

# Improved detection of small deletions in complex pools of DNA

Mark Edgley, Anil D'Souza<sup>1</sup>, Gary Moulder<sup>1</sup>, Sheldon McKay, Bin Shen, Erin Gilchrist, Donald Moerman and Robert Barstead<sup>1,\*</sup>

Biotechnology Laboratory and Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada and <sup>1</sup>Department of Molecular and Cell Biology, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, Oklahoma City, OK 73104, USA

Received December 14, 2001; Revised and Accepted April 16, 2002

## ABSTRACT

About 40% of the genes in the nematode *Caenorhabditis elegans* have homologs in humans. Based on the history of this model system, it is clear that the application of genetic methods to the study of this set of genes would provide important clues to their function in humans. To facilitate such genetic studies, we are engaged in a project to derive deletion alleles in every gene in this set. Our standard methods make use of nested PCR to hunt for animals in mutagenized populations that carry deletions at a given locus. The deletion bearing animals exist initially in mixed populations where the majority of the animals are wild type at the target. Therefore, the production of the PCR fragment representing the deletion allele competes with the production of the wild type fragment. The size of the deletion fragment relative to wild type determines whether it can compete to a level where it can be detected above the background. Using our standard conditions, we have found that when the deletion is <600 bp, the deletion fragment does not compete effectively with the production of the wild type fragment in PCR. Therefore, although our standard methods work well to detect mutants with deletions >600 bp, they do not work well to detect mutants with smaller deletions. Here we report a new strategy to detect small deletion alleles in complex DNA pools. Our new strategy is a modification of our standard PCR based screens. In the first round of the nested PCR, we include a third PCR primer between the two external primers. The presence of this third primer leads to the production of three fragments from wild type DNA. We configure the system so that two of these three fragments cannot serve as a template in the second round of the nested PCR. The addition of this third primer, therefore, handicaps the amplification from wild type template. On the other hand, the amplification of mutant fragments where the binding site for the third primer is deleted is unabated. Overall, we see at least a 500-fold increase in the sensitivity for small deletion

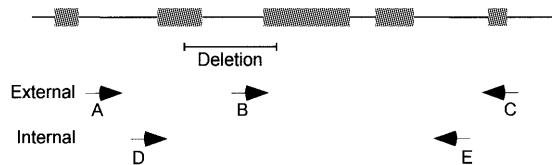
fragments using our new method. Using this new method, we report the recovery of new deletion alleles within 12 *C.elegans* genes.

## INTRODUCTION

Until recently, most genes in the nematode *Caenorhabditis elegans* were known through their mutant phenotypes. Now, however, more *C.elegans* genes are known through sequencing than through genetics (1); of the approximately 19 000 genes in *C.elegans*, only about 1300 have mutant alleles. As genetics is one of the most important tools in the arsenal of *C.elegans* biologists, we and others have devised methods to derive mutations in *C.elegans* genes known only through their sequence (2–6). Our present methods are conceptually simple. We treat worms at the L4 larval stage with a mutagen. The progeny of the mutagenized worms are subdivided into populations that are allowed to reproduce. We then extract the DNA from ~30% of each population. The extracted DNA samples are pooled and subjected to PCR with nested primer sets (Fig. 1). Candidate populations are identified by the presence of a PCR product that is smaller than the size predicted by the genomic DNA sequence. Each candidate population is subdivided and subjected to similar growth and PCR analysis. This process of sib selection continues until we recover a single individual with the deletion. Using this protocol, typically we can recover such an individual in three steps of growth and sib selection.

In principle, our ability to detect a deletion should be a function of the resolution of the gels used for electrophoresis. In practice, however, gel resolution does not determine the size of the deletions that we identify; typically, we do not detect deletions that are less than ~600 bp. This limitation is a consequence of our effort to reduce, as far as possible, the cost and time required to identify a candidate deletion. To accommodate the relatively low frequency at which deletions are induced by chemical mutagens (7), we pool the DNA template samples to minimize the numbers of PCRs that are needed to detect a deletion. Using our present DNA pooling strategy, for every copy of the deletion DNA we have approximately 12 000 copies of wild type DNA. This places the mutant template at a substantial disadvantage in PCR. Deletion fragments that are close to the size of wild type DNA apparently are not able to overcome this initial disadvantage.

