Benchmarks

Identifying Mutations in *Drosophila* Genes by Direct Sequencing of PCR Products

BioTechniques 27:262-264 (August 1999)

Determining the sequence of a mutant gene that causes an abnormal phenotype can contribute greatly to understanding the normal biochemical function of the gene. In the fruit fly *Drosophila melanogaster*, mutations in individual alleles of a given gene have traditionally been identified by: (i) sequencing of clones isolated from phage libraries constructed using genomic DNA from mutant flies (5), (ii) sequencing of cloned polymerase chain reaction (PCR) products generated using genomic DNA from mutant flies (3) and (iii) sequencing of cloned reverse transcription (RT)-PCR products generated using mRNA from mutant flies (2). In the case of homozygous lethal mutations, in which the only source of genomic DNA or mRNA is heterozygous animals, the identification of mutations by these procedures is complicated by the presence of the wild-type (WT) gene. To distinguish clones derived from the mutant as opposed to the WT gene copy, silent sequence polymorphisms between the WT and mutant genes must first be identified. Furthermore, in the PCR-based approaches, multiple clones generated in independent reactions need to be sequenced to eliminate the possibility that identified mutations are merely errors introduced during the PCR. Both of these complications can be circumvented by direct sequencing of PCR products generated from genomic DNA, a procedure that is commonly used for other metazoan organisms (4).

We have adapted this technology to *Drosophila* and here describe a simple, efficient and accurate method for identifying point mutations, deletions or insertions within genes. Figure 1 presents examples of a point mutation and an 8-bp deletion identified using this method. Direct sequencing of a 1:1 (WT:mutant) mixed population of PCR products, using a Model 377 Automated Sequencer (PE Biosystems, Foster

City, CA, USA), reveals mixed sequence in electropherograms at positions that are polymorphic between the WT and mutant genes. Point mutations result in mixed sequence at a single position (compare Figure 1, Panel A with Panels B or C), whereas small insertions and deletions (i.e., <25 bp, which cannot be resolved by agarose gel electrophoresis as different sized PCR products) result in a region of mixed sequence that begins at the point of insertion or deletion and continues to the end of the sequence (compare Figure 1, Panel D with Panel E). This method for identifying sequence mutations in a specific *Drosophila* gene involves four steps, described in detail below.

Designing oligonucleotide primers.

Primers that are designed and synthesized based on the following criteria generally work well for both the PCR amplification and sequencing steps of the method: (i) primers are desalted and are 20 nucleotides in length; (ii) primers contain 50%–75% G or C nucleotides, including a G or C nucleotide at the 3' end; (iii) primers are positioned to generate PCR products that are <400 bp, which results in efficient amplification and makes it possible to use the same primers for complete sequencing of both strands of the PCR product; and (iv) since the first 40 bp of sequence are often difficult to read, primers are positioned 40 bp upstream or downstream of the site where sequence initiation is desired (for example, primers should be placed within introns 40 bp downstream of 5' splice sites or upstream of 3' splice sites).

Isolating genomic DNA from heterozygous mutant flies. Use a disposable plastic micropestle (Research Products International, Mt. Prospect, IL, USA) to homogenize approximately 50 flies in a microcentrifuge tube containing 500 μ L of buffer (0.1 M Tris-HCL, pH 8.5, 0.1 M EDTA, 1% sodium dodecyl sulfate [SDS]). Incubate homogenate at 70°C for 20–30 min, add 70 μ L of 8 M potassium acetate and place on ice for 30 min. Remove insoluble material by centrifugation at 16 000 $\times$ g for 15 min, and transfer the supernatant to a clean tube. Repeat this step until no pellet is visible. Extract the supernatant with 500 μ L of phenol/chloroform/isoamyl alcohol (PCA) 25:24:1,

transfer the aqueous layer to a clean tube and precipitate the DNA and RNA by adding 250 μ L of isopropanol and centrifuging at 16 000 $\times$ g for 10 min at room temperature. Wash the pellet with 70% ethanol and resuspend the pellet in 200 μ L of TE (10 mM Tris-HCL, pH 8.0, 1 mM EDTA). Remove the RNA by incubation with 1 μ L of 10 mg/mL RNaseA at 37°C for 30 min, followed by PCA extraction and ethanol precipitation. Resuspend the DNA in 50 μ L of H₂O and determine the concentration by spectrophotometry at optical density (OD)₂₆₀. Approximately 1–1.5 μ g of genomic DNA should be recovered per fly. This protocol is a modified version of one developed for DNA extraction from single flies (1).

Amplifying coding exons by PCR.

Generate PCR products by combining 0.5–1.0 μ g of genomic DNA, 25 pmol of each primer, 5.0 μ L of 10 $\times$ PCR buffer (Life Technologies, Gaithersburg, MD, USA), 4.0 μ L of 2.5 mM dNTPs, 1.5 μ L of 50 mM MgCl₂, 0.25 μ L of *Taq* DNA Polymerase (Life Technologies) and H₂O up to 50 μ L in a PCR tube. Incubate reactions in a PTC-100™ Thermal Cycler (MJ Research, Watertown, MA, USA) for 1 cycle (5 min at 94°C), 30 cycles (1 min at 94°C, 1 min at 50°C and 1 min at 72°C) and 1 cycle (5 min at 72°C). Resolve PCR products (10–15 μ L of a reaction) by electrophoresis on a 1% agarose gel. Excise bands of the expected size from the gel, and extract the DNA using a QIAEX® II Gel Extraction Kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocols. In the final step, elute the DNA with 20 μ L of H₂O.

