Single-Nucleotide Polymorphism Mapping

M. Wayne Davis and Marc Hammarlund

Summary

Single-nucleotide polymorphism (SNP) mapping is the easiest and most reliable way to map genes in Caenorhabditis elegans. SNPs are extremely dense and usually have no associated phenotype, making them ideal markers for mapping. SNP mapping has three steps. First, recombinant mutant animals are generated over a polymorphic strain (usually CB4856) using standard genetic techniques. Second, the genotype of these animals at SNP loci is determined using one of a variety of SNP detection technologies. Third, linkage between the mutant and one or more SNPs is used to position the mutant on the chromosome relative to the SNPs. This chapter presents a detailed procedure for generating recombinant animals, for assaying SNPs using restriction enzymes, and for analyzing mapping data.

Key Words: C. elegans; single-nucleotide polymorphism; mapping; restriction fragment polymorphism; genetics.

1. Introduction

Single-nucleotide polymorphism (SNP) mapping in Caenorhabditis elegans is a powerful and elegant tool for determining the chromosomal position of a mutant gene. SNP mapping as a method to clone a gene in C. elegans was first used by Jakubowski and Kornfeld (1). More recently, SNP mapping against the polymorphic strain CB4856, a strain isolated in Hawaii (2), was introduced by Wicks et al. (3), who identified a large number of polymorphisms and described the bulk segregant analysis procedure presented in this chapter. Additional polymorphisms in CB4856 were identified by the Waterston group at St. Louis, and by Exelixis (4). Refinements to SNP mapping techniques were developed by Exelixis (4), Zipperlen et al. (5), and Davis et al. (6). These advances have made SNP mapping a robust and rapid method to identify and clone a gene of interest.
SNPs have two advantages over conventional marker mutations. First, unlike conventional visible markers, SNPs in general have no phenotype, allowing a mutation of interest to be scored in a neutral phenotypic background. As a result, many markers can be assayed simultaneously, without worrying about genetic interactions. Second, SNPs between N2 and CB4856 are very dense. The average density is approximately one per kilobase of DNA (3), and to date more than 15,000 SNPs have been identified. Thus, SNP mapping can provide data on very small intervals, and in some cases is capable of mapping to single-gene resolution.

Most SNP mapping experiments begin with chromosome mapping (also known as bulk segregant analysis), and then employ successive rounds of interval mapping until a narrow physical region is identified. SNP mapping can easily and quickly narrow the physical location of a gene to a very narrow region at which point cosmid rescue or sequencing experiments can be done to specifically identify the gene of interest. Both chromosome and interval mapping begin with collecting recombinants over CB4856. In chromosome mapping these recombinants are pooled, so that many recombinations are assayed at once. This allows the proportion of recombination between the allele of interest and individual SNPs to be determined. Chromosome mapping is thus a statistical assay that generally defines the chromosome and approximate position on the chromosome (left arm, middle, right arm) of a gene. Interval mapping, in contrast, involves the analysis of individual recombinants and places physical limits on the position of the allele of interest. Because of the high density of SNPs, analysis of enough recombinants will limit the position of the gene of interest to a very small interval. This chapter provides detailed protocols for chromosome and interval mapping.

SNP mapping presents two major challenges. The first, collecting an adequate number of recombinants, is not unique to SNP mapping. However, the requirement for generating recombinants over CB4856 makes this process more difficult. We discuss several solutions to this problem. Second, because of their silent nature, SNPs cannot be assayed directly. We provide a detailed protocol for detecting SNPs by restriction digest, as introduced by Wicks et al. (3), and discuss some alternatives to this approach.

When contemplating SNP mapping, three other techniques should be considered. First, mobile DNA elements, or transposons, can be used as mutagens (7). Because insertion of an exogenous transposon can simultaneously mutate a gene and tag it with a unique molecular sequence, genes identified in forward transposon screens can be cloned without mapping. Second, the development of RNAi feeding libraries allows whole-genome screens to be performed (8). This method eliminates all mapping because the molecular target of each RNAi construct is known. Likewise, when mutations are created by reverse genetic techniques, mapping is not necessary (9).
However, chemical mutagenesis followed by mapping and cloning is still the preferred approach in many circumstances (10). For example, many genes are resistant to RNAi, and cannot be identified in RNAi screens. Chemical mutagenesis is also the best way to generate an allelic series or to find unusual alleles. Finally, chemical mutagenesis is by far the easiest way of producing mutants, and, therefore, lends itself to complex screens. Fortunately, SNP mapping means that interesting alleles, even those with subtle phenotypes, can be quickly and accurately mapped.

2. Materials

2.1. Tools

1. 0.2 µL 96-pin replicator (hedgehog; Boekel).
2. Gel box (Owl Scientific Centipede).
3. Eight-channel pipet (20 µL vol): Rainin LTS (see Note 1).
4. Repeating pipet (optional) (Gilson Distriman).
5. 96-Well format thermal cycler (any supplier). Protocols in this chapter have been performed on MJ Research DNA Engine™ thermal cyclers.

