Worm Breeding for Super Geniuses:

A guide to genetic mapping in *C. elegans*

Written by David S. Fay, Dan Starr, Andy Spencer, and Wade Johnson
Edited by Amy Fluet and John Yochem

Why we changed the name: We were concerned that a certain publishing company would become displeased upon learning of the enormous profits we were reaping through the use of one of their trademark phrases. In fact, I write these words to you from the sun-drenched deck of my 70-foot luxury yacht (WBFD III), which straddles the little Laramie river just north of town. It's a peculiar site I suppose, given that the river is only 3 feet deep and 15 feet wide at its grandest. Also, the hull is currently encased in late-May ice as winters are especially long here...but it's not as if I mean to complain. Back to the worms….

Contents:

I. Introduction and basics
II. Two-point mapping
III. Three-point mapping
IV. Deficiency and duplication mapping
V. Basics of SNP mapping
VI. Mapping dominant mutations
VII. Mapping suppressor/enhance mutations
VIII. Mapping of synthetic mutations

I. Introduction and basics

What this guide is and isn't. This guide is neither a basic text in Mendelian genetics nor is it in any way a comprehensive description of *C. elegans* biology. Detailed information about the latter can be gleaned from ***C. elegans I*** and ***C. elegans II*** by Cold Spring Harbor Press Inc. This guide does assume a working knowledge of basic genetics and will be of limited use to those who lack some background in this area. It is our hope that this text may serve as a supplement to existing published materials and that it will facilitate the successful breeding of worms by those new to the field.

Corrections and errors. Despite our best efforts, this first version of the guide will inevitably contain mistakes. We request that any errors be directly reported to us (davidfay@uwyo.edu), and we will try to implement corrections in a timely manner. Suggestions about wording, format, presentation, and additional topics are also highly welcome.

In the beginning. Welcome to the world of *C. elegans* genetics. We use genetics in *C. elegans* for two principal purposes: (1) to positional map mutations so that the wild-type copy of the corresponding gene can be cloned; and (2) to generate strains containing multiple mutations for phenotypic analysis. This guide will mainly address the first concern, though the basic techniques will apply to both uses.

Genetics has its good points and bad points. On the good side, genetics usually works and consistently moves us closer to our goal of identifying the affected gene products from our screens. In this way it can be quite satisfying, especially if we have been both creative and successful in the process. On the down side, it can seem like a slow and arduous process , and we are often "slaves" to the developmental time clock of the worms. Moreover, even when a reasonably careful approach is taken, genetics can sometimes fail to provide a clear answer. We may generate pieces of conflicting data that must be resolved by additional experiments.

Probably the best general advice for doing *C. elegans* genetics is to **always take the "sledgehammer" approach.** The bottom line is that it usually takes only a couple of extra minutes to pick a few more animals or to set up additional plates for matings. Contrast this to the days or weeks that can be lost if sufficient animals were not picked to isolate the necessary genotype or generate cross-progeny. *To be an effective C. elegans geneticist you must consistently get things to work the first time.* Failure to do so will vastly reduce any progress. In this sense, *C. elegans* genetics is not substantially different from many other scientific disciplines. The specific problem with *C. elegans* genetics is that given the time required for worms to develop, one can lose a lot of time before discovering that the experiment has failed. Try hard to prevent this from happening to you.

Aside from taking the sledgehammer approach, what is the best way to ensure that your genetics will work the first time? **By drawing out the entire set of crosses before picking a single worm.** Bottom line: If your basic strategy is flawed, then all the experimental diligence in the world won't save you. Each genetic situation will have unique considerations. By drawing out the entire genetic flowchart, complete with all possibilities, one can nearly always guarantee a good result. *Avoid at all costs a faulty scheme. DRAW IT OUT!*

Often a nonsensical piece of genetic data is the result of either experimental error or problems with the strains used for mapping. Both of these are ultimately your responsibility to avoid. Bad or incorrect mapping strains can usually be avoided by a careful examination of the strain before beginning the mapping process. **If any inconsistencies are uncovered between the expected plate phenotypes and those observed, DO NOT SWEEP THESE UNDER THE RUG!** This is a red flag and may be telling you that your mapping strain is not as advertised. Rather than investing weeks or months of your time in trying to map with such a strain, obtain a correct version of the strain from some other source or come up with an alternative mapping strategy. Sometimes it may be difficult or impossible to know if a strain is definitely correct. To some extent we must operate on faith, and we are usually safe in doing so. However, it is always

advisable to have multiple pieces of corroborating data before moving onto either the next step in mapping or cosmid rescue injections.

