Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 10753–10757, November 1992
Genetics

Genomic sequences with homology to the *P* element of *Drosophila melanogaster* occur in the blowfly *Lucilia cuprina*

(transposable elements/evolution/Lu-P1)

HARVEY D. PERKINS AND ANTONY J. HOWELLS

Division of Biochemistry and Molecular Biology, School of Life Sciences, Faculty of Science, Australian National University, P.O. Box 4, Canberra, Australian Capital Territory, 2601, Australia

Communicated by M. M. Green, July 21, 1992

ABSTRACT **We have cloned two DNA elements (Lu-P1 and Lu-P2) from the Australian sheep blowfly *Lucilia cuprina* that are similar to the transposable *P* element of *Drosophila melanogaster* in both structure and sequence but have diverged from it and from each other considerably. Hybridization studies indicate that a third related element probably exists in another, as yet unsequenced, clone. Neither Lu-P1 nor Lu-P2 appears to be active in terms of mobility, and it is not known whether any transposition-competent copies of other related elements occur in the genome of the blowfly. However, the isolation of any *P*-like sequences from a species outside of the family Drosophilidae allows comparisons to be made of more widely divergent *P*-related elements than has been possible previously. We are unaware of any report of the presence of multiple *P*-like family members within a single species. The discovery of Lu-P1 and Lu-P2 in the blowfly fuels the possibility that similar elements may be widespread in insects, and perhaps in other orders of animals.**

The transposable *P* element of *Drosophila melanogaster* has become an invaluable tool for genetic manipulation, experimental mutagenesis, and germ-line transformation in this species (1, 2); however, it is essentially nonfunctional in heterologous species. As a consequence, the basic and applied opportunities afforded by such germ-line transformation systems (3) have stimulated efforts toward their development in other organisms based on endogenous elements with similar properties.

The *P* element belongs to a class of transposable elements called transposons, which transpose at the DNA level and are characterized by the presence of a gene encoding a "transposase" enzyme and by inverted terminal repeat sequences. Other elements of this type are common in prokaryotes (4) and plants (5–7) and have been found in some invertebrates (8), but they appear to be relatively uncommon in insects or the higher orders of the animal kingdom, where most mobile DNA belongs to a class of elements that transpose by way of a reverse-transcribed intermediate.

However, a number of recent observations suggest that elements of the transposon type may be more widespread than has previously been thought. First, *P*-like elements have now been characterized from a range of *Drosophila* species, and a *P*-like element has also been described from *Scaptomyza pallida* (9), a more distantly related member of the family Drosophilidae. Second, elements of the transposon type have now been isolated from a range of insects and include *mariner* (10) and *hobo* (11) elements from various *Drosophila* species and TECth1 from *Chironomus thummi* (12); *mariner*-like elements have also been found in the silkmoth *Hyalophora cecropia* (13). Finally, it has been reported that there are regions of significant homology between the transposase genes of *hobo* and two plant transposons, the *Ac* element of maize, and *Tam3* from *Antirrhinum majus* (14), suggesting a distant but common evolutionary origin.

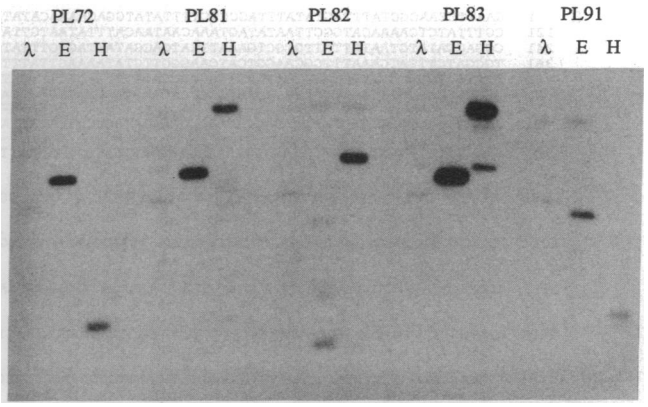

FIG. 1. Southern blot of five genomic clones probed with *P*-element DNA. Multiple hybridizing bands in *Hin*dIII digests of PL81 and PL83 are due to incomplete digestion; however, genuine hybridization occurs to several bands in both digests of PL82. The filter was hybridized and washed at 54°C and then exposed to x-ray film for 24 hr. Lanes: λ, *Hin*dIII-digested λ phage DNA; E, *Eco*RI digests; H, *Hin*dIII digests.

