumped between the two species much more recently than their common ancestor, or there was some kind of contamination artifact in the experiment. Indeed, this observation set an alarm bell ringing for me, because this particular *copia* is also a clone that most fly labs have in their fridges and freezers. However, the *white-apricot copia* is also very successful at transposing in real flies, and it is reasonable to suppose that it was the one that was lucky enough to jump across into *D. willistoni*. The authors were no doubt as worried as I and have carried out a formidable series of controls to prove beyond any reasonable doubt that the *D. willistoni* genome does in fact contain integrated *copia* that is virtually identical to the *white-apricot copia*. The usual negative controls are all there to show that the tubes and solutions are clean. Southern analysis shows that there is a significant amount of *copia* present in the *D. willistoni* genome—not the tiny amount that would suggest contamination—and two further rigorous PCR controls prove that the *D. willistoni* DNA samples contain absolutely no contaminating copies of two *D. melanogaster* genes. Lastly, a *D. willistoni* genomic library has yielded a *copia* clone that is identical to the *white-apricot copia.*

Thus, *copia* has been transferred between two quite distantly related species of *Drosophila*. These two species have shared a host range for only the past 200 years. *D. willistoni* is a New World species and *D. melanogaster* was an African species until it began following humans—or, to be more precise, their rotting fruit—around the world. Moreover, in that time, another transposon has been transferred between these two species. The *P* element of *D. melanogaster*, which gained fame as a vector for germ-line transformation of that species, has been transferred from the *D. willistoni* subgroup, probably in this century (11). The two transfers of mobile elements have thus been reciprocal. Moreover, these two transposons are mechanistically very distinct from each other; the *P* element is a classical DNA transposable element that uses a DNA transposition intermediate. Any mechanism proposed to explain both transfers must take these observations into account.

What possible mechanisms are there? The evolutionary gap between these two species is almost certainly too large for even an abortive mating (M. Ashburner, personal communication); thus, the answer must be that a vector was responsible for the transfer. The most plausible candidate is a parasite with a broad species range, such as the mite *Proctolaelaps regalis*, which was proposed as a vector for the transfer of the *P* element (12). This hypothesis is attractive, because this mite feeds on *Drosophila* eggs by punching holes, sucking out the contents, and then moving on. *P. regalis* is a messy eater that does not kill every egg that it attacks, and it is reasonable to assume that it could transfer small amounts of the contents of one egg into another. Such a procedure could easily induce the transfer of the *P* element, and it is very plausible that transfer of the *copia* virus-like particle would also result in horizontal transfer.

The egg interior is the cytoplasm of a very large single cell, and the cytoplasm is a known intracellular location of the *copia* virus-like particle (13), which contains all of the necessary machinery for introducing a new copy of the retrotransposon into the genome. Unfortunately, this hypothesis is very difficult to substantiate in the laboratory. Nonetheless, it may well be that many of the possible classes of vector have at one time or another managed to ferry transposons from one species to another. Such possible vectors include other kinds of ani-

See companion article on page 12621.

*To whom reprint requests should be addressed.

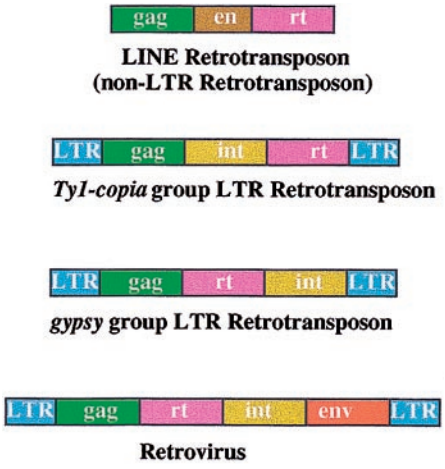

Fig. 1. The simplest common structures of LINE elements (non-LTR retrotransposons), the two groups of LTR retrotransposons and retrovirus proviruses. *copia* is a Ty1-*copia* group LTR retrotransposon. Structural domains: gag, core particle components; en, endonuclease; rt, reverse transcriptase; LTR, long terminal repeat; int, integrase; env, envelope glycoprotein.

mal parasites, such as wasps (14) and nematode worms, or microorganisms like fungi or bacteria (14) or viruses themselves (15). All that the vector would need is a broad host range and the ability to gain access to the germ cells somehow.