Nomenclature. This is in many ways the bane of all genetics and why non-geneticists hate reading our papers. The problem is that the style and rules of nomenclature are different for all the commonly studied organisms. Moreover, unification between the fields is unlikely to ever occur as we are too entrenched in our unique notations and jargon. The general rules for *C. elegans* are described below. Additional information can be found on the *C. elegans* Web site.

Gene names. These are designated by three letters followed by a hyphen and a number. The letters and number are always italicized. The letters chosen are usually either abbreviations of a longer descriptor (such as *lin* for lineage defective or *unc* for uncoordinated) or may be acronym-like (such as *sur* for suppressor of ras). A number then follows the letters (such as *lin-31*) to indicate the approximate order in which the mutations were discovered.

Many, but not all, gene names have arisen through genetic screens where mutant alleles were isolated (see below). In some cases the actual open reading frame (ORF) compromised in these mutants may await identification. In other cases, a gene name may have been given to an ORF (or predicted ORF) for which no mutations have been identified. This most often occurs when ORFs appear to be the obvious homolog of a gene characterized in other systems or are members of large protein families.

There is something of a protocol in our field that should be followed before assigning ones favorite new mutation a novel three-letter name. First, efforts should be made to initially map the mutation, in part to prevent the assignment of a new name to a previously described mutation or gene. For a number of good reasons, it is becoming quite common now for genes to be cloned (the mutant ORF positively identified) before assigning gene names. If the gene or mutant is believed to be novel, a proposed name is submitted to the "worm name czar", Jonathan Hodgkin, who then passes sound judgment on the merits of the suggested name.

Mutant names and alleles. This is both simple and confusing: simple because the name of a mutant strain, *lin-31* for example, is the same as the name of the gene affected in this strain, *lin-31*, and confusing because when we say "*lin-31*" we may mean either the (wild-type) ORF that encodes *lin-31* or mutant *lin-31* animals. Obviously the context will specify which we mean. When we refer to a specific mutation that affects *lin-31* function, we use an allele designation. These are one or two letters (usually two) followed by a number, such as the allele *lin-31(n301)*. The letters that proceed the number are specific to each *C. elegans* lab and allow one to easily identify the origin of the mutant allele (*n* for example is the Horvitz lab's designation). Allele numbers correspond to the order in which they were identified by a given lab The nature and severity of the mutant phenotypes displayed by individual alleles of any given gene may vary greatly

Proteins. The peptide encoded by *lin-31* is LIN-31 (all caps, non-italic).

Phenotype. More complexity. When describing the phenotype of an animal, we capitalize the first letter and do not use italics. Thus animals with an *unc-4* genotype display an **Unc** phenotype. In addition, such animals may display other defects, for example they may be slightly small, or **Sma**. With time one gets to know all the major descriptors, which isn't really too onerous a task given that there are only a limited number of ways to mess up a worm.

Types of mutations

Below is a list of the most common types of mutations used for mapping. One issue to always consider is the penetrance of the allele. If the penetrance is significantly below 100%, the marker may be difficult or even impossible to use for mapping. Another issue is the ease with which the mutation can be scored by its plate phenotype. The back of *C. elegans II* provides information about specific mutations and alleles. A scale of one to three (ES1, ES2, ES3) is commonly used to designate the plate phenotype. ES3 mutations are generally very easy to score, while ES1 mutations require some sort of clairvoyance. ES2-rated mutations range from the reasonable to the ridiculous. For a more complete listing of mutants and phenotypes, see *C. elegans II* or any number of worm-related Web sites.

dpy Produces a "dumpy" (short and fat) phenotype. Different *dpy* mutants can range from severe (small footballs) to moderate in character. The more severe ones will often display a variable **Unc** phenotype as well.