In this paper we report the cloning of two P-like elements from the blowfly *Lucilia cuprina*, providing further evidence that transposons may indeed be widely distributed in insects.*

MATERIALS AND METHODS

Fly Stocks. All fly strains used in this study were obtained from the Blowfly Genetics Group, Commonwealth Scientific and Industrial Research Organisation Division of Entomology (Canberra, Australia). They were originally collected from various geographical locations in Eastern Australia, from Queensland in the north to Flinders Island in Bass Strait in the south.

Library Construction. Genomic DNA, prepared from standard wild-type (SWT) embryos as described (15), was partially digested with *Sau*3A1 and, after partial end filling, ligated into commercially prepared half-filled *Xho* I-cut λGEM-11 vector (Promega).

Subcloning and Plasmid DNA. DNA was subcloned into pBluescript II KS vectors (Stratagene) by standard ligation and transformation techniques, using *Escherichia coli* host strain JPA101. Plasmid DNA was prepared by the alkaline lysis method (16).

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.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M89990 and M89991).

Southern Analysis. DNA was transferred to nitrocellulose in 20× standard saline citrate (SSC) (17) and hybridized in 50 mM Hepes/10× Denhardt's solution/3× SSC/0.1% SDS containing salmon sperm DNA at 50 μg/ml. Washing was in 2× SSC/0.1% SDS. DNA probes were labeled with [α-^{32}P]dCTP (1700 Ci/mmol, Amersham; 1 Ci = 37 GBq) by nick translation.

DNA Sequencing and Analysis. Exonuclease III was used to generate progressively deleted sequencing templates (18). Where a protective 3′ overhang could not be generated, exonuclease digestion was blocked with α-phosphorothioate deoxynucleotides (19) according to the Promega protocol. Sequencing of single-stranded templates was by the chain-termination method (20) as modified for use with Sequenase (United States Biochemical). Manipulation and analysis of sequence data were performed with MACVECTOR (IBI) software.

RESULTS

We have used an internal sequence of the *P* element [2.4-kilobase (kb) *Hin*dIII–Sal I fragment of pπ25.1; see ref. 21] from *D. melanogaster* as a hybridization probe to search for similar sequences in the genome of the Australian sheep blowfly *L. cuprina*. Genomic Southern blots were initially hybridized at normal and reduced stringency. Even under the least stringent conditions (52°C) no specific bands of hybridization were obtained, although some faint diffuse hybrid-

