A Highly Repetitive, mariner-like Element in the Genome of Hyalophora cecropia*

Dan-Anders Lidholm, Gudmundur H. Gudmundsson, and Hans G. Boman
From the Department of Microbiology, University of Stockholm S-106 91 Stockholm, Sweden

Transposable elements are known to cause unstable mutations in a large number of organisms and the extensive literature has recently been summarized (Berg and Howe, 1989). In Drosophila melanogaster a significant part of the genome is made of middle repetitive DNA which is believed to be associated with transposable elements (Young, 1979). Some of these elements have been detected by their ability to cause dysgenesis, that is a type of germ line or somatic instability. The well known P-element of Drosophila (Engels, 1983) which is associated with germ line mutations (hybrid dysgenesis) has been developed into a vector for germ line transformation (Rubin and Spradling, 1982; Spradling and Rubin, 1982). The genetic instability in Drosophila is to a high extent limited to germ line tissue and somatic mutations are rarely detected (Rubin, 1983). In contrast is the fact that nematodes (Bryan et al., 1987; Bryan and Hartl, 1988; Medhora et al., 1988; Hartl, 1989). The element, named mariner, has been found in several Drosophila species but surprisingly not in D. melanogaster.

The discovery of mariner has introduced a novel class of transposable elements that represent some of the smallest eukaryotic transposons known. Until now mariner has only been found in Drosophila. We here report that during our analysis of the gene for preprocecropin A, a precursor to an antibacterial peptide from the lepidopteran H. cecropia (see accompanying paper; Gudmundsson et al., 1991), we found a highly repetitive DNA element homologous to the mariner element of D. mauritiana.

EXPERIMENTAL PROCEDURES

For procedures, see accompanying paper (Gudmundsson et al., 1991) and figure legends.

RESULTS

Structure and Location of the MLE—The gene for preprocecropin B was found to be composed of two exons and one intron of 514 bp (Xanthopoulos et al., 1988). Since preprocecropin A is quite similar to the B form (Lidholm et al., 1987), we expected an overall similarity also in the organization of the two genes. However, during the analysis of the gene for preprocecropin A, restriction mapping indicated an intron size exceeding 2 kilobases, about four times the expected size. We therefore sequenced the entire intron. Computer analysis of the intron sequence lead first to the discovery of two almost perfect inverted repeats (38 bp with two mismatches) separated by 1179 bp and second to the finding that the intron contained a defective reading frame with homology to the mariner transposable element found in D. mauritiana (Jacobson et al., 1986). Fig. 1 shows the position, organization, and restriction sites of the MLE in the gene for preprocecropin A. The element includes 1255 of the 2178 bp long intron, with the 5' end only 53 nucleotides from exon 2. Genomic Southern and expression analyses strongly indicates that the gene for preprocecropin A is functional despite the insertion (Medhora et al., 1988).

Copy of the complete DNA sequence of the MLE, including the inverted repeats. In comparison to mariner, the inverted repeats of the MLE contain 10 additional nucleotides. However, disregarding these extra nucleotides, Fig. 3 shows that positions 7–35 of the inverted repeat of the MLE are 68% homologous to the corresponding 28-bp repeat of mariner. In addition, the overall DNA homology to the mariner element in D. mauritiana is 48% (data not shown). Thus, it seems that the constraint on the primary structure of the terminal repeats has been stronger than for the internal sequence.