unc Uncoordinated. There are many different subclasses of *unc* mutants. These include coiler **Uncs**, kinker **Uncs**, paralyzed **Uncs**, shrinker **Uncs**, **Uncs** that fail to move backwards when touched with a pick on their heads, **Uncs** that display poor forward movement but back well, etc. Recognizing certain types of **Uncs** can initially be challenging, though it usually gets easier over time. Often **Uncs** are somewhat misshapen and are typically smaller or thinner than wild-type animals.

sma Small. **Sma** animals tend to be more proportional in shape than **Dpy** animals; less fat, more like wild type.

lon Long. Animals often on the thin (stringy) side. While **Dpy** and **Sma** animals can in some cases be very small compared to wild type, even the longest **Lon** is only about 50% greater in length than wild-type animals.

egl Egg-laying defective. This can lead to the **Bag** (bag of worms) phenotype where embryos hatch within the mother leaving a cuticle sack containing multiple wriggling larvae. **Egl** animals can be recognized before bagging as animals that seem to be bloated with eggs. However, caution must be employed, as aging wild-type animals can often appear somewhat **Egl** over time. An individual bag will only last for about 24 hours on the plate. Once worms become starved, the incidence of **Egls** and **Bags** in genetically wild-type animals

increases substantially. Unambiguous identification must always be carried out on non-starved plates.

let Lethals. These can range from embryonic lethals that never hatch to lethals that die as larvae. The latter category is easier to recognize, especially when the worms display a distinct larval lethal phenotype such as a "paralyzed rod" or a severe **Dpy** phenotype. Dead or dying eggs can be difficult to distinguish from healthy wild-type eggs on first viewing with a dissecting scope. To "see" embryonic lethals, one must allow a parent to lay eggs for a set period of time (usually 3-12 hours). The parent is removed to a new plate and the fates of the eggs are followed. The presence of lethals can usually be identified unambiguously after about 18 hours (at 20ºC) when the vast majority of wild-type embryos would have already hatched. Several other designations for embryonic lethal mutations include ***emb*** and ***zyg***.

ste Sterile animals come in several varieties. The most useful for mapping are those where the sterility is obvious because the adult worms are devoid of eggs. Care must be taken to avoid mistaking a sterile animal for one that is merely a young adult that does not yet contain obvious eggs. If in doubt, transfer the suspected sterile animal to a new plate and follow its fate. In some cases, sterile animals may contain a protruding vulva (**Pvl**-sterile) which makes identification very rapid. So called "maternal-effect" lethal mutants are really just sterile animals that contain dead eggs.

rol The roller phenotype. Animals form a horseshoe shape and tragically twist in place about their long axis. The **Rol** phenotype can be masked by strong *unc* or *dpy* mutations, which prevent the animals from carrying out the classic roller moves.

bli Blister phenotype. Adult animals have a variably blistered cuticle, which can resemble a large bubble on the surface of the worm. The **Bli** phenotype can be suppressed by a number of *dpy* and *rol* mutations.

lin Lineage defective. These can display any number of distinct phenotypes depending on the specific nature of the lineage defect (see other sources for specifics).

Genetic shorthand. There are undoubtedly numerous "correct" ways to convey genetics in writing. Some standard *C. elegans* conventions that I use throughout the text are shown in Figure 1.

