On the specificity and effects on transcription of P-element insertions at the *yellow* locus of *Drosophila melanogaster*

Geoff Howes, Mark O'Connor and William Chia*

Department of Biochemistry and Unit of Molecular Genetics, University of Bristol, University Walk, Bristol BS8 1TD, UK

Received December 11, 1987; Revised and Accepted March 4, 1988

ABSTRACT

Fourteen P-element insertion mutants at the *yellow* (*y*) gene of *D. melanogaster* have been analysed by restriction mapping and DNA sequencing. The distribution of the insertion sites is nonrandom, with a preferred target site in the 5'-transcribed but non-translated leader of the gene. Transcription analysis indicates that P-elements inserted into the transcribed region of *y* can terminate *y* transcription. While the great majority of the P-element insertions into the preferred target site abolish *y* activity, this loss of activity is not due to disruption of the insertion site since imprecise excision events which do not restore the wild type sequence at the insertion site can restore *y* function. We further demonstrate that a mutant caused by a 221bp P-element insertion into the 5' transcribed but non-translated region, which show partial *y* activity, produces a larger mutant transcript consistent with it being a product of transcriptional readthrough. We present results which suggest that imprecise excision events which restore (or partially restore) *y* activity can occur by deleting P-element internal sequences which normally act to terminate *y* transcription.

INTRODUCTION

In *Drosophila melanogaster* there are at least three different transposable element families which can be mobilised by interstrain (dysgenic) crosses (1,2,3,4, 5). The P-element family is perhaps the best characterised member of this class. When males from P-strains are crossed to females from M-strains, P-elements are mobilised in the germline of the F1 progeny. One of the consequences of these dysgenic crosses is the generation of new mutations, many of which are caused by P-element insertions and P-mediated rearrangements (reviewed in Ref. 6). However the frequency with which P-element insertion mutations occur varies depending on the target gene.

The structure of the P-element family was described by O'Hare and Rubin (7). Its largest members are 2.9kb in size, encode a transposase responsible for the mobility of this family (8,9,10, 11), and are called

P-factors. The smaller P-elements, which are heterogeneous in size, are derived from the P-factors by internal deletion. The mobility of these elements requires, in addition to transposase in trans, the 31bp inverted repeat found at their ends (12). These elements produce a 8bp target site duplication and there appears to be some degree of target site specificity at the sequence level (7).

Recent studies of P-element insertion mutations at the Rudimentary (13), Rp11215 (14), and Notch (15) loci have raised two interesting points concerning the properties of these elements. Firstly, that these elements appear to insert preferentially near the transcriptional start sites of these genes and, at least in the case of the insertions at Notch, this preference cannot be attributed to the distribution of the sequences related to the P-element consensus target. Secondly, that P-elements may be more efficient at generating control mutations than traditional methods of mutagenesis (13). To extend on these observations we report here the characterisation of a set of PM hybrid dysgenesis induced mutations of the yellow locus and their revertants.

The y locus is required for the pigmentation of larval and adult cuticle structures. Mutant alleles of y can either abolish pigmentation in all of these cuticle structures (type1 alleles), or effect mosaic pigmentation patterns in which some cuticle structures are wild type while others are mutant (type2 alleles). Despite the phenotypic complexity of some of its mutants, the y locus contains a single gene which has previously been cloned (16,17,18, 19). The nucleotide sequence of this gene and the structure of the single 2.1kb transcript it encodes has been determined (19, 20). We have reported the preliminary characterisation of fourteen PM dysgenic mutants of y and a number of their revertants (19). There appears to be a minimum of four separable insertion sites at y; however there is a region located in the 5' leader within 30bp of the transcriptional start site where eight out of thirteen insertions map. Whilst the great majority of P-element insertions at y result in a type1 phenotype, we have reported an exceptional P-element insertion mutation which exhibit mosaic pigmentation pattern. Furthermore, it appears that phenotypically type1 alleles can revert to give both wild type and type2 revertants. In this paper we present data obtained from the structural mapping, transcriptional and DNA sequence analysis of these mutants and some of their revertants. These data, along with

the available information on the sequence and transcription of y, allow us to comment both on the specificity of P-element insertions and the effects of these insertions on the expression of y.