2.2. Reagents

1. Single worm lysis buffer (SWLB) + proteinase K: 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 2.5 mM MgCl₂, 0.45% IGEPAL CA-630 (or NP40), 0.45% Tween-20, 0.01% (w/v) gelatin. Autoclave, then store in aliquots at –20°C. Add fresh proteinase K to 60 µg/mL just before use.
2. 4X SWLB + proteinase K: 200 mM KCl, 40 mM Tris-HCl, pH 8.3, 10 mM MgCl₂, 1.8% IGEPAL CA-630 (or NP40), 1.8% Tween-20, 0.04% (w/v) gelatin. Autoclave, then store in aliquots at –20°C. Add fresh proteinase K to 240 µg/mL just before use.
3. 10X PCR buffer: 22.5 mM MgCl₂, 500 mM Tris-HCl, 140 mM (NH₄)₂SO₄, pH 9.2 at 25°C.
5. Taq DNA polymerase (any supplier).
6. dNTPs (any supplier). 10 mM each dNTP (40 mM total concentration of dNTP) in 50% sterile glycerol. This can be made from 100 mM stock solutions as: 10 µL dATP, 10 µL dCTP, 10 µL dGTP, 10 µL dTTP, 60 µL 80% glycerol (see Note 2).
7. PCR primers (any supplier). Protocols in this chapter have been preformed using primers from Integrated DNA Technologies (see Note 3).
8. Gel-loading buffer: 0.2% orange G, 0.05% xylene cyanol FF, 60% glycerol, 60 mM EDTA (see Note 4).

2.3. Disposables

1. 96-Well PCR plates (BD Falcon, cat. no. 352133).
2. PCR plate sealing mats (Genemate, cat. no. T-3161-1; distributed by ISC-BioExpress).
3. Methods

We provide a detailed method for chromosome mapping in Subheading 3.1., and for interval mapping in Subheading 3.2., using restriction-length polymorphism detection to assay SNP genotypes. These methods are also described in Davis et al. (6). Using these methods, a mutant gene can be mapped to an interval of 5–10 map units. To map more finely requires the design of additional SNP primers and the collection of many recombinant animals; brief protocols for these steps are given in Subheadings 3.2.5. and 3.2.6. You may also want to consider using one of the other SNP detection techniques that have been employed in C. elegans (Subheading 3.3. contains a description of these). Finally, Subheading 3.4. discusses the mapping of challenging genotypes.

3.1. Chromosome Mapping

The first step in mapping a mutant gene is to assign the gene to a broad region of one chromosome. The method presented here uses 48 SNPs evenly spaced along the chromosomes, allowing the whole genome to be scanned simultaneously. The 48 primer sets are designed to amplify under identical conditions, and the SNPs are detected using identical restriction digestion conditions. The primers are arrayed in a 96-well plate to maximize reproducibility and to minimize set-up time. Two DNA templates are generated, one from a mixture of 50 homozygous mutant animals and a control template from nonmutant animals (or an equal mixture of N2 and CB4856 animals). Unlinked SNPs should have equal N2 and CB4856 content. Thus, mutant templates and control templates should produce an equal proportion of N2 and CB products at unlinked loci. Linkage to a SNP will result in an increase in the ratio of N2 DNA/CB DNA in the mutant relative to the control. The final output of this method is an agarose gel that has mutant and control PCR products side by side, making the differences between the mutant and the control clearly visible.

3.1.1. Genetics

1. Obtain homozygous CB4856 males (see Note 5).
2. Cross them into homozygous mut/mut hermaphrodites (see Note 6).
3. After 24 h the hermaphrodites should have a mating plug. Clone each mated hermaphrodite to a fresh plate (see Note 7).
4. In 3 d clone 8–12 L4 hermaphrodites; these should be mut/CB4856 (see Note 8).
5. Wait 3–4 d, or until Mut and non-Mut progeny can be distinguished (see Note 9).
6. Put 50 Mut animals into a tube containing 20 µL SWLB + proteinase K. Put 50 non-Mut animals in a second tube (see Note 10).
7. Freeze tubes at –80°C for at least 10 min. This is a convenient place to stop if so desired.

3.1.2. Chromosome Mapping Primers

Primer pairs are premixed in a separate 96-well plate and added to the reaction by pin replication. This reduces both reaction preparation time and the potential for errors. The primers are formatted to produce a gel that has each mutant digest next to its control, moving in order from LG IL to LG XR. To accomplish this, the primers are set up in a 96-well format so that row 1 contains the eight primer pairs for LG I, with well 1H containing primers for the left-most marker. Row 2 is a duplicate of row 1. Rows 3 and 4 are duplicates and contain the LG II primer pairs, rows 5 and 6 have primers for LG III, and so on. When the reactions are loaded onto the gel using a multichannel pipet duplicate rows will be interleaved, placing each mutant reaction directly adjacent to its corresponding control.

Dissolve primers in 10 mM Tris-HCl, pH 8.0, and add primers to each well of the round-bottomed storage plates at 10 µM each primer. The primer stock plates can be frozen in a nonfrost-free freezer and reused many times; however, multiple plates should be made at the same time with only 20 µL per well because plates can go bad for a number of reasons (left out of freezer, hot pin replicator, cross-contamination, dropped, and so on). The plates should always be briefly spun in a tabletop centrifuge before removing the lids in order to prevent cross-contamination owing to condensation in the lid.