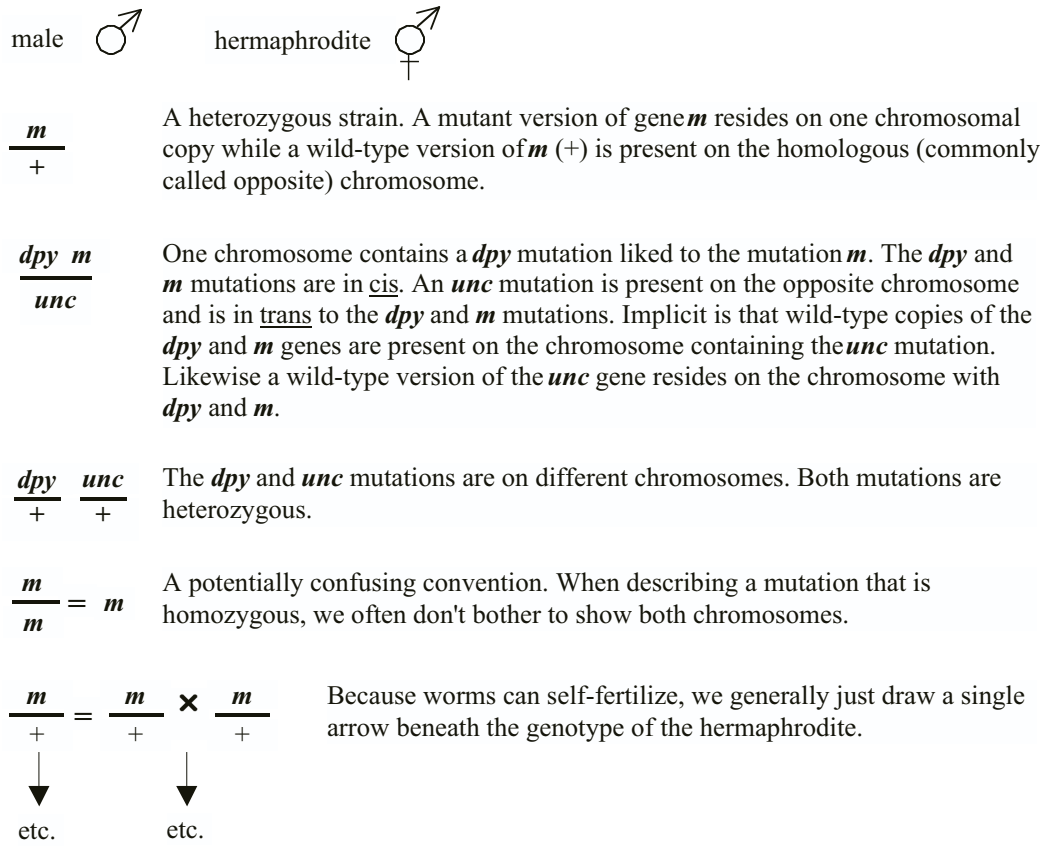


Figure 1.

Feeding, growing, and maintaining worms

Maintaining a worm stock is relatively simple. Worms are generally grown on NGM plates containing the bacterial (*E. coli*) strain known as OP50. They crawl around the plate, eat off the bacterial lawn, and reproduce. The plates are secured with a rubber band and are stored upside down to prevent them from drying out. Usually worms are grown at either 15ºC or 20ºC. It generally takes about three days at 20ºC for a fertile adult to develop from a one-cell embryo. At 15ºC this process takes about twice as long, and varying the incubation temperature (between 15ºC and 20ºC) is pretty much the only way to control the rate of worm growth and development. Higher temperatures (20ºC-25ºC) can further expedite the rate of development but can cause a drop in fertility and poor health, especially in certain mutant backgrounds. Temperatures >25ºC are usually harmful and under normal circumstances should be avoided.

Embryogenesis itself normally takes about 14-16 hours at 20ºC. This is followed by four larval stages during which the majority of growth occurs. Wild-type worms at 20ºC will begin manufacturing and laying eggs 3-4 days into their life cycle and will produce on average 200 or more self-fertilized progeny. After about two generations, the OP50 bacteria will be completely consumed and the worms will become starved. Starvation in worms does not have the same

connotation as it might in other organisms. Worms are tough and can survive without food for a period of time. They do this in part by forming "dauer" larvae, which are dark and thin and often lie motionless. Neglected worms can survive for up to several months provided the plates do not become badly contaminated or dry out. Wrapping plates in Parafilm and storing at 10ºC to 15ºC can help to increase long-term survival rates. However taking a lackadaisical attitude is not to be encouraged. It is important to be highly vigilant, especially with precious strains, in order to prevent loss.

Avoid contamination! There are two general types of contamination, bacterial and fungal/mold. Though the fungus (generally a fuzzy growth) may appear especially sinister and will require a fairly rapid response, it is the easier of the two to get rid of. Normally a fungus can be defeated by transferring animals to a clean plate, and then moving them to a second clean plate after several minutes or an hour. Bacterial infestations occur when strains other than OP50 colonize the plate. Getting rid of bacteria can be problematic. This is because the worms have been eating the stuff and it's in their intestines. The only way to get rid of a nasty bacterial infestation is to dissolve gravid (fertile, egg-containing) worms in a mixture of sodium hypochlorite (bleach) and sodium hydroxide (see *C. elegans I*), which will kill everything but the internal eggs, which are protected by their chitin shell.