MATERIALS AND METHODS

Nucleic Acids Genomic DNA was prepared from adult flies as previously described (24). RNA was prepared essentially as described by Henikoff (25). Tissues collected from various developmental stages were frozen at -70C in 1.5 ml centrifuge tubes. Using a tight-fitting glass pestle, the samples were homogenized in 0.5 ml of 0.1M Tris:HCl (pH 7.5), 0.01M EDTA, 0.35M NaCl, 7M urea and 2% (w/v) SDS. Following two extractions with equal volumes of neutralized phenol/chloroform (1:1, v/v), nucleic acids in the aqueous phase were precipitated by the addition of two volumes of absolute ethanol. The nucleic acid pellets were twice washed in 75% ethanol, dissolved in sterile distilled water and stored at -70C.

Southern and Northern

RNA blots were performed essentially as described by Maniatis et al., (26). RNA samples were denatured in 2.2M formaldehyde, 5 mM Na acetate, 20mM morpholinopropanesulfonic acid, pH 7.0 and 50% formamide for 5 min at 65C. Samples were electrophoresed in agarose gels containing 2.2M formaldehyde, 20mM morpholinopropanesulfonic acid and 5 mM Na acetate and transferred to either nitrocellulose or nylon membranes. The filters were prehybridized in 3x SSC, 10x Denhardt's solution, 0.2 mg/ml of sonicated, denatured salmon sperm DNA, 0.1% SDS and 9% (w/v) dextran sulphate at 65C for at least 4h. Hybridisations were carried out using the same conditions, with labelled probe present at a concentration of between 5 to 10 ng/ml, for up to 48 h. Following hybridisation, the filters were washed 3x in 1x SSCP with 0.1% SDS and 3x in 0.1x SSCP with 0.1% SDS at 65C (30 min per wash). Autoradiography was performed at -70C with an intensifying screen. Southern hybridisation conditions were identical to those used for the Northern filters.

DNA sequencing and preparation of single-stranded hybridisation probes

Single-stranded templates were obtained by cloning the appropriate fragments into the vectors M13mp10 and M13mp11. The dideoxy chain termination method of Sanger et al. (27) was used for DNA sequence determination. The single-stranded probes used for the Northern hybridisations were synthesized by primer extending template prepared

Figure 1. Restriction maps of P-element insertions at *y*. (A) The restriction map of the wild type *y* region along with the positions of the P-element insertions described in this paper are shown on the top line. The positions of the various insertions are indicated above the line. The structure and orientation of the 2.1kb wild type *y* transcript is shown below the line. The deduced translational start lies 20bp to the left of the *Bam*HI cleavage site. (B) The restriction maps of the P-elements associated with the *y* alleles *hd1*, *hd2*, *hd4* and *hd14* were derived from genomic Southern mapping experiments. The maps of remaining P-element insertions were determined from cloned DNA. Abbreviations* B, *Bam*HI; Ba, *Bal*I; G, *Bgl*III; H, *Hind*III; O, *Xho*I; P, *Pst*I; R, *Eco*RI; S, *Sal*I; T, *Sst*I; V, *Pvu*II. The *Xho*I and *Pvu*II sites have not been mapped for all of the P-element insertions.

from M13 subclones containing appropriate restriction fragments from the y coding region. Details of this method have been previously described (28).

Construction of genomic banks

The phage vector EMBL4 was used for the isolation of y DNA from the insertion mutant strains described in this work. The y region from the mutants were initially mapped using genomic Southern experiments. For each of the mutants we cloned, an enzyme (EcoRI, SalI or BamHI) was chosen which did not cleave within the insertion. EcoRI, SalI, or BamHI fragment banks were generated and screened with y DNA as previously described (26). The relevant fragments from the phage recombinants containing y DNA were subcloned into the plasmid vectors pUC12 and pUC13 and further analysed.

RESULTS

P-element insertion sites at y

The genetic crosses used to generate the mutants studied here have been described elsewhere (19). The mutants alleles hd1 to hd14 are all phenotypically type1. The precise relationship between hd8-1 and hd-8 is not clear. However, we believe that hd8-1 arose spontaneously from a hd8 culture. The mutant hd8-1 is phenotypically a type2 allele. Physical maps were constructed for the mutants either by Southern blotting using cloned wild type y DNA as probe or by restriction mapping genomic y clones obtained from the mutant strains (see Figure 1. legend). Thirteen of the mutants contained insertions which were either shown to share homology to the P-factor clone p π 25.1 or had restriction maps which are consistent with those expected for P-elements. The size of these P-elements range from 0.22kb to 2.9kb and they are inserted in both orientations with respect to y. The mutant hd12 is a deletion (19) and hd3 appear to be an insertion which is not a P-element; the maps of these two mutants have not been included.