3.1.3. SNP Genotyping

1. Lyse worms. Incubate 65°C for 1 h, 95°C for 15 min. After lysis, keep template on ice and store at –80°C when not in use to prevent degradation of DNA.
2. Set up PCR master mixes. For each mutant you are mapping, make two tubes of the following: 52 µL 10X PCR buffer, 10.4 µL 10 mM dNTPs, 3.12 µL Taq (5 U/µL), 424 µL water. Add 20 µL of your Mut lysis to one tube, and 20 µL of your non-Mut lysis (or control DNA) to the other. Mix thoroughly.
3. Add master mixes to the plate. Using a multichannel or repeating pipet, aliquot 9.8 µL of the Mut mix into rows 1, 3, 5, 7, 9, and 11 of a 96-well PCR plate. Aliquot 9.8 µL of the non-Mut (or control) mix into rows 2, 4, 6, 8, 10, and 12.
4. Pin replicate primers. Dip replicator into primer plate and then into PCR plate (see Note 11).
5. PCR. Cycling parameters: 94°C for 2 min, (94°C for 15 s, 60°C for 45 s, 72°C for 1 min) 35 times, 72°C for 5 min. Because the volumes are small, a heated lid is critical to prevent evaporation.
6. Digest. Mix 168 µL 10X DraI buffer, 26.25 µL (262 U) DraI, 435.75 µL water. Aliquot 6 µL into each PCR well, and centrifuge the plate to mix. Incubate at 37°C for 4 h overnight. Evaporation during this step can be a problem. Incubation is best accomplished in a thermocycler with a heated lid. Alternatively, plates can be sealed carefully with aluminum sealing film, making sure each well is fully sealed, followed by incubation in a 37°C air incubator.

7. Add gel-loading buffer. Using a multichannel or repeating pipet add 4 µL of loading buffer to each well (see Note 4).

8. Run the gel. Pour a 2.5% agarose gel using TAE or TBE and put two 50-well combs into the tray. When the gel is ready, load the gel using an eight-channel pipettor. Load row 1 (1H–1A), starting in the first well of the 50-well comb. Load row 2 (2H–2A) interleaved with row 1, starting in the second well. Leave a space, and load rows 3 and 4 interleaved, followed by another space and then rows 5 and 6.

9. This should fill the 50 wells. Load the bottom lanes with rows 7–12. Put 50 bp ladder in the four empty spaces.

3.1.4. Analyzing SNP Data

Unlinked SNPs should have equal N2 and CB4856 content and thus, mutant templates and control templates should produce an equal ratio of cut-to-uncut products. Linkage to a SNP is determined by an increase in the ratio of N2 DNA to CB4856 DNA in the mutant product relative to the control. Depending on the SNP, this may be an increase in cut or in uncut DNA. The ratio is the relevant parameter, not the absolute amount of product. When using eight SNPs per chromosome, a mutation will show tight linkage to at least one SNP and noticeable linkage to adjacent SNPs. Weak linkage to one SNP, but not to adjacent SNPs, is almost always a spurious result.

3.2. Interval Mapping

Once your mutation has been assigned to a broad region of a chromosome using chromosome mapping, interval mapping can be used to narrow the interval to a region small enough for molecular cloning techniques to be applied. Briefly, individual recombinant animals are genotyped for SNPs in the region to which the gene has been mapped. The closest break point on either side of the mutation then defines the physical limit of the mutation’s position. To reduce the size of this interval, additional SNPs within the interval can be assayed. Only those animals that are recombinant within the interval (i.e., have different SNP genotypes at the ends of the interval) are likely to provide additional information. Once the limits are moved so close that there are no more informative recombinants, additional recombinants must be collected, as described in Subheading 3.2.6.

3.2.1. Genetics

1. Generate recombinants. Follow steps 1–5 of Subheading 3.1.1. These animals can be generated from the same crosses used for chromosome mapping (see Note 10).
2. Single at least 200 homozygous Mut hermaphrodites onto individual seeded plates.
3. Take some time to examine the self-progeny brood and verify that each plate is indeed homozygous mutant. Discard any plates that appear questionable.
4. Label the plates clearly, corresponding to a 96-well plate format.
5. Just before the plates starve, add 250 \( \mu L \) sterile water to each plate, tip the plate, and transfer 100 \( \mu L \) containing worms to the well of a 96-well plate. The amount of water can be adjusted based on the size of the plate used and the dryness of the agar. It is recommended to work in sets of four, six, or eight plates at a time. Make sure that the well and the plate label correspond. Use a new pipet tip for each plate to avoid cross-contaminating plates or wells. Also make a well with N2, a well with CB4856, and a well with an equal mix (see Note 12).
6. Store the starved plates at 15\(^\circ\)C for later reference.
7. Place the 96-well plates at 4\(^\circ\)C for 15 min to overnight to allow the worms to settle.
8. Check all of the wells to verify that they have approximately equal volumes and a clearly visible pellet of worms.
9. Carefully remove 55 \( \mu L \) from the top of each well, leaving the worm pellet undisturbed. The remaining volume should be 45 \( \mu L \).
10. Add 15 \( \mu L \) 4X SWLB + proteinase K to each well.
11. Freeze at –80\(^\circ\)C with a foil lid. This is a convenient place to stop. However, freeze even if proceeding immediately, as it helps with lysis.