Contamination will come from three sources: 1) the plates themselves may contain the dreaded exploding "footballs" or some other unwanted microbe; 2) the OP50 used to spot the plates may itself become contaminated; and 3) air-born nasties, which are usually of the fungal or mold-like variety, can fly onto your plates. Obviously, one wants to do everything possible to avoid using inherently bad plates. It is advisable to let plates sit out for a day or two before spotting with OP50 so that contaminants can manifest themselves. In our experience, we will go through periods where plates are questionable, and this is usually due to an inexperienced or careless work study student who does the pouring. Bad OP50 is another source of contamination. Often this is due to lack of proper sterile technique. Always inoculate liquid LB cultures by picking OP50 colonies from a reasonably fresh LB plate. Never inoculate a new OP50 liquid culture from a preexisting OP50 liquid stock. This will nearly ALWAYS lead to contaminants. To avoid fungal infestations, keep plates covered whenever possible. **Basically, use good sense and be meticulous about your plate pouring and spotting techniques.** A bad contamination can literally ruin an experiment, cause undo hassles, or at the very least make the work far less pleasant.

Maintaining a worm stock can be significantly more difficult if the strain is not "balanced". Roughly speaking, a balanced strain is one that contains distinct mutations on each copy of a particular chromosome. Balancing a mutation is usually only an issue if the mutation causes lethality or sterility when it is homozygous. A sterile mutation, for example, could be balanced by a set of ***dpy*** and ***unc*** mutations on the homologous chromosome. Usually the best configuration for balancing is when the markers are close together and flank the mutation that needs to be balanced. This decreases the likelihood that the mutation will be lost due to a single recombination event. Still, even having close

flanking markers does not guarantee that the strain cannot be lost over time, and diligence must be exercised during each passage of the stock to make sure that this does not occur. Other than having a homozygous mutation, the most stable situation is where the mutation is balanced over a chromosomal translocation or deficiency. In this case, the" balancer" chromosome is homozygous lethal and prevents recombination from occurring in the region of the mutation.

Worm chromosomes

C. elegans has a total of six chromosomes, which include five autosomes (I-V) and an X chromosome. Hermaphrodites are diploid for all six; males are diploid for the autosomes but haploid for X (X/Ø). A variety of visible markers for mapping exists on all six chromosomes. Markers are distributed throughout the chromosomes, but there is a markedly higher density occurring in the central region of each autosome. For this reason (and others) it is generally easier to map and clone mutations that reside in the central or "cluster" regions of the autosomes. (For more details, see *C. elegans I* and *II*.)

The genetic distance separating two genes is determined by the frequency of meiotic recombination that takes place between them. During meiosis each pair of homologous chromosomes will experience at least one recombination event. The nearer the two genes are to each other, the less likely that recombination will have occurred between them. One map unit (1.0) is equal to a 1% meiotic recombination frequency. In other words, if **1%** of the gametes (sperm or oocyte) coming from a double-mutant hermaphrodite have undergone a recombination event causing a separation of these mutations, the mutations (genes) are then considered to be **1.0** map units apart.

Note that recombination frequency has been reported to change somewhat with temperature and age of the parent. While the frequency of meiotic recombination does not substantially vary between 16ºC and 20ºC, rates increase significantly at temperatures greater than 20ºC and decrease at temperatures below 15ºC. Therefore one should maintain worms at temperatures where recombination rates will be consistent during the course of an experiment. As animals age, recombination frequencies decrease, though this variability may be hard to control during most mapping procedures.