Evidence is accumulating slowly and steadily that many classes of transposable element have been transferred horizontally between species. One of the most primitive mobile elements, the group I intron, has probably invaded the mitochondrial genome of plants from fungi many times (16). Other DNA transposons, in particular the *mariner/Tc1* superfamily seem to have been particularly aggressive in their spread across the animal kingdom (17). LINE elements may have been transferred between snakes and mammals (18) and SINE elements (pseudogenes of small RNA genes that are propagated via RNA intermediates) have probably moved between fish species (19).

Why is the horizontal transfer of retrotransposons among multicellular eukaryotes particularly significant? All of the retroelements (genetic elements that use reverse transcription and RNA intermediates in their propagation) transpose replicatively, which leads to high levels of accumulation of these elements in many classes of organism. This phenomenon is particularly noticeable in mammals, birds, amphibians, and plants, where huge numbers of ancient relics of retrotransposons, retrovirus copies, and SINE elements accumulate, sometimes to the point where they take up large fractions of the total genome (20–22). In addition, it is commonly seen that the introduction of a transposon into a new host genome stimulates the transposition of the mobile element many-fold (23–25). Finally, transposition of retroelements results in gene mutation, causing a variety of genetic diseases, for example, in humans (26). Therefore, it is very important to understand the extent to which horizontal transfer of retroelements has occurred in the past histories of higher eukaryotes, the effects that this transfer has had on the genomes of those organisms, and the mechanisms of the transfers.

Jordan *et al.* (6) favor the view that horizontal transfer is a strategy for transposons to escape an "arms race" between the natural tendency of the transposons to transpose and the evolution of host-mediated mechanisms for repressing transposition (6, 27). It is surely true that "bailing out" of an increasingly antagonistic ecological niche could lead to an overall amplification in the number of transposons in existence. However, this situation does nothing for the poor wee transposons that are left behind, and, in that donor species, the rules stay the same: namely, multiply by transposition; perish by random mutation; or, if you are extremely fortunate, become useful for the host (as has been the case for the LINE elements that preserve the telomeres of *Drosophila*; refs. 28 and 29). The situation is even worse, because every transposon must compete for genomic space against the other mobile elements in the host. Finally, natural selection against the accumulation of transposon-induced deleterious mutations is constantly eliminating all but the fittest individuals. *Drosophila* may be a particularly antagonistic environment for transposons, because it has a small genome in which any transposition has a reasonable chance of knocking out a gene. This antagonistic environment is probably why *Drosophila* has relatively few retrotransposons in euchromatic regions (areas of open chromatin where most of the genes reside), but many of these transposons are apparently capable of transposition (Wei Li and A.F., unpublished observations). One of the biggest mysteries for me is how so many different types of transposons manage to coexist in such a difficult environment. There are at least 23 different LTR retrotransposons and retroviruses in the genome of *D. melanogaster* alone (see http://fly.ebi.ac.uk:7081/transposons/lk/melanogaster-transposon.html). Maybe horizontal transfer is useful for the mobile element, not so much to allow these mobile elements to escape their antagonistic host environment, but to give them time to mull over a few good ideas for resisting that environment and perhaps give the host time to forget how to deal with them. Perhaps we should follow the lead of those who model host–pathogen interactions (30) to come up with hypotheses for how horizontal transmission affects the evolution of retroelements and their host genomes.