\*To whom correspondence should be addressed. Tel: +1 405 271 1766; Fax: +1 405 271 3153; Email: barsteadr@omrf.ouhsc.edu



**Figure 1.** Protocol schematic. Our current protocols use nested PCR to identify smaller than normal PCR fragments derived from a mutagenized population of worms. Typically, we use two external (A and C) and two internal PCR primers (D and E). As shown, the poison primer method adds a primer (B) in the first round of the nested PCR.

However, recent work has shown that a significant proportion of deletions induced by trimethylpsoralen (TMP) treatment followed by UV irradiation are <600 bp, falling in the range of 50–600 bp (E.Gilchrist, M.Edgley, G.Mullen, B.Shen, T.Rogalski, T.Szczygielski and D.Moerman, manuscript in preparation) (8). Based on these data, it is likely that we do not detect a significant number of TMP-induced deletion mutants using our standard screening strategy. The challenge then became to develop methods that would allow us to detect small deletions. This paper reports such a method.

## MATERIALS AND METHODS

### Reconstruction experiments

Our initial tests were done with a known 133 bp deletion in the *C.elegans* gene *dim-1(gk54)* (GenBank locus U39667). We used the following oligonucleotide primers in nested PCR as shown in Figure 1: first round, 5'-CTCAGTCGATCACAGT-ACA-3', 5'-TCCACCAACAAGCTTTTGCC-3'; second round, 5'-ACACTTCCCACAACAACCAG-3', 5'-CGGTAAGCTT-CAGGTGAAG-3'; internal poison, (A) 5'-CGAACAAGGG-AAGCGACAGC-3', (B) 5'-GTTGGCACTGAAGCGTCCAG-3'.

Each primer was used at a final concentration of 1  $\mu$ M. We did PCR using template DNA purified from wild type or *dim-1* mutant strains. Each of the necessary deoxynucleotide triphosphates was at a concentration of 250  $\mu$ M. The PCR cycling parameters were as follows: 94°C, 30 s; 60°C, 30 s; 72°C, 90 s. We did 35 total cycles. PCR fragments were analyzed on 1% agarose gels.

### Generation and screening of mutant libraries

To induce deletion mutations, young adult hermaphrodites were treated with trimethylpsoralen and UV light as described (7) (E.Gilchrist, M.Edgley, G.Mullen, B.Shen, T.Rogalski, T.Szczygielski and D.Moerman, manuscript in preparation). F<sub>1</sub> progeny of mutagenized animals were cultured in 1152 groups of 50 worms each. After one generation, DNA was prepared from each population by proteinase K lysis. PCR was used to identify populations with animals carrying deletions using the cycling parameters described below. Populations carrying a deletion were repeatedly subdivided until homozygotes carrying the deletion were obtained. Deletion endpoints were determined by sequencing PCR products that spanned the deleted region.

### PCR conditions

External-round screening PCRs were 10  $\mu$ l and contained 2  $\mu$ l library DNA, 4 pM each external primer and poison primer, 200  $\mu$ M each dNTP, 1 $\times$  standard PCR buffer containing