*This work was supported by Grant BU 2453 from the Swedish Natural Science Research Council and from the Swedish Board for Technical Development. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

1 The abbreviations used are: bp, base pair; MLE, mariner-like elements.
Copy Number and Genomic Distribution of the MLE—The genomic copy number of the MLE was tested by slot blot analysis using purified MLE DNA as a probe (Fig. 4). Densitometer evaluation of the autoradiographs indicate a copy number of about 1000/genome. This should be compared to a copy number of 20–30 in *D. mauritiana*. The genome size of Cecropia is estimated to be 2.5 × 10^9 bp or 1.7 times the size of that of *Drosophila* (Burke et al., 1976). As shown in Fig. 1, the inverted repeats of the MLE contain two restriction sites, unique for the element. The *SspI* site is almost at the border of the element while the *NdeI* site is located in the middle of each repeat. Assuming preservation of the inverted repeats, the *NdeI* restriction site was used in order to investigate the size distribution of MLEs in the Cecropia genome. Using isolated MLE DNA as a probe, genomic Southern analyses showed that the majority of the MLE copies fall into one major size class and also that these elements are similar in size to the MLE copy presented here (Fig. 5, left lane). Several Cecropia pupae were tested, and they all indicated the same copy number and distribution (data not shown). Our result shows that the population of MLEs is highly homogeneous with respect to size and that most or all of the elements include the inverted repeat structure presented here. It should be noted that the size resolution in the Southern experiment cannot detect copies having very small deletions. In the case of EcoRI restriction, it seems to be a more or less random distribution of the MLEs (Fig. 5, right lane) indicating numerous locations in the genome. This was also found to be the case for the mariner element (Jacobson et al., 1986).

The MLE Contains a Defective Reading Frame—The MLE contains a defective reading frame in opposite direction to the transcript of the gene for preprocecropin A. Fig. 6, upper sequence, shows the translated reading frame of MLE compared to the one in the mariner element. The predicted translation product of MLE seems to be defective in at least three ways: (i) three nonsense and/or frameshift mutations (indicated at positions 73, 114, 168); (ii) a deletion of 36 nucleotides producing a 12 amino acid deletion (position 132–143); and (iii) the absence of a stop codon (position 327). In addition, the potential promoter elements and the polyadenylation signal recognized in mariner are not convincingly detected in the MLE. A computer algorithm including conservative replacements shows that the overall homology of the predicted translation products is 34% identity or 56% similarity. Thus, the MLE found here is clearly related to mariner but defective in respect to its ability to encode a putative transposase.

Northern blot analysis was done in order to test the presence of intact MLEs (data not shown). Several developing pupae were tested for the presence of MLE transcripts. In no case was it possible to detect any transcript.

**Discussion**

The Cecropia gene for preprocecropin A (see accompanying paper, Gudmundsson et al., 1991) was found to have an insertion element in the intron. A data-bank search showed that the element was homologous to the mariner transposable element responsible for the unstable white mutation in *D. mauritiana* (Jacobson et al., 1986; Hartl, 1989), one of the smallest eukaryotic transposons known. We have designated the Cecropia element MLE (for Mariner-Like Element). The overall structure of the two elements is very similar but at the molecular level there are several differences. The MLE is 1255 bp including 38-bp terminal inverted repeats (Fig. 2) while mariner spans 1286 bp including the 28-bp terminal inverted repeats. The smaller size of the MLE might be due to a deletion since an alignment of the two sequences produces a 36-nucleotide gap flanked by good homology. The overall homology at the DNA level is 48%. Although the inverted repeats of MLE are longer than in mariner, excision of the first six and last three nucleotides of the MLE repeat leaves a sequence that is 68% homologous to the mariner repeat (Fig. 3). The terminal six to eight nucleotides of the MLE repeats could also include the duplication of the target site upon insertion of the element. Mariner contains an open reading frame that could produce a 345-amino acid protein and putative control elements such as consensus promoter and termination sequences are found at appropriate positions. The reading frame of the MLE seems to be defect because nonsense, frameshift, and deletion mutations have accumulated in a few positions (Fig. 6, upper sequence). Translation of the MLE reading frame provides further evidence for the relationship between the two elements. Considering conservative replacements, the overall similarity between MLE and mariner is 56% which includes short regions with almost full identity (e.g. MLE residues 44–54, 114–121, and 258–270 in Fig. 6).