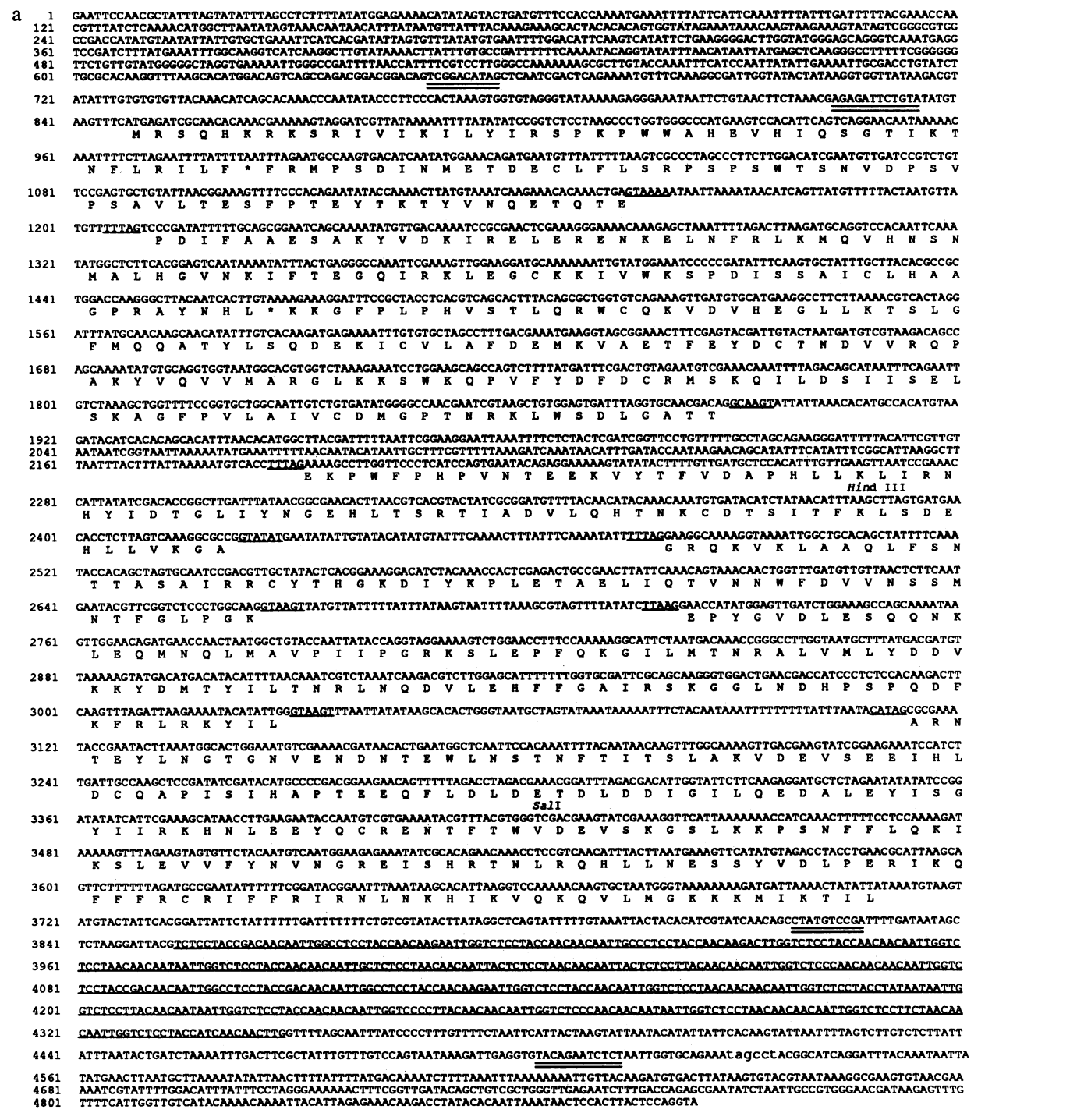

FIG. 2. (*Figure continues on the opposite page.*)