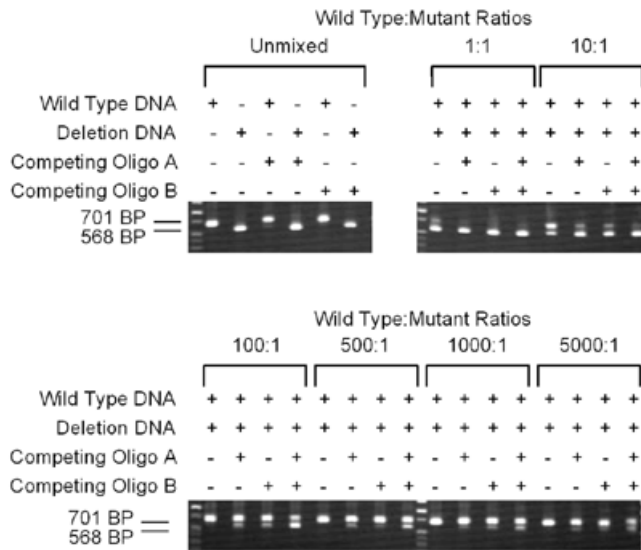
1.5 mM MgCl<sub>2</sub> and 0.2 U *Taq* polymerase (Roche). Internal-round PCRs were identical in all respects except for input DNA (0.2  $\mu$ l from external round) and primers (4 pM each internal primer). Replication of the external round into the internal round was done with a Robbins Hydra 96 pipetting station. Thermal cycling was done on MJ Research PTC-200 DNA Engines or Biometra Uno II thermal cyclers using the following conditions: 94°C/30s, 35 cycles of 94°C/30 s – 61°C/30 s (external round) or 55°C/30 s (internal round) – 72°C/60 s, followed by cooling to 4°C. Loading buffer (10  $\mu$ l; 15% Ficoll with bromophenol blue and xylene cyanol) was added to each reaction, and 1  $\mu$ l of this mix was loaded into each lane of 2% agarose gels. Gels were stained for 30 min with fresh 1:10 000 SYBR® Green (BMA) and imaged with a Molecular Dynamics Fluorimager. The molecular weight marker is BioRad 100 bp PCR Molecular Ruler (Cat. 170-8206). The wild type amplification product is 1151 bp and the deletion product is 810 bp.

## RESULTS

Our initial experiments were done with a known small deletion in the *C.elegans* gene *dim-1* (T.Rogalski and D.Moerman, unpublished data). We first tested different pooling strategies, hoping to reduce the competition between the deletion and wild type fragments. We found that a 133 bp deletion could be detected only if the initial ratio of wild type to deletion template did not exceed 10:1 (Fig. 2). Together with the forward frequency for PCR detectable deletions generated by TMP (probably <1  $\times$  10<sup>-5</sup> for the average gene) this indicates that an unreasonably large number of PCRs would be required to detect a small deletion. It was not feasible, therefore, simply to reduce the pooling depth to improve our ability to detect small deletions.

### Reconstruction experiments with known deletion alleles

To improve the sensitivity of the PCR assay for deletion fragments that are close to the wild type size, we developed the strategy shown in Figure 1. This strategy is a modification of our original protocol, in which we used a two-step nested PCR series to detect deletions between the primer sets. In the modified version, a third functional PCR primer that falls between the two external primers is included in the first round of nested PCR. Amplification from the wild type template leads to the production of two fragments, one full length and the other relatively short. In practice, the shorter fragment is produced much more efficiently than the longer. Amplification from a mutant template, in which the site for the third internal primer is deleted, leads to the production of a single mutant fragment from the external primers. In the second round of PCR, we use two primers placed just inside the external first round primers. The shorter wild type band from the first round cannot serve as a template for the second round PCR because it does not include one of the second round primer binding sites. The longer wild type fragment can serve as a template, but because its production was limited by competition in the first round, its production in the second round is limited correspondingly. The internal functional primer is called a 'poison' because it interferes with production of the full-length wild type fragment. The effectively lower level of wild type product gives the deletion fragment an advantage. Our data show that we can detect a 133 bp deletion at wild type to mutant ratios of 5000:1 (Fig. 2).



**Figure 2.** Poison primers improve sensitivity for small deletions. We tested the poison primer method on a known deletion mutant. We did nested PCR using two different internal poison primers, designated competing oligos A and B that were known to fall within the deleted sequence. Different amounts of wild type and deletion mutant DNAs were mixed at the indicated ratios. With no poison primer, wild type gives a 701 bp PCR fragment. The deletion mutant gives a 568 bp fragment. Without the poison primer, we lose the ability to detect the deletion at wild type:mutant ratios between 10:1 and 100:1. With the poison primer, we can detect the deletion fragment at ratios of 5000:1.

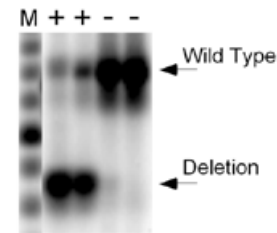
### Screens for new alleles

We tested our new strategy in screens for new deletion mutants. We cultured the  $F_1$  progeny of mutagenized parents in sets of 10 or 50 worms. We harvested a portion of each  $F_2$  population and prepared DNA for PCR. Although, as described above, we were able to detect a small 133 bp deletion in pools where the ratio of wild type to mutant DNA reached 5000:1, we did not exceed a ratio of 1200:1 in screens for new unknown deletions. Figure 3 shows an example of a 341 bp deletion identified using the poison primer method and a pooling ratio of 240:1. The detection of this deletion required the presence of the poison primer, as the deletion band was clearly visible when it was in the reaction mix but was not detected otherwise (Fig. 3). This example clearly illustrates the enhanced detection that can be obtained using this strategy.

