Transcription Initiation From Within P Elements Generates Hypomorphic Mutations in Drosophila melanogaster

Matthew C. LaFave* and Jeff Sekelsky*†

*Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, North Carolina 27599 and  †Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599

ABSTRACT Numerous transposable element insertions in Drosophila melanogaster cause hypomorphic mutations. We report that transcription initiation within a region found in many P-element constructs provides an explanation for why some gene function is retained. We detected evidence of this transcription in four different types of P constructs, regardless of whether the insertion was in a coding exon, intron, 5’ untranslated region, or upstream of the gene span.

Results and Discussion

Genetically engineered P elements have been used as tools for experimental manipulation of the Drosophila genome (Bellen et al. 2004; Thibault et al. 2004). These elements tend to insert near transcription start sites, frequently resulting in hypomorphic mutations (Spradling et al. 1995). This result is sometimes surprising, as insertions within early exons might be expected to produce null mutations rather than hypomorphs. Here, we report that transcription initiation from within a P element may be responsible for some mutations being hypomorphic rather than null.

We first observed this phenomenon in a P insertion allele of nbs. The nbsEY15506 allele contains a 10.9-kb P{EPgy2} element inserted into the coding sequence in the second exon. Although the insertion separates two conserved domains, it creates a hypomorphic, separation-of-function allele, rather than a null allele. The nbsSM9 derivative has an internal deletion of ∼7 kb of this P element, but retains the hypomorphic nature of nbsEY15506. We previously reported that nbsSM9 has a longer transcript than wild-type nbs on a Northern blot and that the hypomorphic character of nbsSM9 was due, at least in part, to transcription that initiated from

within the P element (Figure 1) (Mukherjee et al. 2009). Using 5’ RACE, we found that transcription initiated in the region downstream of the 3′ end of the white (w) gene carried on the P element—a region that is located 3′ to the endogenous w locus, but is not part of the annotated w gene span. In addition, we detected that three introns, with canonical splice donor and acceptor sequences, had been spliced out (Figure 1, B and C). Two of these introns are located in the region downstream of w, but the third is within the P-element 5′ end, which is transcribed in the direction opposite to P transposase transcription. The terminal 3 bp of the P end form a start codon that is in-frame with the nbs-coding sequence. Results from our genetic assays suggested that nbsSM9 does produce an N-terminally truncated protein that lacks the FHA domain, resulting in loss of some NBS functions but retention of others.

The region from which this intra-P transcription appears to initiate corresponds with the genomic location X:2,683,995…2,684,631, directly upstream of the gene CG32795 (D. melanogaster genome release 5.34; Tweedie et al. 2009). The proximity to CG32795 led us to investigate the possibility of promoter elements in this region. There are no EST cDNAs in GenBank for this region and no annotated exons; nevertheless, there is evidence for transcriptional activity. Importantly, transcripts from the region between w and the first annotated exon of CG32795 have been detected in white pre-pupae and adult male flies (Celniker et al. 2009). On the basis of the position of RNA-seq reads in the modENCODE genome browser, these transcripts appear

Copyright © 2011 by the Genetics Society of America
doi: 10.1534/genetics.111.129825
Manuscript received March 4, 2011; accepted for publication April 21, 2011
Sequence data from this article have been deposited with the GenBank Data Libraries under accession nos. JF834162–JF834166.
†Corresponding author: Department of Biology, CB #3280, University of North Carolina, Chapel Hill, NC 27599. E-mail: sekelsky@unc.edu
to contain exons in the same relative position as those within the P-element transcripts. A notable difference is that the downstream intron of the endogenous transcript is longer and appears to splice to the second exon in the 5' downstream intron of the endogenous transcript is longer and produces both Slbp transcript and SLBP protein, as determined by in situ hybridization to mRNA and Western blot, respectively (Sullivan et al. 2001).

We performed PCR on cDNA from Slbp10/TM6B adults and second instar larvae and found that there is also transcription from the region downstream of w in this P(Epy) element. In adults, the sequence and splice sites are identical to those that we detected in nbs5M9 (Figure 1, B and C; Figure 2B). Intriguingly, splicing of the intron in the P element was not detected in cDNA isolated from larvae, raising the possibility that splicing of the intron-P transcript may be regulated by tissue and/or developmental timing.