From the summary of the mapping data shown in Figure 1, it is evident that three of the mutants hd5, hd10, and hd11 contain insertions within the amino acid coding region of the gene; sequence analysis of clones obtained from these mutants indicate that the three mutants contain P-element insertions at three distinct insertion sites; in each case a 8bp target sequence is duplicated (Figure 2). The insertions associated with the two mutants, hd2 and hd4, map less than 100bp upstream of the

HD5	+274 CCAGCTGGACTT
HD10	+309 AGGACCAAGCTC
HD11	+3337 AATCCGTGCCCC
HD6	+20 CGGTCCACAGAA
HD7	+20 CGGTCCACAGAA
HD8-1	+14 CGTTACCGCGCC
HD8	+14 +20 CGTTACCGCG(5bp)GTCCACAGAA
HD14	-62 +20 TTTCATGGC(80bp)GTCCACAGAA

Figure 2. The DNA sequence of the P-element insertion sites. The DNA sequence of y has previously been published (19, 20). The deduced start of translation is at nucleotide +174. The positions of the insertions relative to the y transcription unit are indicated by the nucleotide numbers above the sequence. The 8bp target site duplications associated with six of the insertions are underlined. For hd8 and hd14, there is no target site duplication and the 8bp immediately flanking the insertions are underlined; in addition, the hd8 insertion is associated with a 5bp deletion of y sequence (nucleotides +15 to +19) and the hd14 insertion is associated with a 80bp deletion of y sequence (nucleotides -61 to +19). With the exception of hd8-1, the 31bp inverted terminal repeats are conserved in all of the P-element insertions.

BalI site which is located 4bp upstream from the start site for transcription. Their precise location has not been determined. However, the majority of mutants including the phenotypically type2 allele hd8-1 map just downstream of the BalI site, probably within the 5' transcribed but non-translated leader. To localise these insertion sites more precisely five of these alleles, hd6, hd7, hd8, hd14 and hd8-1, were cloned and the sequence of their insertion sites determined (Figure 2); the insertion sites of the remaining three alleles, hdl, hd9 and hd13, were more precisely mapped using the Southern transfer technique.

The P-elements associated with hd6 and hd7 are inserted at the same site and both insertions are flanked by the same 8bp duplication of the target site (spanning nucleotides +20 to +27). There are three additional

HD8-1

+7
CGTTACCGCGCATGATGAAATAATGAAATAATAATTTTAC

TCCAGTCACAGCTTTGCAGCAAAATTTGCAATATTTTCATT

TTTTTTTATTCACGTAAGGGTTAATGTTTTCAAAAAAA

ATTCGTCGCACACAACCTTTCCTCTCAACAAGCAACGT

GCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTTT

 +14
CGACGGGACCACCTTATGTTATTTTCATCATGTTACCGCGC

Figure 3. The nucleotide sequence of hd8-1. The 8bp target site duplication is underlined. P-element sequence is boxed. The position of the insertion relative to the y transcription unit is indicated.

alleles hd1, hd9 and hd13 containing P-elements which are inserted in the same orientation as the hd6 and hd7 insertions. Although these mutant alleles have not been cloned, it is clear that the insertion sites of hd1 and hd13 lie within a few bp of and could in fact be the same as the hd6 and hd7 insertion site since the HindIII/BamHI fragments (217bp) spanning the right hand junction of each of these insertions are indistinguishable in size. The HindIII/BamHI fragment spanning the right hand junction of the hd9 insertion is smaller, 207bp in size, possibly due to a more proximal insertion site.