3.2.2. SNP Genotyping

Inspect your chromosome mapping results and select the three to five adjacent markers that are most tightly linked to your mutation. You will test each of your recombinants for these markers and generate a snapshot of the recombinant chromosomes in the region of your mutant. Make sure the plates are well labeled, including registration marks indicating the orientation of the plate.

1. Lyse worms. Mix the settled worm pellet and SWLB by vortexing the 96-well plate. Incubate at 65\(^\circ\)C in a thermocycler with a heated lid for 30 min. Vortex again. Incubate at 65\(^\circ\)C for 30 min, then heat inactivate the proteinase by heating to 95\(^\circ\)C for 15 min. After lysis, keep template on ice and store at –80\(^\circ\)C when not in use to prevent degradation of DNA.
2. Make a separate master mix for each SNP you wish to genotype. The mixes will differ only in their primers. The volume of mix depends on the number of SNPs and the number of recombinants. See Table 1 for volumes. Add 9.8 \( \mu L \) master mix to each well of the plates.
3. Add templates to the 96-well plates by pin replication (see Note 11). More reliable results have been obtained when the pin replication process is repeated a second time.
4. PCR. Use the same conditions as described in Subheading 3.1.3., or conditions empirically optimized for each new primer set.
5. Digest. Add 6 \( \mu L \) restriction digestion mix to each well. See Table 2 for volumes.
6. Run TAE agarose gel as described in Subheading 3.1.3.
3.2.3. Analyzing SNP Data

We have found that the easiest way to analyze interval mapping SNP data is to put the data into a simple spreadsheet, with a row for each singled plate (the F₂ generation) and a column for each SNP. It is helpful to color each cell depending on the genotype (N₂, CB4856, or heterozygote) by using the conditional formatting function in Excel. If you have selected SNPs that are close to your mutant, most animals will be homozygous N₂ at all SNPs. Some animals will carry recombinant chromosomes and be heterozygous at one or more adjacent SNPs. These recombinant animals provide information about the position of the mutation, because the mutation cannot be in the region that is heterozygous. By comparing all the recombinants, the minimal interval containing the mutation can be identified.

The key is to find two SNPs that are near your mutation, but flank it. At this point, for further mapping experiments you need only genotype each animal for the two nearest flanking SNPs. Animals homozygous N₂ at both of these SNPs are uninformative and need not be assayed further. This will immediately cut your large number of SNP assays to a small number of informative

---

### Table 1
**Interval Mapping PCR Mix Volumes**

<table>
<thead>
<tr>
<th></th>
<th>For each 10 µL rxn</th>
<th>1 plate</th>
<th>2 plates</th>
<th>3 plates</th>
<th>4 plates</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>8.50 µL</td>
<td>850 µL</td>
<td>1700 µL</td>
<td>2550 µL</td>
<td>3400 µL</td>
</tr>
<tr>
<td>10X PCR buffer</td>
<td>1.00 µL</td>
<td>100 µL</td>
<td>200 µL</td>
<td>300 µL</td>
<td>400 µL</td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>0.20 µL</td>
<td>20 µL</td>
<td>40 µL</td>
<td>60 µL</td>
<td>80 µL</td>
</tr>
<tr>
<td>Taq polymerase</td>
<td>0.06 µL</td>
<td>6 µL</td>
<td>12 µL</td>
<td>18 µL</td>
<td>24 µL</td>
</tr>
<tr>
<td>Fwd primer 100 µM</td>
<td>0.02 µL</td>
<td>2 µL</td>
<td>4 µL</td>
<td>6 µL</td>
<td>8 µL</td>
</tr>
<tr>
<td>Rev primer 100 µM</td>
<td>0.02 µL</td>
<td>2 µL</td>
<td>4 µL</td>
<td>6 µL</td>
<td>8 µL</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9.80 µL</td>
<td>981 µL</td>
<td>1962 µL</td>
<td>2943 µL</td>
<td>3924 µL</td>
</tr>
</tbody>
</table>