P elements also frequently insert into introns (Spradling et al. 1995). To determine whether transcription from the region downstream of w can contribute to the mutagenic properties of intronic insertions, we analyzed an insertion of a P(SUPOr-P) element into the first intron of GlcAT-S (Figure 2). By performing PCR on cDNA obtained from GlcAT-S5GO1446 adults, we found that transcripts were produced in the region downstream of the P-element w gene, but with a slightly different splicing pattern than in the two previous examples. The first intron was spliced out as described above, but the second intron was found to splice to the next downstream exon of GlcAT-S (Figure 1B and Figure 2B). This may be due to the fact that the P element was in an intron, or it
could be related to the su(Hw) binding region between the end of the w span and the 5' P end of P(SUPor-P).

We also tested a P element inserted into an intergenic region to determine if it could promote transcription of a nearby gene. We examined P(wHy)DG30201, which is inserted upstream of CG9775 (Figure 2A). PCR on cDNA from P(wHy)DG30201 adult flies revealed a situation similar to that of GlcAT-SG001446 and the endogenous w locus. Transcription initiated in the region downstream of w, and the first intron was spliced out. The second intron spliced into the first splice acceptor sequence of the first exon of the 5' UTR of CG9775, 23 bp downstream of the annotated transcription start site. This exon can be found in both transcript CG9775-RA and transcript CG9775-RD (Figure 2B). Thus, this P(wHy) insertion is responsible for additional transcription of a downstream gene.

These findings provide insight into the nature of hypomorphic alleles that arise from insertion of certain engineered transposable elements. For example, there is no guarantee that transcription from within the transposable element will match the expression pattern of the gene's native promoter. If this is the case, altered expression may result in different severities, potentially ranging from wild-type function in some tissues to complete absence of function in other tissues. It is interesting to note that there are two potential ORF-initiating start codons at the end of the 5' P end: the final three bases, and another that begins 13 bp upstream (Figure 1C). As the resulting ORFs are in different reading frames, one would expect as many as 2/3 of correctly oriented w-containing P elements in exons to produce at least a partial protein product.

Features of these cryptic transcripts that affect translation efficiency may play a role in the reduced function as well. While Slbp10 mutants produce SLBP protein, they do so at reduced levels (Sullivan et al. 2001). This may be due to characteristics of the cryptic transcript. Among the exons on the P-element–Slbp transcript, the exon in which translation begins—a fusion of the 5' P end and the remaining sequence of the 5' UTR and start of the coding region of Slbp—is the only exon that does not contain a short open reading frame. This suggests that translation can successfully take place only when ribosomes bind to that key exon, as binding to the upstream exons would lead to premature termination of translation and potentially to nonsense-mediated decay (NMD) (Peltz et al. 1993; Gatfield et al. 2003). This feature may contribute to the hypomorphic character of some mutants with intra-P transcription: even if transcripts were being produced at the normal rate, transcript degradation may ensure that not all transcripts would survive to be translated, leading to a reduced level of protein. The cryptic transcript of nbsSM9 does not share this characteristic. Rather, we found that levels of the nbsSM9 transcript appeared to be similar to that of wild type (Mukherjee et al. 2009). This suggests that NMD may not be involved in reducing the levels of nbs transcript. This is curious, as the P elements in Slbp10 and nbsSM9 have identical sequence in the region of the cryptic transcript. This suggests that the stability of these transcripts is affected by additional factors that are not immediately apparent.

The position and orientation of a P element insertion may have additional ramifications for mutagenesis. Insertions in introns may invoke alternative splicing of the native transcript, possibly resulting in impaired functionality of the transcript or protein. Additionally, it may be that insertions that take place in a gene span, but which transcribe the region downstream of w in the opposite direction as the gene does, may produce antisense RNA. These transcripts could initiate RNA interference of the native transcript, providing another means of producing a hypomorph.

The w gene is a common marker in engineered transposable elements in Drosophila. We have demonstrated that the region downstream of w can transcribe into the gene into which the P element is inserted or is near. The examples presented here encompass four different P-element constructs in four different gene regions (coding sequence, 5' untranslated region, intron, and upstream). Thus, this is likely to be a common effect, and transcription initiation from within these elements is likely to contribute to the hypomorphic mutations generated by many of these insertion alleles.

**Acknowledgments**

We thank Deirdre Tatomer for providing Slbp mutant flies; the laboratory of Lillie Searles for assistance with RNA anal-
ysis; and Bob Duronio, William Marzluff, and members of the Sekelsky and Duronio labs for helpful discussions of the manuscript. This work was supported by a grant from the National Institutes of Health (GM061252) to J.S.

**Literature Cited**


Mukherjee, S., M. C. LaFave, and J. Sekelsky, 2009 DNA damage responses in Drosophila nbs mutants with reduced or altered NBS function. DNA Repair (Amst.) 8: 803–812.


*Communicating editor: J. A. Birchler*