We identified 12 deletion strains from a UV/TMP mutagenized library of mutants using the poison primer method. The deletions ranged from 78 to 850 bp. Examples of these are shown in Figure 4A. Figure 4B shows the position of a small deletion that we recovered in a gene that resides entirely within the intron of another gene. DNA sequencing showed that 11 of the 12 deletions encompass the poison primer binding sites. The poison primers were required to detect 10 of these 11 deletions in the initial mutant library screens (data not shown). The exceptional case was a deletion of 850 bp in a 1335-bp wild type interval.

### DISCUSSION

Genome sequence data provide a platform for standard mutational genetics, a powerful strategy to examine gene



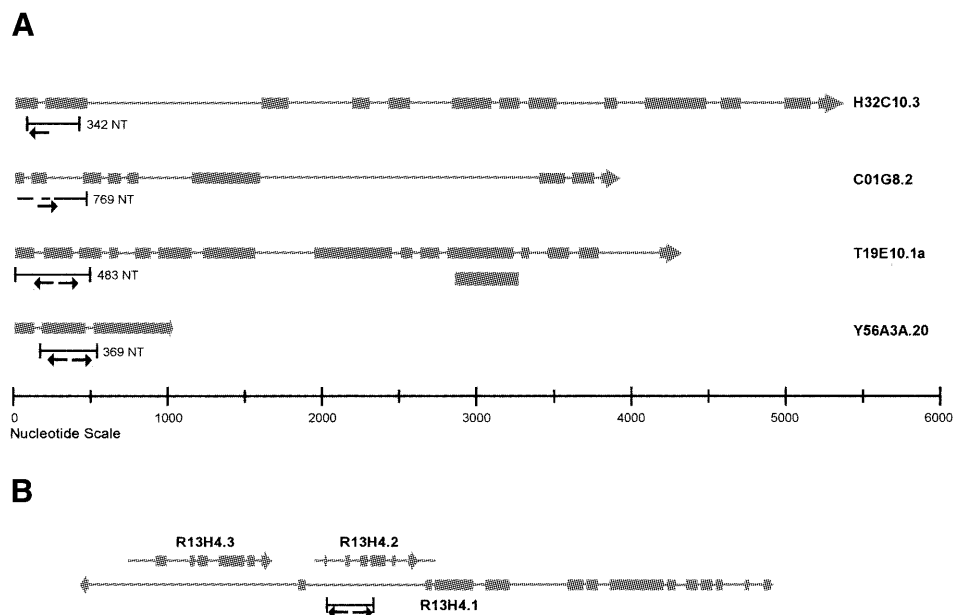
**Figure 3.** Poison primer strategy works to identify new deletion alleles. Gel image of nested PCR amplification products illustrating the effect of a poison primer in the external amplification round on deletion detection. Twelve samples of pooled library DNA, one of which contained a 341-bp deletion in the gene H32C10.3 at a complexity of one deletion chromosome in 240, were subjected in duplicate to PCR amplification with and without a poison primer that lies within the deleted region. Starting after the marker lane (M), the two leftmost lanes contain products of reactions with the poison primer, and the two rightmost lanes contain products of reactions without the poison. The wild type and deletion fragment positions are as indicated.

function. To exploit genetic methods, however, one must identify mutations at target loci. Using our original protocols, we could detect deletions >600 bp in an interval of 3000 bp. Using the poison primer protocol, we have detected deletions as small as 78 bp. We also tested whether we could improve the method through the simultaneous use of two closely spaced poison primers annealing to opposite strands of the target. Although our data set is small, two poisons appeared to be somewhat better at reducing the level of the competing wild type PCR product (data not shown).