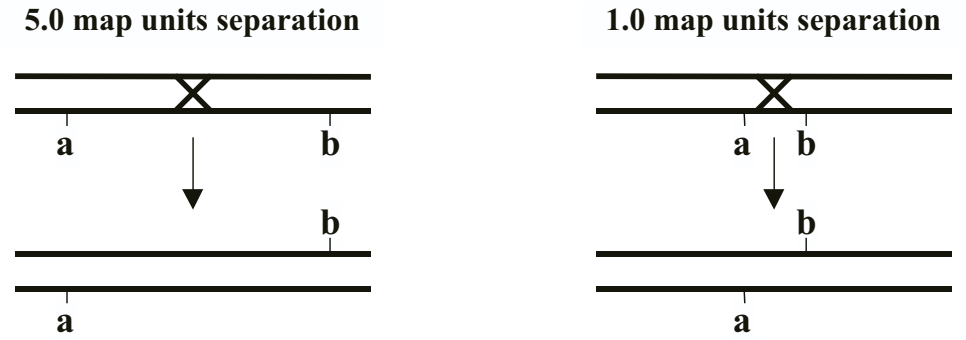


Figure 2.

In the examples shown in Figure 2, the recombination event on the left will occur in 5% of the gametes while the one on the right in 1%. Both lead to the mutations becoming genetically (and physically) unlinked from each other. Most chromosomes are on average about 50 map units long. **This means that mutations on opposite ends of a chromosome will appear genetically to be unlinked, since they will be separated during meiosis 50% of the time.**. The clusters or gene-rich regions in the center of chromosomes usually span a distance of about 5-8 map units.

One thing you will hear about is the concept of "genetic" versus "physical" distances. As we have seen, genetic distance is based on the frequency of meiotic recombination between two genes. Physical distance is the amount of DNA between them. Although the arrangement of the genes on the genetic map always agrees with the arrangement on the physical maps, the distances may not correlate. This is because the frequency of meiotic recombination is not uniform along the physical chromosome. Sometimes fairly small physical regions can be quite large genetically while large physical regions can be relatively small genetically. This can be an important factor when deciding how much mapping to carry out before attempting cosmid rescue experiments.

Setting up matings/crosses

Getting matings to work is one of the most critical aspects of successful genetic mapping. To begin with, all matings will require males. Unfortunately, males occur at only a low frequency (~0.02%) in wild-type populations. Therefore anyone doing serious genetics will maintain his or her own stock of males by placing about a dozen N2 males on a plate with several N2 hermaphrodites. Usually several plates are kept going, and the process is repeated every few days. There are several general things to keep in mind here. 1) Do not use old hermaphrodites! They are past their prime and will not work. The best hermaphrodites to use are very young adults that have few or no eggs. It is better to use an L4 than an aging gravid adult. 2) Males should also be on the young side (though this is somewhat less critical). 3) Matings will usually work best if the bacterial spot is not too large and does not contact the edge of the plate. 4) If you are in desperation, it is permissible to set up matings with animals that may be somewhat starved. Males seem to recover quite rapidly once placed on plates with food, and hermaphrodites also do reasonably well, provided they are picked as L4s or very young adults.

In some cases it may be advantageous to have a male stock for a homozygous strain. The males can first be generated following a mild heat shock (34ºC for 3-4 hours). Once several males are obtained, these can be mated back into the homozygous mutant strain. As for all male stocks, you must remember to set up crosses regularly. Should your homozygous male stock become contaminated, transfer several dozen males and hermaphrodites to a single plate, incubate overnight, and hypochlorite treat the hermaphrodites the next day.

Another source for male animals comes from ***him*** (high incidence of male) mutant strains. These are mutants that normally produce 20-40% males due to defects in segregation of the X chromosome. *him-5* and *him-8* strains are most popular for this use. One downside to using *him* mutants is that the strain you generate may itself contain the *him* mutation. Depending on your intended use for this strain, it may not be convenient to have your strain throwing large numbers of male self-progeny. Another potential source for generating homozygous males is through the inactivation of *him* genes by RNAi.

Beginning then with a stock of male animals, you will be able to set up matings between mapping strains and your mutants. There are always two ways to go here, as shown in Figure 3. You can first cross N2 animals into your mapping strain, and then mate the male cross-progeny obtained into your mutant strain (scheme #1); or you can first cross N2 males to your mutants, and then mate the male cross-progeny into your mapping strain (scheme #2).