### Table 2
**Interval Mapping Digest Mix Volumes**

<table>
<thead>
<tr>
<th></th>
<th>For digesting each 10 µL PCR</th>
<th>1 plate</th>
<th>2 plates</th>
<th>3 plates</th>
<th>4 plates</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>4.15 µL</td>
<td>415 µL</td>
<td>830 µL</td>
<td>1245 µL</td>
<td>1660 µL</td>
</tr>
<tr>
<td>10X buffer</td>
<td>1.60 µL</td>
<td>160 µL</td>
<td>320 µL</td>
<td>480 µL</td>
<td>640 µL</td>
</tr>
<tr>
<td>Enzyme</td>
<td>0.25 µL</td>
<td>25 µL</td>
<td>50 µL</td>
<td>75 µL</td>
<td>100 µL</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>6.00 µL</td>
<td>600 µL</td>
<td>1200 µL</td>
<td>1800 µL</td>
<td>2400 µL</td>
</tr>
</tbody>
</table>
recombinant DNA templates. A sample that is heterozygous at the flanking SNP should be reassayed for intermediate SNPs. As you assay these different SNPs, fewer samples will be heterozygous, until you have found two SNPs that each have a single heterozygote. These SNPs mark the left and right boundaries of the genetic interval that contains your gene of interest. If the boundary SNPs are sufficiently close to each other or bracket a sufficiently promising candidate gene, other molecular techniques can be used to clone the gene (transgene rescue, sequencing). If they are too far apart, a new collection of recombinants should be collected and assayed for the SNPs that mark the left and right of your interval.

In some cases, the problem arises that the initial set of SNPs are all on one side of your mutation. In this case, the problem can be solved by assaying new SNPs until you have found an SNP on the other side. Another common problem is the occurrence of rare genotypes that are inconsistent with a unique map position. The genotypes may be because of misscoring the phenotype. When the plates corresponding to these SNP reactions are regrown and re-examined for the phenotype it is almost always discovered that they have indeed been misscored. Finally, it is possible to see no obvious linkage from the genotypes. In these cases, it may be that the chromosome mapping gel has been misinterpreted. It is also possible that the phenotypes that were selected are actually represented by multiple genotypes.

3.2.4. Designing New SNP Primers

When using SNP mapping to clone a gene, you will almost always have to go through several rounds of designing SNP detection primers that provide information in subsequently narrower genetic intervals. There are several web-based tools that assist worm researchers in identifying verified and potential SNPs that have been found using shotgun sequencing of CB4856. Most of this data has been provided by two groups: one at the Genome Sequencing Center at Washington University at St. Louis (3), and one at Exelixis (4). They have both made their primary data available on the web. The St. Louis data are available at website: http://genome.wustl.edu/projects/celegans/index.php?snp=1 and the Exelixis data are available at website: http://www.exelixis.com/index.asp?secPage=elegans. These data have been combined and made searchable at the WormBase website (http://www.wormbase.org/db/searches/strains). This website is geared towards mapping against visible markers but can provide SNP data in an interval. It can filter the SNPs for verified or unverified and for snip-SNP (contains a polymorphism that changes a restriction site) or nonsnip-SNP (contains a polymorphism that can only be identified by sequence). For verified SNPs, the primers used by the verifying lab are given. At the time of writing, the site was still very much a work in progress, with inconsistent results that fre-
quentely listed snip-SNPs as only detectable by sequencing, and listing numerous unavailable restriction enzymes.

Alternatively, a Filemaker (Filemaker Inc., http://www.filemaker.com) database file that contains data from both sequencing groups analyzed for polymorphic restriction sites has been compiled by the authors and is available from http://www.biology.utah.edu/jorgensen/SNPs.

When a promising SNP has been identified, primers that amplify the SNP fragment must be designed. A wide variety of primer design methods and utilities have been described, but the authors have found that the web-based primer3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) produces reliable results. Exact primer properties seem to be unimportant in achieving SNP results, but designing all SNP primers with a uniform melting temperature can often assist in minimizing PCR optimization steps and often allows multiple SNPs to be assayed in a single batch of PCR. Obviously, primers should be designed to provide a detectable difference in restriction fragment length. We have found that using 2.5% agarose gels bands with 30 bp differences can be detected, with 50 bp the smallest absolute size. If possible, the two bands produced by restriction digestion should be of different sizes. In addition, some people prefer to have a nonpolymorphic site present in the PCR product as a positive control for digestion. However, a nonpolymorphic site too close to the SNP site will make the SNP impossible to detect by restriction digestion.

Even if you chose to use primers previously designed and tested by others, it is a good idea to test the primers on at least N2, CB4856, and a mixture of N2 and CB4856 DNA before embarking on a massive PCR screen. If time and equipment provide, a gradient PCR machine should be used to empirically find an optimal annealing temperature for each new primer set.

In some cases, the known SNPs will not be sufficiently dense to provide satisfactory map information for cloning a gene. In this case, the researcher will have no other option but to look for SNPs of their own, by direct sequencing of CB4856. The density of known SNPs suggests that an SNP can be found in approximately every 1 kb (3). The actual SNP frequency may be very different in different regions of the genome, but a few sequencing reads from PCR products concentrated in noncoding regions should provide a SNP. Approximately 50% of new SNPs should be snip-SNPs, and 60% or more of these can produce restriction fragments with large enough size differences for detection on agarose gels (3).

3.2.5. Directed Mapping

It is an unfortunate truth that the closer one gets to one’s gene, the more difficult mapping becomes. This is a simple consequence of informative recombinants comprising a smaller and smaller fraction of the whole. For example, if
SNP Mapping

in the first round of interval mapping you position your gene between two SNPs that are five map units apart, you can expect that only 10% of future recombinants will fall in that interval and, thus, provide information. In many cases, the application of brute force will solve this problem by collecting many recombinants, genotyping them at the flanking SNP loci, and saving for further analysis those animals that have different genotypes at these loci. These are recombinant in the region of interest, and can be analyzed further.