Most probably the element in Cecropia is nonfunctional with respect to its ability to express the putative transposase.
A mariner-like element in H. cecropia

Fig. 3. Comparison of the inverted repeats of MLE and mariner. The MLE-inverted repeat was reduced to include nucleotides from position 7 to 35. Under this condition the homology is 68%. The sequences were computer aligned by the GAP program of the University of Wisconsin Genetic Computer Group.

Copy number of MLE

| MLE   | Genomic DNA |
|-------|-------------|
| 100 ng| 7 μg        |
| 50 ng | 3.5 μg      |
| 10 ng | 0.7 μg      |
| 1 ng  | 0.35 μg     |
| 0.1 ng| 0.07 μg     |
| 0.01 ng|            |

Fig. 4. Copy number estimation of the MLE. Various amounts of Cecropia genomic DNA and purified NdeI fragments corresponding to the MLE (see Fig. 1) were applied on a Hybond-N membrane (Amersham Corp.) by the use of a slot blot apparatus (Schleicher & Schuell). The amounts of DNA were recorded spectrophotometrically at A260. The membrane was hybridized at high stringency (see accompanying paper, Gudmundsson et al., 1991) with 32P-labeled MLE (SspI fragment seen in Fig. 1), and the film was analyzed by a densitometer (LKB). The estimation of the copy number presupposes that the genome size of Cecropia is 2.5 × 109 bp (Burke et al., 1976).

suggested for mariner. Still, the mobility of the MLE cannot be ruled out since the mobilization could be conferred by a trans-acting transposase provided by an autonomous equivalent. The fact that the absolute majority of the MLEs share a NdeI restriction site in the middle of their inverted repeats suggests that the repeats are highly preserved (Fig. 5). Thus, functional copies of MLE may still exist or the elements may just have become inactivated. However, a proper conclusion needs to be based on the comparison of several copies of the element or evidence for transcriptional activity.

To investigate the presence of functionally active elements, we checked several developing pupae for production of RNA corresponding to the MLE, but in no case could we detect MLE transcripts. Three interpretations are possible. (i) There are no functionally active MLEs present in the animals tested. (ii) The quantity of MLE RNA is below the level of detection. (iii) The MLE gene(s) is seldomly expressed and/or dependent or some interaction with the host. Thus, the expression of the MLE transcript could be at an embryonic or a larval stage rather than in pupae.

The copy number and distribution of the MLE differs in some respects to mariner. The MLE has a high copy number of about 1000/genome, and the absolute majority of the MLE copies have the same size and the same inverted repeats (Fig. 5). In contrast, the mariner element in D. mauritiania has a copy number of 20–30, and copies occur with internal deletions (Hartl, 1989). The reason for this vast difference in copy number cannot be simply explained. The genome of Cecropia is only slightly larger than the genome of Drosophila, but Cecropia (Burke et al., 1976) contains more repetitive DNA (30%) than Drosophila (10–15%) (Young, 1979). Although repetitive DNA has been associated with the presence of mobile DNA elements (Young, 1979), the size and composition of the Cecropia genome can hardly alone be responsible

Fig. 5. Genomic distribution of the MLE. 10 μg of genomic DNA from one Cecropia pupae was digested with either NdeI (N) or EcoRI (E) and separated by electrophoresis on a 0.7% agarose gel. The DNA was transferred to a Hybond-N membrane and probed at high stringency with 32P-labeled MLE (SspI fragment in Fig. 1). Molecular weight markers are indicated to the left.

Fig. 6. Amino acid comparison of the predicted translation products of MLE and mariner. The MLE sequence is translated (top) to give optimal homology to the predicted translation product encoded by mariner (bottom). Nonsense and/or frameshift mutations are indicated as a star (positions 73, 114 and 168) and the 12 amino acid deletion as a gap (position 132–143). The carboxyl-terminal 28 amino acids encoded by the MLE are excluded from the figure. Sequences were aligned by the GAP program of the University of Wisconsin Genetic Computer Group. Conservative replacements defined by the program are denoted.
A mariner-like element in H. cecropia

for the increase in copy number. More likely, the difference in transposition control involves chromosomal and/or cytoplasmic factors that are specific to the MLE and its host.