The hd8 and hd14 insertions are unusual; neither insertion is flanked by a target site duplication; both insertions are associated with the deletion of wild type sequences immediately flanking the insertion site (Figure 2). The 8bp which constitute the target site of hd6 and hd7 also flank the right hand side of the hd8 and hd14 insertions. The 5bp

(spanning nucleotides +15 to +19) lying immediately to the left of the hd6 target site are deleted in hd8. The deletion associated with hd14 removes the 80bp sequence spanning nucleotides -61 to +19. It is unlikely that these deletions are generated as a result of cloning; Southern hybridisation experiments using total genomic DNA from hd14 confirms the absence of the BalI cleavage site (located at nucleotide -4) which is removed by the hd14 deletion (data not shown). Since all of the P-element insertion sites sequenced to date are flanked by target site duplications, it is unlikely that our hd8 and hd14 strains represent the original mutational event; one likely explanation for these observations is that the four mutants hd6, hd7, hd8 and hd14 were all originally caused by P-element insertions into the same target site; and, our hd8 and hd14 strains resulted from P-element mediated unidirectional deletions which removed sequences lying to the left of the insertion site, including one of the two duplicated target sites (see Discussion). However an alternative explanation for the generation of the hd8 and hd14 insertion sites is also possible; for example in the case of hd8, the element could have inserted at exactly the same site as the hd8-1 element, duplicating TTACCGCG, and then lost 13bp (one copy of the 8bp target site duplication plus 5bp of y) to the right of the insertion.

Of all the dysgenically induced mutants analysed here, hd8-1 is the only one which is phenotypically type2. It is caused by the insertion of a 221bp P-element; this insertion is flanked by an 8bp target site duplication (spanning nucleotides +7 to +14). Since this element is inserted into the 5' non-translated leader and the other seven P-element insertions into the same region are phenotypically type1, we decided to examine this allele in greater detail.

The effect of P-element insertions into the 5 non-translated leader on y transcription

The nucleotide sequence of the hd8-1 insertion and the y sequences flanking it are shown in Figure 3. Previous studies of P-element insertions into the coding region of the white (w) gene have shown that when a P-element is inserted such that its transcription is in the same direction as that of the target gene, the transcription of the target gene terminates within the insertion (21). Since hd8 is inserted 14bp downstream from the start of transcription, our result which indicate that hd8 does not accumulate detectable amounts of mutant y transcript

Figure 4. Mutant *y* transcripts produced by *hd8-1* and *hd11*. (A) The positions of the insertions are indicated above the restriction map. The single stranded probes used for these experiments are represented by the solid arrows above the restriction map. The 2.1kb wild type *y* transcript is shown below the map. (B) 10 μ g of total RNA prepared from embryonic (track1), first instar larval (track2) and second instar larval (track3) stages of *hd8-1* and the pupal stage of wild type *Drosophila* was fractionated on a 0.8 agarose/formaldehyde gel and transferred to nylon membrane. The filter was hybridized to probeII to produce the autoradiogram. (C) 10 μ g of pupal total RNA prepared from *hd11* (track1) and wild type (track2) was fractionated as above, transferred to a nylon membrane and hybridized to probeI. (D) A filter identical to that used in panel C hybridized to probeII.

is consistent with the above finding (data not shown). The insertion of *hd8-1* is, however, in the opposite orientation to *y*; furthermore, since it is only 221bp and does not contain a putative polyadenylation signal, we considered it likely that *y* transcription can read through

the inserted sequences thereby producing a larger transcript with an extended 5' non-transcribed leader. The results of Northern hybridisation experiments using total RNA prepared from various developmental stages are consistent with this possibility. The data (shown in Figure 4b) indicate the presence of a transcript which is about 200 bases larger than the 2.1kb wild type y transcript.

With the exception of hd8-1, all of the other P-element insertions into the 5' non-translated leader are phenotypically typel (null) regardless of the orientation of the inserted element. This suggests that even insertions in which the directions of transcription of the element and y are divergent can abolish the expression of y. There are several possible ways in which the loss of y function can be effected. Firstly, it is possible that P-elements, even inserted in the "divergent" orientation, can terminate y transcription; this possibility is at least consistent with the sequence data on the P-factor P*25.1 which indicate the presence of putative polyadenylation signals on both strands (7); the second possibility is that since the majority of the typel insertions have probably inserted into the same insertion site, the loss of function may be due to the insertional inactivation of a necessary cis acting control element. However, since these insertions are located close to the start of transcription, it is difficult to ascertain whether they terminate y transcription because the putative mutant transcripts would share little homology to y. To answer the question of whether P-element insertions which are oriented in the "divergent" orientation to y can terminate y transcription, we analysed RNA from hd11 which contains a 2.9kb P-factor inserted in the appropriate orientation within the coding region of y. The results (shown in Figure 4C) demonstrate the presence of a mutant transcript (approx. 1.5kb long) when the Northern filter is probed with a probe located upstream of the insertion; however when an identical filter is probed with a probe lying within the transcribed region but downstream of the insertion, no homology was observed (Figure 4d). These results indicate that P-elements inserted such that its direction of transcription is opposite to that of the target gene can potentially terminate the transcription of the target gene.