However, a more elegant solution also exists. This solution relies on using recombination between visible markers to enrich for recombination events in the region of interest. We describe next three experimental approaches for directed mapping. Only the genetics part of SNP mapping is covered. Once recombinants have been collected, the mapper should assay their SNP genotypes using the methods described in Subheading 3.2.2.

3.2.5.1. METHOD 1: TWO-FACTOR MAPPING

You have established that your gene is flanked by SNP a and SNP b, and you wish to enrich for recombinants within the ab region. To use a two-factor approach, you need to construct a double mutant containing your gene (mut-1) and a visible marker (vis-1). The marker should lie outside but close to the ab interval, and you will obtain map data only on the side the marker is on. Also, you must be able to reliably distinguish at least one of the recombinant phenotypes (Mutant non-Marker or Marker non-Mutant) from the parental phenotypes (Mutant Marker and wild-type).

Hermaphrodites with the genotype vis-1 mut-1/CB4856 are allowed to self. These heterozygotes will segregate rare progeny with the recombinant phenotypes. If vis-1 is close to SNP a, many of these crossovers will be in the a-mut-1 interval and thus informative about the limit of mut-1 on the side containing SNP a.

This approach can be useful if there are good visible markers near your SNP boundaries. Because it only generates data on one side, two parallel experiments are usually necessary.

3.2.5.2. METHOD 2: THREE-FACTOR MAPPING

Three-factor mapping has two large advantages over two-factor mapping. First, information is obtained on both sides of the mutation from a single experiment. Second, it is not necessary to score the mutant phenotype when collecting recombinants. The initial procedure is similar to two-factor mapping, except the mutant is marked on both sides. To use a three-factor approach, you need to construct a triple mutant containing your gene and two visible markers, vis-1 and vis-2. The markers should flank the ab interval as closely as possible. You should
be able to reliably distinguish at least one of the recombinant marker phenotypes (Marker 1, non-Marker 2, or Marker 2, non-Marker 1) from the parental phenotypes (Marker 1, Marker 2, and wild type).

Hermaphrodites with the genotype \( \text{vis-1 mut-1 vis-2/CB4856} \) are allowed to self. These heterozygotes will segregate rare progeny with the recombinant phenotypes Marker 1, non-Marker 2, or Marker 2, non-Marker 1. These animals will have at least one chromosome that is recombinant in the \( \text{vis-1-vis-2} \) interval. These animals can be scored for \( \text{mut-1} \) immediately, or cloned and scored in the next generation as a population.

Again, this approach requires good visible markers, and also requires the construction of a fairly complex strain. If these conditions are met, this approach is very powerful.

3.2.5.3. Method 3: Using the CB4856 Background

In this approach, mutations are generated in the CB4856 background and mapped against N2, rather than the reverse. Either two- or three-factor mapping can then be done, with the significant advantage of not having to mark the mutant strain. Instead, the mutant is simply placed over a marked N2 strain. Recombination either between the mutation and an N2 marker or between two N2 markers can be used to enrich for informative recombinants.

This approach is attractive because it does not require strain building. There are many existing double-marker strains in the N2 background, in lab collections or available from the CGC. However, having the mutation of interest in the CB4856 background could prove problematic for later study.

3.3. SNP Detection Technology

Several SNP detection techniques have been shown to work well in \( \text{C. elegans} \). The first technique described for SNP detection in \( \text{C. elegans} \), and still the most widely used, is restriction digestion detection after PCR (3)—the strategy used in the previously mentioned protocols for chromosome and interval mapping. In 2002, Swan et al. (4) published an application of fluorescence polararization-template directed incorporation (FP-TDI) for use in SNP detection in \( \text{C. elegans} \).

More recently, a method involving detection of small insertion/deletion (InDel) polymorphisms using capillary sequencing has been developed by Zipperlen et al. (5). Many other techniques, including oligonucleotide microarray hybridization (11), locked nucleic acid (LNA) hybridization (12), or allele-specific extension (13), rolling-circle amplification (14), Taqman (15), and mass spectroscopy (16), are used to detect SNPs in other systems and could be applied to \( \text{C. elegans} \) after investment in development of the necessary worm-specific reagents. The choice of a particular technique for SNP detection will ultimately depend on consideration of your specific requirements and resources.
3.3.1 Restriction Fragment-Length Polymorphism Detection

Restriction digestion SNP (known as snip-SNP) detection is currently the most commonly used technique in the worm field, probably because it has low equipment and expertise requirements. Because it depends on the SNP changing a recognition site for a restriction enzyme, a major limitation of this technique is that not all SNPs can be detected. However, snip-SNPs between N2 and Hawaiian are dense enough to give resolution finer than one map unit in most genomic regions. This is usually dense enough to allow cloning the gene, or to limit the number of animals bearing informative recombinant chromosomes to a small enough number that additional SNPs can be assayed by DNA sequencing. Another drawback of snip-SNP detection is that each SNP in a genomic region usually requires a different restriction enzyme. This means that frequently each SNP requires different digestion conditions. This adds additional hands-on processing time and becomes a potential source of confusion if multiple SNPs are assayed at the same time. Further, some SNPs can only be detected with relatively expensive or difficult to obtain enzymes.