Although the poison primer method allows for the detection of small deletion alleles in complex DNA pools, it changes the target for deletion relative to our typical strategy. With the poison primer method the deletion must eliminate the internal poison primer site. In a 3000 bp interval, therefore, with a poison annealing to a single site one can detect only ~7% of the total number of 100 bp deletions, 14% of the 200 bp deletions, etc. The poison primer method, however, gives greater control over the position of the deletion within the target, thereby allowing for the recovery of more precise 'designer' deletions. Cases where this method might be useful include the deletion of a single exon in alternatively spliced genes, the deletion of parts of promoters, or the deletion of genes that fall within introns of other genes. The data in Figure 4B validate the poison primer method for such uses.

Although the exact endpoints of the deletions are not subject to our control, we generally target the 5' end of the coding sequence to enhance the odds that the recovered deletion alleles will eliminate all functions of the gene. We have developed a public interactive web site that can be used to identify primer sequences that meet our design criteria (<http://ko.cigenomics.bc.ca/oligos.shtml>).

We have used the nematode *C.elegans* to develop a new method to detect deletion mutations in known genes. Our new method enhances the reverse genetic tools available for *C.elegans* and other genetic model systems and, further, adds to the growing list of genetic tools developed to screen for mutations in humans in appropriate clinical or research populations.



**Figure 4.** Detection of deletions with the poison primer technique. The gene structure for each of the genes listed was predicted from genomic DNA sequence. Grey boxes and thin lines represent exons and introns, respectively. The positions of the deletions and their sizes are shown below each gene. Arrows represent the poison primers designed to interfere with the amplification of a full-length wild type PCR product. Particular regions of the gene can be targeted based on the placement of such primers. **(A)** Representative deletion mutations. Deletion breakpoints were determined by DNA sequencing of the internal PCR products amplified from DNA isolated from single worms. The deletions shown are 342 NT, 769 NT, 483 NT and 369 NT in the coding regions of H32C10.3, C01G8.2, T19E10.1a and Y56A3A.20, respectively. Note that the deleted regions contain all or part of the poison primer-binding site. **(B)** Precise excision of a nested gene. R13H4.1 is a large gene whose coding sequence is on the minus strand of clone R13H4 of chromosome II of *C.elegans*. The gene R13H4.2, one of two genes nested within introns of the larger gene on the opposite strand, was targeted with two oppositely oriented poison primers. The deletion mutation detected with these primers excised the entire 3' end of R13H4.2, including the target exon, but did not extend outside of the intron of the larger gene, leaving its coding sequence intact.

## ACKNOWLEDGEMENTS

We thank Marco Marra and Steven Jones of the Genome Sequence Center in Vancouver BC for their advice and enthusiasm for this project, and for the use of their fluorimager to scan gels. This work was supported by a National Institutes of Health grant R01 HG01843 to R.B., and a grant from the Canadian Institute for Health Research to D.G.M. Additional support was provided by the Biotechnology Laboratory and U.B.C. to the *C.elegans* Reverse Genetics Core Facility.

## REFERENCES

1. The *C.elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C.elegans*: a platform for investigating biology. *Science*, **282**, 2012–2018.
2. Barstead,R.J. and Waterston,R.H. (1991) Vinculin is essential for muscle function in the nematode. *J. Cell Biol.*, **114**, 715–724.
3. Rushforth,A.M., Saari,B. and Anderson,P. (1993) Site-selected insertion of the transposon Tc1 into a *Caenorhabditis elegans* myosin light chain gene. *Mol. Cell. Biol.*, **13**, 902–910.
4. Jansen,G., Hazendonk,E., Thijssen,K.L. and Plasterk,R.H. (1997) Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nature Genet.*, **17**, 119–121.
5. Dernburg,A.F., McDonald,K., Moulder,G., Barstead,R., Dresser,M. and Villeneuve,A.M. (1998) Meiotic recombination in *C.elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell*, **94**, 387–398.
6. Liu,L.X., Spoerke,J.M., Mulligan,E.L., Chen,J., Reardon,B., Westlund,B., Sun,L., Abel,K., Armstrong,B., Hardiman,G. *et al.* (1999) High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.*, **9**, 859–867.
7. Anderson,P. (1995) Mutagenesis. *Methods Cell. Biol.*, **48**, 31–58.
8. Gengyo-Ando,K. and Mitani,S. (2000) Characterization of mutations induced by ethyl methanesulfonate, UV and trimethylpsoralen in the nematode *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.*, **269**, 64–69.