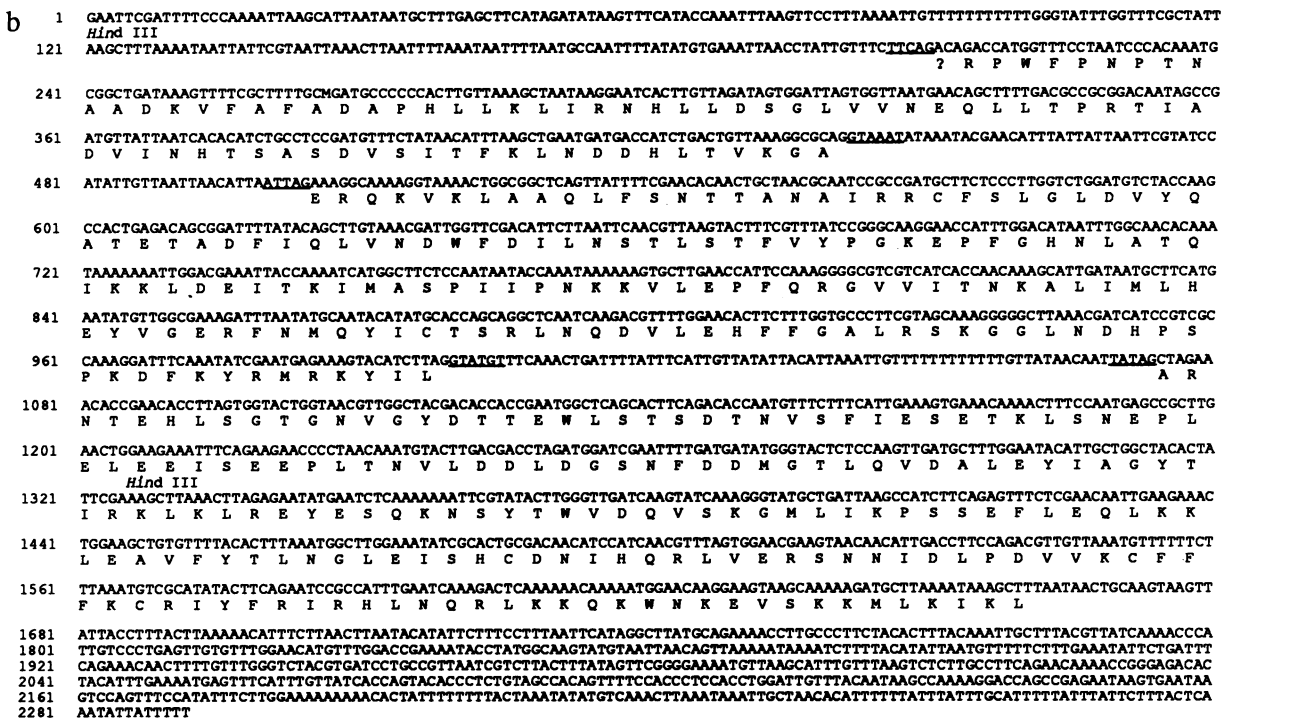

FIG. 2. Nucleotide and derived amino acid sequence for Lu-P1 (*a*) and Lu-P2 (*b*). Splice consensus sequences are indicated by single underlines; inverted repeats mentioned in the text are shown by double underlines. Restriction sites used to generate internal probes (see text and Fig. 4) are shown above the sequence. Also indicated by underlining is the 495-base-pair (bp) region of repeated sequence mentioned in the text. A short stretch of ambiguous sequence from PL83 (positions 4529–4534) is indicated by lowercase letters. The sequence from PL83 (Lu-P1) was obtained by cloning a 5.6-kb *Eco*RI fragment as overlapping 3.6-kb *Eco*RI–*Sal* I and 2.8-kb *Xho* I–*Eco*RI subclones, which were sequenced from both strands. The sequence ends about 550 bp upstream of the second *Eco*RI site. The 4.0-kb *Eco*RI fragment from PL91 (Lu-P2) was sequenced from one strand only for the first 2.3 kb.

ization was observed that appeared to correspond to satellite bands visible in the gel (results not shown). However, work in our laboratory aimed at the cloning of eye color genes (22) and retrotransposons from *L. cuprina* by cross-species hybridization has shown that, due to the amplification of the particular sequences, appropriate clones can often be isolated from genomic libraries despite a lack of clear signals on genomic Southern blots. Consequently a *L. cuprina* genomic DNA library was screened with the *P*-element probe under conditions of low stringency (54°C). From a screen of 35,000 clones (approximately one genome equivalent of *L. cuprina* DNA), about 30 positively hybridizing clones were detected. Five of these were plaque-purified and analyzed further. Preliminary restriction mapping indicated similarities between only two of the five clones (PL72 and PL81), and Fig. 1 shows the results obtained when DNA from all five was Southern blotted and probed with *P*-element DNA. Distinct bands of hybridization occur in all cases and the intensity of the signal varies amongst the five clones, with clone PL83 having the strongest.

DNA sequence was obtained from the *P*-hybridizing region of PL83 (Fig. 2*a*) and later, for comparison, from a smaller region of PL91 (Fig. 2*b*). These obviously related but different elements have been named Lu-P1 and Lu-P2, respectively. Analysis of the sequences produced a number of open reading frames and amino acid sequences with striking similarity to those of the *P* element of *D. melanogaster*. Alignments based on this similarity have enabled some structural features of Lu-P1 and Lu-P2 to be deduced (Fig. 3*a*) and a basis on which to assess the similarities and differences between these elements. While sequence from Lu-P1 appears to cover the corresponding full length of the *Drosophila P* element, amino acid homology does not begin until part way into the putative exon 1, then continues to the end of exon 3. No homology is detected between exon 0 of the *P* element and the putative exon 0 of Lu-P1. The incomplete sequence data obtained for Lu-P2 begin in the second intron, whereafter homology can be traced throughout exons 2 and 3. The upstream part of Lu-P2 is presumably present in the adjacent 6-kb fragment, which in *Eco*RI digests remains attached to the right arm of the vector and hybridizes to the *P*-element probe (see Fig. 1). Interestingly, in Lu-P1, "exon 2" is interrupted by two small introns not present in the *P* element, while Lu-P2 possesses only one of these additional introns. Otherwise the insertional position of intron sequences for the three elements is identical. Comparisons of the exon 2 sequences from pπ25.1, PS18 from *S. pallida* (9), Lu-P1, and Lu-P2 (summarized in Table 1) show that pπ25.1 and PS18 are more closely related than Lu-P1 is to Lu-P2, and that despite the ≈34% difference between them the Lu-P elements are about equally divergent from both pπ25.1 and PS18.