3.3.2. FP-TDI Detection

Detection of SNPs by FP-TDI relies on the fact that a fluorescently labeled nucleotide in solution has different fluorescence polarization characteristics than one incorporated in a DNA strand. Dye terminator nucleotides are used in a primer extension reaction at the SNP, and fluorescent polarization information is used to detect which base has been added and thus identify the SNP genotype. FP-TDI overcomes some of the limitations of snip-SNP detection, but requires an initial investment in a moderately expensive piece of equipment (a plate-reading fluorescence polarimeter). Unlike many other methods, FP-TDI can be used to detect any sequence polymorphism. In addition to many confirmed C. elegans FP-TDI primers identified by Exelixis, many C. briggsae FP-TDI primers have been identified by Washington University. These primers are available at http://snp.wustl.edu/snp-research/c-briggsae/. Because all of the detection reactions can be carried out in a single well of a 96- or 384-well plate, FP-TDI can be automated, reducing the hands-on time and the potential for mistakes. Unlike snip-SNP detection, no SNP-specific enzymes are required for FP-TDI. However, because of the expense of some of the reagents used in FP-TDI the cost per reaction is comparable or higher than snip-SNP detection. In addition, the plate reading fluorescence polarimeter costs tens of thousands of dollars.

3.3.3. InDel Detection

The InDel detection technique described by Zipperlen et al. (5) has a mix of advantages and disadvantages over snip-SNP detection and FP-TDI detection.
Like snip-SNP detection, InDel detection is limited to a subset of SNPs. In this case, it is limited to SNPs that insert or delete one or two base pairs, estimated to be about 20% of all *C. elegans* SNPs (5). As in the other techniques, PCR is used to amplify the region containing the polymorphism. However, in this technique one of the primers must be fluorescently labeled. The fluorescent PCR product is then run on an ABI capillary sequencer to resolve the PCR products that vary by 1 or 2 bp. This approach has the advantages of automation found in FP-TDI and simplicity of snip-SNP detection. However, the availability of the appropriate equipment and the cost of reagents may limit the access of some labs to this technique. A variation of this technique, called SNPwave, which uses an initial ligation mediated detection step to detect single base pair changes, has been published by van Eijk et al. (17).

### 3.4. Mapping Challenges

#### 3.4.1. Complex Genotypes

Previous sections discussed mapping of simple recessive loss of function alleles. However, an added benefit of SNP mapping is that genes contributing to complex phenotypes can also be mapped relatively easily. For example, semidominant or dominant alleles, phenotypes that require mutant alleles at more than one locus, and second-site enhancer and suppressor alleles can all be mapped using SNPs.

The key to SNP mapping these complex genotypes is to identify a phenotype segregating from a heterozygous animal that results from a single, unique genotype. This will almost always be a homozygous genotype. In the simple case of a single recessive allele, the Mut phenotype results only from the homozygous *m/m* genotype. However, for dominant alleles, the Mut phenotype is found in both the heterozygote and homozygous mutant genotypes. Thus, in the case of a dominant mutation, the phenotype that should be selected is wild-type, because this results only from the homozygous *+/+* genotype. It may seem counter-intuitive to go through the entire cross and then pick entirely wild-type animals, but this is the best way to obtain unambiguous mapping results with dominant alleles. Because the wild-type allele is present in the Hawaiian genome, linkage will be detected as an increase in Hawaiian DNA (the opposite of mapping a recessive allele).

In the case of some allelic interactions, the first generation of animals segregating from a heterozygote (the F₂ generation) will not contain a phenotype associated with a unique genotype. For example, this is the case for a recessive suppressor of a recessive mutation if the suppressed phenotype is indistinguishable from wild-type. In these cases, one can often first self F₂ Mut animals and then identify suppressed segregants in the F₃ generation, which will now have a unique homozygous suppressor genotype. In other cases of allelic interac-
tions, the double homozygote has a unique phenotype in the F2. In this case the animals can be found in the F2, albeit at only one-sixteenth of the population.

A different problem arises when a phenotype can only be reliably assayed as a population. A viable way to proceed in this case is to blindly single many F2 progeny of a heterozygote and assay each of their broods. For a recessive mutation, broods that are 100% Mut should be saved and their DNA collected for SNP analysis as described in Note 10 and in Subheading 3.2.1.