In D. mauritiana, the presence of a factor named Mos, has been shown to influence the excision of mariner from the white-eye locus (Medhora et al., 1988; Jacobson, 1990). The Mos factor, which is a special copy of mariner, is a transposable element regulating not only the excision and/or transposition of other mariner elements but also of itself. The activity of Mos is assumed to be the result of nucleotide sequence alterations or position effects mediated by the host genome. The occurrence of the Mos factor has also been suggested as a possible evolutionary mechanism for a rapid increase and decrease in the number of copies of mariner in the genome. Clearly the high copy number of the MLE in Cecropia indicates a very high transposition rate. If this is a result of the presence or absence of regulatory derivatives corresponding to the Mos factor remains to be investigated.

It is surprising that the mariner element is present in the lepidopteran Cecropia but not in in the dipteran D. melanogaster. A route of horizontal transfer between diptera and lepidoptera rather than vertical transmission is an interesting possibility that could also account for the high copy number in Cecropia. The wide host range of the mariner transposon suggests that it could be developed into a vector for interspecies transformation.

Acknowledgments—We thank Dan Hultmark and Jong-Youn Lee for critical readings of the manuscript.

REFERENCES

Berg, D. E., and Howe, M. M. (1989) Mobile DNA, American Society for Microbiology, Washington D. C.
Bryan, G. J., and Hartl, D. L. (1988) Science 240, 215-217
Bryan, G. J., Jacobson, J. W., and Hartl, D. L. (1987) Science 235, 1636-1638
Burke, D. D., Rubin, D., and Willis, J. H. (1976) J. Insect Physiol. 22, 791-798
Engels, W. R. (1983) Ann. Rev. Genet. 17, 315-344
Fedoroff, N. V. (1989) in Mobile DNA (Berg, D. E., and Howe, M. M., eds) pp. 375-413, American Society for Microbiology, Washington D. C.
Gudmundsson, G. H., Lidholm, D.-A., Åsling, B., Gan, R., and Boman, H. G. (1991) J. Biol. Chem. 266, 11510-11518
Hartl, D. L. (1989) in Mobile DNA (Berg, D. E., and Howe, M. M., eds) pp. 531-537, American Society for Microbiology, Washington D. C.
Haymer, D. S., and Marsh, J. L. (1986) Dev. Genet. 6, 281-291
Jacobson, J. W. (1990) Genet. Res. 55, 153-158
Jacobson, J. W., Medhora, M. M., and Hartl, D. L. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 8684-8688
Lemeunier, F., and Ashburner, M. (1984) Chromosoma 89, 343-351
Lidholm, D. A., Gudmundsson, G. H., Xanthopoulos, K. G., and Boman, H. G. (1987) FEBS Lett. 226, 8-12
Medhora, M. M., MacPeek, A. H., and Hartl, D. L. (1988) EMBO J. 7, 2185-2189
Moerman, D. G., and Waterston, R. H. (1989) in Mobile DNA (Berg, D. E., and Howe, M. M., eds) pp. 537-557, American Society for Microbiology, Washington D. C.
Rubin, G. M. (1983) in Mobile Genetic Elements (Shapiro, J. A., ed) pp. 329-361, Academic Press, New York
Rubin, G. M., and Spradling, A. C. (1982) Science 218, 348-353
Spradling, A. C., and Rubin, G. M. (1982) Science 218, 341-347
Xanthopoulos, K. G., Lee, J.-Y., Gan, R., Kockum, K., Faye, I., and Boman, H. G. (1988) Eur. J. Biochem. 172, 371-376
Young, M. W. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 6274-6278