Despite the availability of 850 bp of sequence upstream of the proposed exon 0, and nearly 1200 bp downstream of the end of exon 3 of Lu-P1, no good evidence for the presence of inverted repeat termini could be found. (The 31-bp perfect inverted terminal repeats of the *P* element lie within 200 bp of the start and end of the transcription unit.) The best candidates are a 10-bp sequence (TCGGACATAG) and a 12-bp sequence (AGAGATTCTGTA) indicated in Fig. 2*a*; interestingly the 12-bp repeat shows some similarity to the 12-bp inverted repeat (CAGAGAACTGCA) of the *hobo* transposable element of *D. melanogaster* (11). However, neither of these inverted repeats is flanked by a direct repeat (transposable elements characteristically produce a "target site" duplication as a result of insertion) and neither is present in the 3′ sequences of Lu-P2. In the case of the 12-bp repeat, the situation is further obscured by the presence of a stretch of DNA between the end of exon 3 and the downstream 12-bp repeat which appears to encode a series of 24 tandemly repeated 7-amino acid motifs similar to those

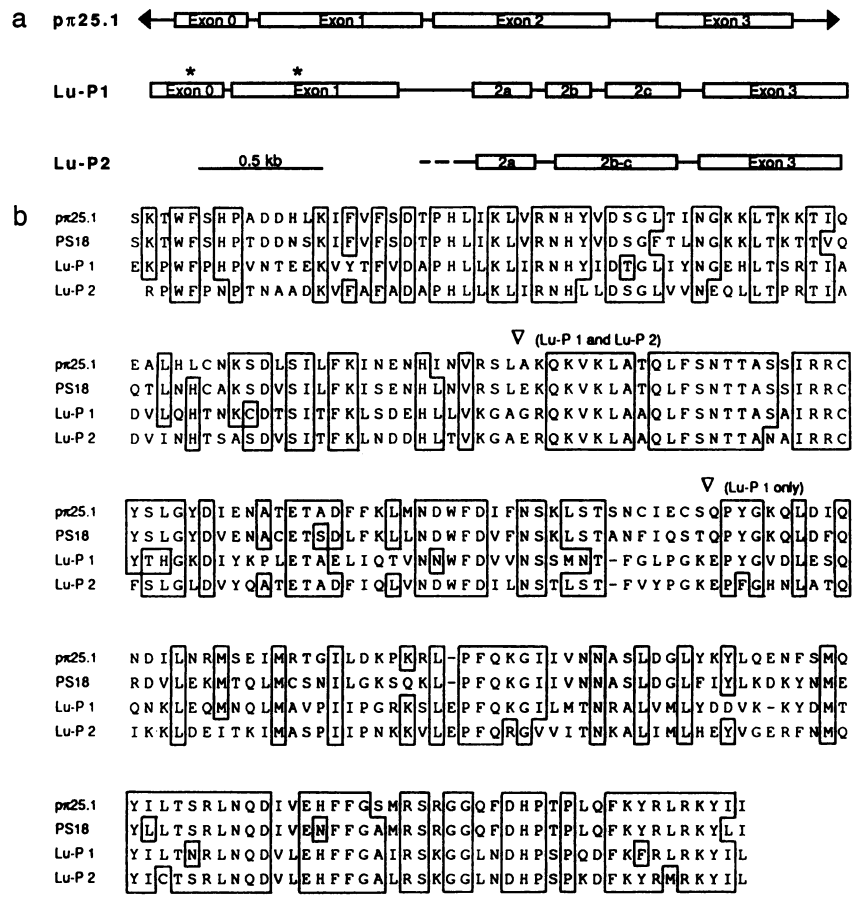

FIG. 3. Structural and amino acid sequence similarities between the *P* element (pπ25.1), Lu-P1, and Lu-P2. (*a*) Schematic representation of the structures of the three elements. The 31-bp inverted repeat termini of the *P* element are represented by arrowheads; stars mark the positions of stop codons in putative exons 0 and 1 of Lu-P1. Minor variations in exon size occur among the three elements; however, the main structural differences are the small introns interrupting exons 2 of the Lu-P elements. (*b*) Amino acid alignments for the exon 2 regions of the *P* element, PS18 (a P-like element from *Scaptomyza*; see ref. 12), Lu-P1, and Lu-P2. Dashes indicate gaps introduced to optimize the alignment, and the positions of introns in Lu-P1 and Lu-P2 are marked by arrowheads. Identical amino acids shared by at least three of the four sequences are boxed. Similar levels of homology are also found for exons 1 and 3.

present at the 3′ end of the largest subunit of RNA polymerase II (23, 24); although possible, it seems unlikely that these are part of the Lu-P1 structure. Hence the ends of these elements remain undefined.