3.4.2. CB4856 Background Effects

It should be noted that CB4856 is not phenotypically identical to N2 in all respects. It is currently known that CB4856 has alleles of npr-1, plg-1, and ppw-1 that are phenotypically different from N2 (18–20). In some cases the differences may be detrimental to phenotyping, and thus SNP mapping, of some mutations. For example, the npr-1 allele may interfere with many types of assays that involve observing taxis behaviors, as npr-1 mutant animals have a strong tendency to crawl at the borders of a bacterial lawn and a strong tendency to clump together (18). ppw-1 makes CB4856 animals resistant to RNAi induced by feeding, but not by dsRNA injection (20). This fact complicates mapping any phenotype that requires RNAi as part of the phenotypic assay. Likewise, the mating plug deposited by plg-1 mutants may interfere with male mating assays. It should not simply be assumed that if your phenotype would not obviously be affected by these alleles that you are safe. CB4856 may still have an interfering effect on your phenotype. For example, CB4856 animals have an altered response to ethanol because of their modified npr-1 allele (21). Further, it remains to be seen whether there are additional CB4856 polymorphisms with phenotypic consequences, a fact which should be considered if unusual phenotypes appear to arise in your assays after outcrosses to CB4856.

4. Notes

1. It is critical that the tips align perfectly in the multichannel pipettor, a problem solved by the square geometry of the LTS design. Pipettors using conical tips have been found to perform poorly for loading gels.

2. dNTPs are very sensitive to freeze–thaw cycles. They should always be stored at −20°C in a nonfrost-free freezer. The glycerol in the solution prevents multiple freeze–thaw cycles from degrading the dNTPs. Alternatively, a 10 mM dNTP stock can be made in water and frozen in single use aliquots. All enzymes and dNTPs should be maintained at less than 0°C using a benchtop freezer block when in use.

3. PCR primers can be sensitive to low pH present in ddH2O. Primer stocks should always be dissolved in buffered Tris-HCl, pH 8.0.

4. Any preferred gel-loading buffer can be used. The additional step of adding the gel loading buffer after digestion can be eliminated by adding 1 µL of 30% sucrose
plus 0.5 µL of 0.1% cresol red in the restriction digest mix. Although many restriction enzymes work after sucrose addition, it is not known whether all enzymes work as efficiently in this condition.

5. These animals can be difficult to recover from a mating plate because of the strain’s tendency to burrow, but vortexing the plate will bring worms to the surface. Alternatively, you can remove the P₀s after 24 h of mating or use agarose plates.

6. For single matings, use 3 hermaphrodites and 10 males. Use more hermaphrodites if they are difficult to mate to, or if singling the mated animals.

7. This step enriches for cross-progeny. It is not necessary unless cross-progeny cannot be distinguished from self-progeny.

8. If you have singled mated worms, clone only from those plates that have segregated approx 50% males. It is not a bad idea to clone 8–12 additional mut/CB4856 animals and put them at 15°C, just in case. If you want to collect a very large number of recombinants, clone more worms.

9. Plates not segregating Mut, and plates segregating males, should be discarded. Some phenotypes are easier to score than others. If your phenotype cannot be reliably distinguished on a mixed plate, see Note 10.

10. If you cannot score your phenotype in individual worms, clone 200 self-progeny of mut/CB4856 hermaphrodites to individual plates. Wait until you can distinguish the homozygous mutant plates; discard the rest. For chromosome mapping, multiply the number of confirmed homozygous plates by 4 µL. Into that volume of SWLB + proteinase K pick about 10 worms from each plate. This will be your Mut lysis. For your non-Mut control, put 25 N2 and 25 CB4856 worms into a tube containing 20 µL SWLB+ proteinase K. If you can distinguish Mut from non-Mut animals, single 200 Mut F₂ animals at this point for later use in interval mapping.

11. Clean pin replicator by dipping the pins, about half way up the length of the pin, in ethanol, then flaming it. Be aware that flaming alcohol can sometimes drip off of the replicator or flare upward off of the replicator. Never put your hands, face, or hair above or below the flaming replicator. Before starting, make sure to clear your work area of flammable items like paper, especially if it is hanging from walls or shelves. When the ethanol tray catches fire, do not panic. At worst the ethanol will burn out eventually if you do not spread it around. Simply cover the tray with a nonflammable item like a plate or beaker to smother the flame. Allowing the pins to cool to room temperature is critical, because primers and templates are severely damaged by hot pins. After the ethanol has burned off, waving the replicator in a broad arc speeds the cooling—just make sure you do not drop it or hit someone behind you. Pay close attention to keeping the orientation of the plates consistent, so that you will have the same alignment in your assay plate as in the PCR primer plate. Also make sure that you are placing the pins into all of the wells; it is easy to get out of register and miss the first or last row of the plate. Rocking the replicator slightly after placing it into the sample can assure that all of the pins have hit the bottom of the wells.
12. In some cases, the plates will not all reach starvation at the same time. In addition, some plates may burrow. Some degree of judgment may be required to harvest at the perfect time to optimize the yield of worms. It is usually best to wait until at least most of the plates have almost starved, even if that means many plates have been starved for a day or so. Adequate yield can easily be obtained from recently starved plates. In the case of burrowed plates, vortexing the plate for 5–10 s will usually drive most of the worms to the surface after a few minutes, allowing worms to be washed off. Yield from contaminated plates is often low. In the case of contamination of a few of the plates with slime or mold, it is often easiest to discard the plates rather than risk spreading the contamination. Minimizing contamination also aids in the storage of the plates for later reference.

References