Sequencing of PCR products. Sequencing of PCR products using dRhodamine or BigDye™ Terminator Cycle Sequencing Ready Reactions (PE Biosystems) was performed according to the manufacturer's protocols with the following modifications, which improved signal intensity and read length. Use 8 μ L of eluted DNA template (ca. 300–600 ng) and 33–120 pmol of primer per reaction. Note that the BigDye reaction solution does not work as well when diluted.

To illustrate the utility of the method, we have used it to identify mutations in five alleles of the *Drosophila* *TAF110*

Benchmarks

gene. *TAF110* has a coding region of 2763 bp that spans 9 exons. Twelve sets of primers (24 primers) were needed to generate PCR products of <400 bp that included all of the *TAF110* coding regions. Thus, sequencing of each *TAF110* allele required only 12 PCRs and 24 sequencing reactions. By sequencing multiple *TAF110* alleles, each allele served as a control for the other alleles. Natural sequence variations were attributed to polymorphisms present in all of the alleles, while genuine mutations were attributed to polymorphisms unique to individual alleles. However, if only a single mutant allele is available, comparison to sequences determined for multiple WT strains, such as *CantonS* and *OregonR*, should serve as an appropriate control.

Because each strand of a PCR product can be completely sequenced using the same primers used to generate the PCR product, polymorphisms were confirmed as being present on both strands, as illustrated in Figure 1, B and C. In cases in which insertions or deletions are large enough to result in PCR

Figure 1. Examples of electropherograms containing sequence polymorphisms detected by direct sequencing of PCR products amplified using genomic DNA from heterozygous adult flies. (A–C) The sequence of a region of the *TAF110* gene from (A) WT or (B and C) *TAF110^{S-466}* mutant flies. A point mutation, indicated by an arrow, is revealed upon sequencing of both the sense (B) and antisense (C) strands of the PCR product. We have observed that in most cases, peak sizes at polymorphic bases are not equivalent, but both peaks are clearly above background. Furthermore, in most cases, the sequence analysis software does not call an ambiguous base (i.e., an N) at these polymorphic positions, so electropherograms must be visually analyzed after the run. (D and E) The sequence of a region of the *TAF110* gene from (D) WT or (E) *TAF110^{XS-1101}* mutant flies. A deletion is revealed by a stretch of mixed sequence that initiates abruptly, as indicated by the arrow. It is interesting to note that the sequence of the WT and the mutant gene can be made out clearly within the region of mixed sequence. The 3' end of the deletion was determined by sequencing the antisense strand of the PCR product (data not shown).

products that can be resolved from the WT gene product by agarose gel electrophoresis, the nature of the polymorphism can be determined by comparing the sequence of the aberrant and normal PCR products.

The application to *Drosophila* of a direct sequencing method for identifying mutations in heterozygous animals is a significant advancement over traditional methods described above (2,3,5). With the completion of the *Drosophila* genome sequence in the coming year, it will be possible to determine the intron/exon structure of all *Drosophila* genes and thus make this method for identifying sequence mutations applicable to alleles of any gene.

REFERENCES

1. Ashburner, M. 1989. Preparation of DNA from single flies, p. 108-109. In M. Ashburner (Ed.), *Drosophila*, A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.
2. Brunner, D., N. Oellers, J. Szabad, W.H. Biggs, S.L. Zipursky and E. Hafen. 1994. A gain-of-function mutation in *Drosophila* MAP kinase activates multiple receptor tyrosine kinase signaling pathways. *Cell* 76:875-888.
3. Chang, H.C., N.M. Solomon, D.A. Wassarman, F.D. Karim, M. Therrien, G.M. Rubin and T. Wolff. 1995. Phyllopod functions in the fate determination of a subset of photoreceptors in *Drosophila*. *Cell* 80:463-472.
4. Mori, A., A. Takagi, Y. Ikeda and A. Yamamoto. 1997. Development of a direct DNA sequencing method for detecting heterozygous mutations of the human lipoprotein lipase gene. *Clin. Biochem.* 30:315-324.
5. Therrien, M., H.C. Chang, N.M. Solomon, F.D. Karim, D.A. Wassarman and G.M. Rubin. 1995. KSR, a novel protein kinase required for Ras signal transduction. *Cell* 83:879-888.

Address correspondence to Erin M. Schlag, *Cell Biology and Metabolism Br., National Institute of Child Health and Human Dev., Bldg. 18T, Rm. 101, National Institutes of Health, Bethesda, MD 20892, USA. Internet: eschlag@helix.nih.gov*

Received 31 March 1999; accepted 26 May 1999.

Erin M. Schlag and David A. Wassarman
National Institute of Child Health and Human Development, NIH Bethesda, MD, USA