The presence of single stop codons in both exon 0 and exon 1 of Lu-P1 indicates that it cannot produce an active transposase. While there are no obvious defects in Lu-P2, the lack of the entire sequence prevents us from speculating on its potential activity. To gain some indication of whether there are transposition-competent copies of either of these elements in the genome of the blowfly, we probed genomic Southerns of five different wild-type strains of *L. cuprina* with internal sequences from each element. The results (Fig. 4*a*) indicate that there is only a single copy of Lu-P1, which appears unchanged in all five strains, whereas Lu-P2 is apparently present in 5–10 copies, again with little difference in the pattern of hybridizing bands between strains. This indicates that neither of these elements is transpositionally active; similar findings were reported by Capy *et al.* (26) for the *mariner* element in *Drosophila sechellia*. The same probes from both Lu-P1 and Lu-P2 were also used to probe DNA from the five PL clones to determine the relationships between them. The results shown in Fig. 4*b* indicate that, apart from Lu-P1 and Lu-P2 (present in PL82 and PL91), a third element also exists (in PL72 and PL81) that is apparently phylogenetically intermediate between the other two but, given the stronger signal, more closely related to Lu-P2. Notably there is no hybridization between Lu-P1 and Lu-P2 under these high-stringency conditions (63°C), further illustrating the distinctness of these two elements.

Table 1. Matrix of the proportion of nucleotide and amino acid differences between the various *P*-like elements

	pπ25.1	PS18	Lu-P1	Lu-P2
pπ25.1	—	*0.219*	*0.496*	*0.475*
PS18	0.219	—	*0.533*	*0.492*
Lu-P1	0.410	0.426	—	*0.354*
Lu-P2	0.429	0.436	0.327	—

Values have been determined from the exon 2 region of each element only. The *P*-element sequence used was pπ25.1 (21), and PS18 is the *Scaptomyza pallida P*-like element (9). Nucleotide differences are shown in roman type, and amino acid differences in italics.

DISCUSSION

The discovery of the Lu-P sequences in *L. cuprina* is of considerable interest because they represent *P*-like elements isolated from a species outside of the family Drosophilidae and because, as far as we are aware, the isolation of three related but strongly divergent *P*-like elements from one species has not been reported before.