In order to test the possibility that the typel phenotype associated with the great majority of the P-element insertions into the non-translated leader is due to the insertional inactivation of a control sequence necessary for expression of y, we destabilised hd8 using dysgenic crosses,

Figure 5. Nucleotide sequence of hd8 and its revertants. P-element sequences are boxed. The sequences internal to the hd8 insertion, not shown, are indicated by dots. The 5bp, nucleotides +15 to +19, from the transcribed region of y are absent in hd8 and its revertants. The arrows above the nucleotide sequence of the revertants indicate the point of discontinuity where sequences from the left hand terminal repeat are joined to those of the right hand repeat. For rev19 the breakpoint can be at three possible positions. For rev17, the C beneath the arrow lies at the breakpoint but cannot be accounted for.

obtained several revertants which either totally or partially restored the wild type phenotype, cloned the y sequences from these revertants and analysed these clones by restriction mapping and DNA sequencing. All four reversion events are due to the imprecise excision of the P-element of hd8 leaving part of its 31bp inverted repeat at the original insertion site. As expected, the 5bp deletion associated with the hd8 insertion remain absent in these revertants, and like their parental insertion, the P-element sequences remaining in these revertants are not flanked by 8bp duplications (see Figure 5). Rev6 retained 29bp of P-element sequence in which the 15bp from the left end of the hd8 P-element is joined to the 14bp at its right end. Rev17 and rev19 can similarly be explained by one break in each of the inverted repeats; rev17 retained 35bp of P-element sequence and rev19 retained 32bp of P-element sequence. Since all of these imprecise excision events restore, at least in part, the wild type phenotype, it is clear that the integrity of the insertion site, per se, is not necessary for y expression.

DISCUSSION

Specificity of P-element insertions at y

P-element insertions have been extensively analysed for a number of Drosophila genes. The specificity of P-element insertions appear

to operate at three different levels, at the choice of the target gene, the position within the target gene, and the DNA sequence of the target site. That the frequency of P-element insertion mutations vary widely for different genetic loci is well documented (reviewed in Ref.6). Furthermore Kelley *et al.* (15) noted that the distribution of insertion mutations at N (13/14), r (5/5) and RpII215 (7/9) show a distinct preference for insertion sites located in close proximity to the transcriptional start sites of these genes. Since all of these genes are probably active in germline cells where the mobilisation of P-elements occur, these authors suggest a possible correlation between the non random nature of P-element insertion sites with the structural changes which accompany gene activation. y encodes a function required for the pigmentation of cuticle structures. Its expression coincides with developmental periods during which pigmentation of the epidermal derived structures occur (22). Therefore the clustering of P-element insertions near the transcriptional start site of y (10/13) appear anomalous in this context. Sequences which show 6/8 or 5/8 matches (there are no 7/8 or 8/8 matches) to the 8bp P-element consensus target site (7) show an essentially random distribution within the published y sequence (19,20). Hence it is clear that the clustering of P-element insertion sites near the transcriptional start site of y is not due to the non random distribution of sequences related to the P-element consensus target site.

The effect of P-element insertions on y expression

Our data suggest that P-elements can terminate transcription of y when inserted in either orientation into the transcribed region of the gene. This effect is presumably dependent on sequences located within the P-factor and only P-elements in which these sequence(s) have not been deleted will show this effect. Hence, small P-elements like the hd8-1 insertion which do not contain putative polyadenylation signals will not terminate y transcription; as expected RNA prepared from this strain contain a mutant transcript of a size consistent with it being the product of transcriptional readthrough. Since its insertion does not disrupt the amino acid coding region of y, it is not surprising that this mutant shows y activity, although not at the wild type level. The fact that phenotypically type1 mutations caused by insertion into the 5' non-translated region can revert via imprecise excision events to restore wild type or partial y expression may be attributed to the deletion of these portions of the P-element sequence which causes transcriptional

termination. Our sequence analysis of revertants obtained from the hd8 insertion which indicate the retention of between 29bp to 35bp of sequence from the inverted terminal repeats supports this assertion.