I thank Margaret Kidwell for many helpful suggestions with this manuscript, Michael Ashburner for his useful opinions concerning the sexual practices of fruit flies, and Amar Kumar for information on the variety of LTR retrotransposons in *D. melanogaster*.

1. Xiong, Y. & Eickbush, T. H. (1990) *EMBO J.* **9,** 3353–3362.
2. Kim, A., Terzian, C., Santamaria, P., Pelisson, A., Prudhomme, N. & Bucheton, A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1285–1289.
3. Flavell, A. J., Pearce, S. R., Heslop-Harrison, J. S. & Kumar, A. (1997) *Genetica* **100**, 185–195.
4. McClure, M. A. (1999) in *Origin and Evolution of Viruses*, eds. Domingo, E., Webster, R. & Holland, J. (Academic, London), pp. 163–195.
5. Garfinkel, D. J., Curcio, M. J., Youngren, S. D. & Sanders, N. J. (1989) *Genome* **31,** 909–919.
6. Jordan, I. K., Matyunina, L. V. & McDonald, J. F. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 12621–12625.
7. Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42,** 1053–1063.
8. Alberola, T. M. & de Frutos, R. (1993) *J. Mol. Evol.* **36,** 127–135.
9. Lindsley, D. L. & Grell, E. H. (1968) *Genetic Variations of Drosophila melanogaster* (Carnegie Inst., Washington, DC) Publ. No. 627.
10. Mount, S. M. & Rubin, G. M. (1985) *Mol. Cell. Biol.* **5,** 1630–1638.
11. Daniels, S. B., Peterson, K. R., Srrasbaugh, L. D., Kidwell, M. G. & Chovnick, A. (1990) *Genetics* **124,** 339–355.
12. Houck, M. A., Clark, J. B., Peterson, K. R. & Kidwell, M. G. (1991) *Science* **253,** 1125–1128.
13. Shiba, T. & Saigo, K. (1983) *Nature (London)* **302,** 119–124.
14. Heat, B. D., Butcher, R. D. J., Whitfield, W. G. F. & Hubbard, S. F. (1999) *Curr. Biol.* **9,** 313–316.
15. Miller, D. W. & Miller, L. K. (1982) *Nature (London)* **299,** 562–564.
16. Cho, Y., Qiu, Y.-L., Kulman, P. & Palmer, J. D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14244–14249.
17. Hartl, D. L., Lohe, A. R. & Lozovskaya, E. R. (1997) *Annu. Rev. Genet.* **31,** 337–358.
18. Kordis, D. & Gubensek, F. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10704–10709.
19. Hamada, M., Yukiharu, K., Himberg, M., Reist, J. D., Ying, C., Hasegawa, M. & Okada, N. (1997) *Genetics* **146,** 355–367.
20. Smit, A. F. A. (1996) *Curr. Opin. Genet. Dev.* **6,** 743–748.
21. SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z., *et al.* (1996) *Science* **274,** 765–768.
22. Marracci, S., Batistoni, R., Pesole, G., Citti, L. & Nardi, I. (1996) *J. Mol. Evol.* **43,** 584–593.
23. Engels, W. R. (1983) *Annu. Rev. Genet.* **17,** 315–344.
24. Bucheton, A. (1990) *Trends Genet.* **6,** 16–21.
25. Blackman, R. K., Grimaila, R., Koehler, M. M. D. & Gelbart, W. M. (1987) *Cell* **49,** 497–505.
26. Kazazian, H. H. (1998) *Curr. Opin. Genet. Dev.* **8,** 343–350.
27. Maruyama, K. & Hartl, D. L. (1991) *J. Mol. Evol.* **33,** 514–524.
28. Sheen, F.-M. & Levis, R. W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12510–12514.
29. Biessman, H. & Mason, J. M. (1997) *Chromosoma* **106,** 63–69.
30. Lipsitch, M., Siller, S. & Nowak, M. A. (1996) *Evolution* **50,** 1729–1741.