One of the most interesting aspects of the Lu-P elements, given their divergence and apparent immobility, is their high degree of integrity (e.g., absence of deletions or duplications and preservation of splice junction sequences). In Lu-P1 the helix–turn–helix motif and two of three leucine zipper motifs of the transposase also appear to have been preserved. This would argue either that the Lu-P elements have only recently become immobile or that they have been selectively maintained in the genome as non-mobile elements. The latter is perhaps more likely given the apparent lack of inverted repeat termini (at least for Lu-P1), a situation that is reminiscent of the *P*-like elements of *Drosophila guanche* and *Drosophila subobscura*, for which a role in the suppression of transposition has been suggested (27, 28). If the Lu-P elements also have some regulatory function, they may act by a mechanism different from that of the *P* element. Certainly, the production of a protein product equivalent to the 66-kDa *P*-element repressor (29) would require some read-through

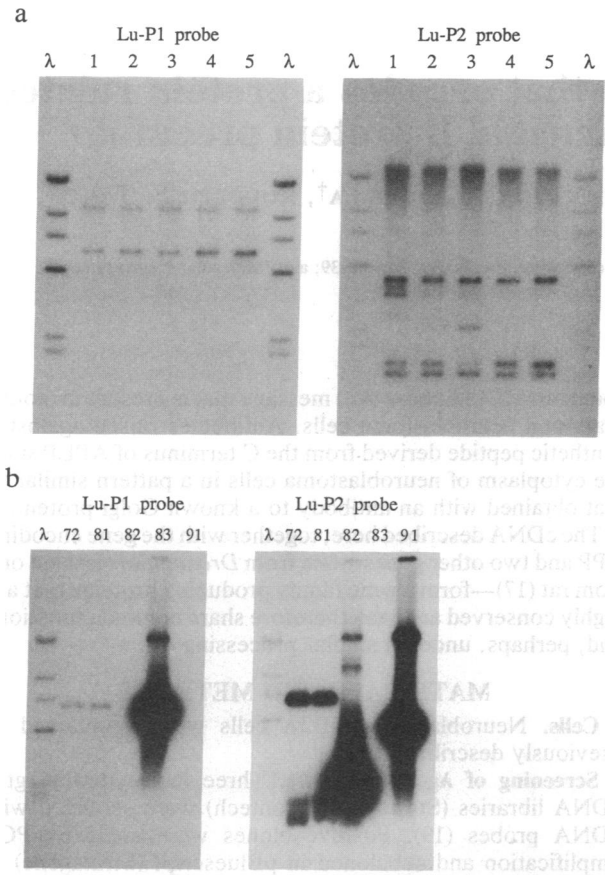

FIG. 4. Southern blots of genomic and cloned *L. cuprina* DNA probed with the two Lu-P elements. (*a*) Genomic DNA from five wild-type strains of *L. cuprina* digested with *Eco*RI and probed at 63°C with internal fragments of either Lu-P1 or Lu-P2. The Lu-P1 probe was a 1-kb *Hin*dIII–*Sal* I fragment, and the Lu-P2 probe a 1.2-kb *Hin*dIII fragment, both from regions spanning exons 2 and 3 (see Fig. 2). The presence of two bands in each digest probed with Lu-P1 appears to be due to polymorphisms between individual flies within each strain [the DNA used for these blots was prepared by the method of Lifton as described by Bender *et al.* (25), using five adult flies per preparation]. The slight variability in banding pattern with Lu-P2 is also likely to be due to polymorphism; however, the less distinct bands may represent cross-hybridization to other related elements. Lanes: λ, *Hin*dIII-digested λ phage DNA; 1, LBB; 2, SWT (standard wild type); 3, FBWT; 4, Weller; 5, Llandilo. (*b*) DNA from the five PL clones digested with *Eco*RI and probed as for *a*.

mechanism to permit translation of the stop codons in exons 0 and 1 of Lu-P1.

The extent of the sequence differences between the Lu-P elements, the *Drosophila P* element, and the *P*-like elements of *Scaptomyza* suggests that these sequences have been evolving independently for a substantial period of time, and comparisons of the nucleotide sequences from the various elements fit well with the phylogenetic relationships of the respective host species. It therefore seems likely that *P*-like elements were present in the common ancestor of *Drosophila* and *Lucilia* and have been evolving independently since their divergence, estimated to be about 110 million years ago (30). This suggests that the apparent horizontal transmission of the *P* element from *D. willistoni* to *D. melanogaster* (31) may be an isolated incidence of such transfer and that transmission of *P*-like sequences occurs primarily by normal vertical inheritance. Certainly our data provide no supportive evidence to suggest that lateral transfer has occurred between the *Drosophila* and *Lucilia* genera in the recent past.