The nature of the sequences retained in the hd8 revertants are very similar to those described for revertants of r (13) and RP11215 (14). While it is obvious that only genetic events which will restore y function will be recognised, this factor alone cannot account for the restricted nature of the sequences retained in the revertants. Because we know that insertions as large as the hd8-1 insertion (221bp) fail to abolish y activity, there is no reason, based on phenotypic considerations, why revertants of hd8 should not be structurally similar to hd8-1. The fact that all the revertants retain at most 35bp of terminal repeat sequence suggest that the 31bp inverted repeats act as hot spots for internal deletion breakpoints.

ACKNOWLEDGEMENTS

We thank the SERC for financial support.

*To whom correspondence should be addressed

REFERENCES

1. Rubin, G.M., Kidwell, M.G. and Bingham, P.M. (1982). *Cell* **29**, 987-994.
2. Bucheton, A., Lavigne, J.M., Picard, G. and L'Heritier, P. (1976). *Heredity* **36**, 305-314.
3. Bucheton, A., Paro, R., Sang, H.M., Pelisson, A. and Finnegan, A.J. (1984). *Cell* **38**, 153-164.
4. Yannopoulos, G., Stamatis, N., Monostiriotti, M., Hatzopoulos, P. and Louis, C. (1987). *Cell* **49**, 487-495.
5. Blackman, R.K., Grimaila, R., Koehler, M. and Gelbart, W.M. (1987). *Cell* **49**, 497-505.
6. Engels, W.R. (1983) *Ann. Rev. Genet.* **17**, 315-344.
7. O'Hare, K. and Rubin, G.M. (1983). *Cell* **34**, 25-36.
8. Spradling, A.C. and Rubin, G.M. (1982). *Science* **218**, 341-347.
9. Rubin, G.M. and Spradling, A.C. (1982). *Science* **218**, 348-353.
10. Laski, F.A., Rio, D.C. and Rubin, G.M. (1986). *Cell* **44**, 7-19.
11. Rio, D.C., Laski, F.A. and Rubin, G.M. (1986). *Cell* **44**, 21-32.
12. Karess, R.E. and Rubin, G.M. (1984). *Cell* **38**, 138-146.
13. Tsubota, S. and Schedl, P. (1986). *Genetics* **114**, 165-182.
14. Searles, L.L., Greenleaf, A.L., Kemp, W.E. and Voelker, R.A. (1986) *Mol. Cell. Biol.* **6**, 3312-3319.
15. Kelley, M.R., Kidd, S., Berg, R.L. and Young, M.W. (1987). *Mol. Cell. Biol.* **7**, 1545-1548.
16. Campuzano, S., Carramolino, L., Cabrera, C., Ruiz-Gomez, M., Villares, R., Boronat, A. and Modolell, J. (1985). *Cell* **40**, 327-338.
17. Biessman, H. (1985). *Proc. Natl. Acad. Sci. USA.* **82**, 7369-7373.

18. Parkhurst, S.M. and Corces, V. (1986). *Mol. Cell. Biol.* 6, 47-53.
19. Chia, W., Howes, G., Martin, M., Meng, Y.B., Moses, K. and Tsubota S. (1986). *EMBO J.* 5, 3597-3605.
20. Geyer, P.K., Spana, C. and Corces, V.G. (1986). *EMBO J.* 5, 2657-2662.
21. Levis, R., O'Hare, K. and Rubin, G.M. (1984). *Cell* 38, 471-481.
22. Nash, W.G. (1976). *Devel. Biol.* 48, 336-343.
23. Grindley, N.D.F. and Reed, R.R. (1985). *Annual Rev. Biochem.* 54, 863-896.
24. Chia, W., Karp, R., McGill, S. and Ashburner, M. (1985). *J. Mol. Biol.* 186, 689-706.
25. Henikoff, S. (1983). *Nucl. Acids. Res.* 11, 4735-4752.
26. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
27. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
28. Burke, J.F. (1984). *Gene* 30, 63-68.