In conjunction with the currently known distribution of *P*-like elements, these results suggest that the *P*-element family is much more diverse and widespread than previously thought and indicate that further work on the distribution of such elements is warranted. The sequences of Lu-P elements, when taken together with the sequences of the *P* element of *Drosophila* and the *P*-like elements from *Scaptomyza*, provide us with a better idea of the highly conserved regions of these elements, which enables the design of oligonucleotide primers to use in searching for similar elements in other species.

We thank P. East and P. Atkinson for comments on the manuscript. This work was supported by a postgraduate scholarship from the Australian Wool Corporation.

1. Cooley, L., Kelley, R. & Spradling, A. (1988) *Science* **239**, 1121–1128.
2. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218**, 348–353.
3. Cockburn, A. F., Howells, A. J. & Whitten, M. J. (1984) *Biotechnol. Genet. Eng. Rev.* **2**, 69–99.
4. Berg, D. E. & Howe, M. M., eds. (1989) *Mobile DNA* (Am. Soc. Microbiol., Washington).
5. Fedoroff, N., Wessler, S. & Shure, M. (1983) *Cell* **35**, 235–242.
6. Sommer, H., Carpenter, R., Harrison, B. J. & Saedler, H. (1985) *Mol. Gen. Genet.* **199**, 225–231.
7. Kikuchi, S., Liu, X. J., Frommer, W. B., Kostertopfer, M. & Willmitzer, L. (1991) *Mol. Gen. Genet.* **230**, 494–498.
8. Moerman, D. G. & Waterston, R. H. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington), pp. 537–556.
9. Simonelig, M. & Anxolabéhère, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 6102–6106.
10. Jacobson, J. W., Medhora, M. M. & Hartl, D. L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8684–8688.
11. Streck, R. D., MacGaffey, J. E. & Beckendorf, S. K. (1986) *EMBO J.* **5**, 3615–3623.
12. Wobus, U., Baumlein, H., Bogachev, S. S., Borisevich, I. V., Panitz, R. & Kolesnikov, N. N. (1990) *Mol. Gen. Genet.* **222**, 311–316.
13. Lidholm, D. A., Gudmundsson, G. H. & Boman, H. G. (1991) *J. Biol. Chem.* **266**, 11518–11521.
14. Calvi, B. R., Hong, T. J., Findley, S. D. & Gelbart, W. M. (1991) *Cell* **66**, 465–471.
15. Perkins, H. D., Bedo, D. G. & Howells, A. J. (1992) *Chromosoma* **101**, 358–364.
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
17. Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517.
18. Henikoff, S. (1984) *Gene* **28**, 351–359.
19. Putney, S. D., Benkovic, S. J. & Schimmel, P. R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7350–7354.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
21. O'Hare, K. & Rubin, G. M. (1983) *Cell* **34**, 25–35.
22. Elizur, A., Vacek, A. T. & Howells, A. J. (1990) *J. Mol. Evol.* **30**, 347–358.
23. Allison, L. A., Wong, J. K.-C., Fitzpatrick, V. D., Moyle, M. & Ingles, C. J. (1988) *Mol. Cell. Biol.* **8**, 321–329.
24. Nawrath, C., Schell, J. & Koncz, C. (1990) *Mol. Gen. Genet.* **223**, 65–75.
25. Bender, W., Spierer, P. & Hogness, D. S. (1983) *J. Mol. Biol.* **168**, 17–33.
26. Capy, P., Maruyama, K., David, J. R. & Hartl, D. L. (1991) *J. Mol. Evol.* **33**, 450–456.
27. Miller, W. J., Hagemann, S., Reiter, E. & Pinsker, W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4018–4022.
28. Paricio, N., Perezalonso, M., Martinezsebastian, M. J. & Defrutos, R. (1991) *Nucleic Acids Res.* **19**, 6713–6718.
29. Rio, D. C. (1991) *Trends Genet.* **7**, 282–287.
30. Beverley, S. M. & Wilson, A. C. (1984) *J. Mol. Evol.* **21**, 1–13.
31. Daniels, S. B., Peterson, K. R., Strausbaugh, L. D., Kidwell, M. G. & Chovnick, A. (1990) *Genetics* **124**, 339–355.