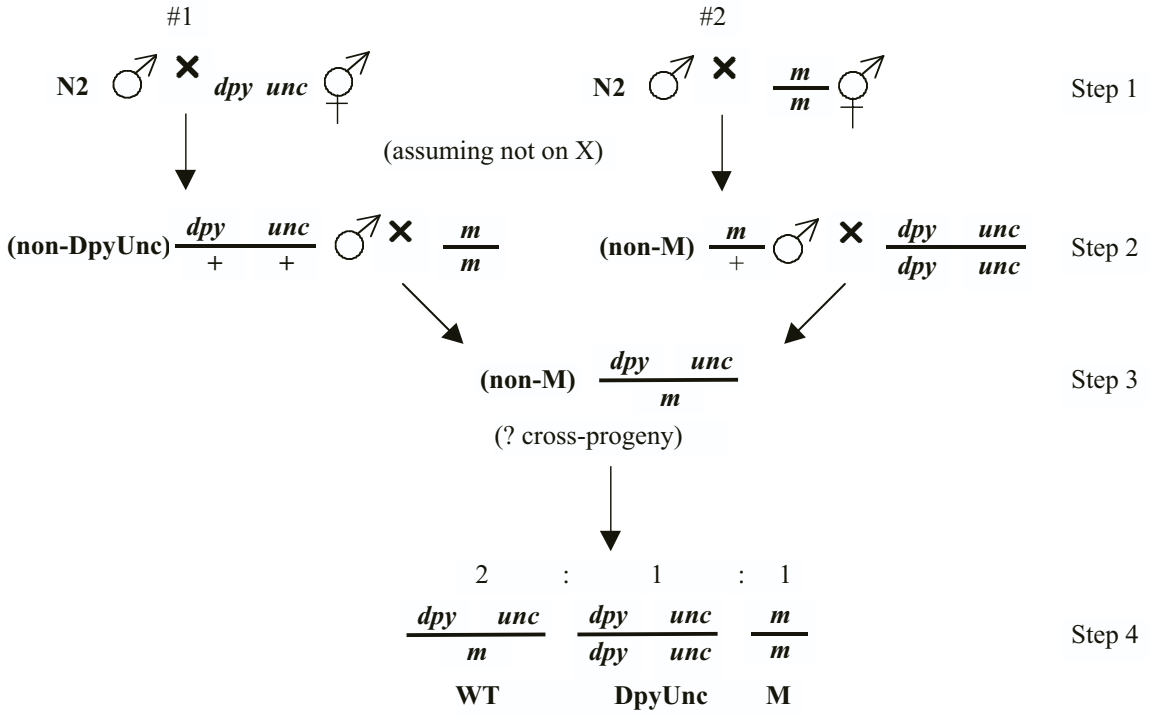


Figure 3.

The way you choose to do this will depend on several factors. For example, if your mutation or markers are on the X chromosome, then cross-progeny males (***dpy unc*/Ø** or ***m*/Ø**; generated in step 2) may be incapable of mating since they will be mutant. Usually though, it is most convenient to mate into the mapping strain last. This is because cross-progeny from these animals should be easy to identify while cross-progeny from ***m*/*m*** animals (especially if ***m*** has to be maintained as a het ***m*/+**) will usually be more difficult. The basic goal is

to minimize the amount of cross-progeny that you will have to pick "blindly" (also see below) thereby tilting the odds in your favor. If, however, your mutant phenotype is difficult to score easily by its plate phenotype, scheme #1 would be preferable since all cross-progeny generated in step 3 will be guaranteed to be heterozygous for ***m***.

A few more comments about setting up matings:

1) As already stated, take the sledgehammer approach! Having too many males is not a problem. Having too few males is a big problem! Having too many cross-progeny is not a problem. Having too few cross-progeny can be a big problem! When setting up matings with strains that normally have low brood sizes such as **DpyUncs** adopt the more-the-merrier philosophy. For such matings you can put 15 males on a plate with an equal number of **DpyUnc** animals. Since you will be picking out **non-DpyUnc** cross-progeny, you need not worry much about the plates starving too quickly as the wild-type cross-progeny will develop very rapidly compared to the **DpyUnc** self-progeny.

2) For many matings it will be extremely important that you **DO NOT inadvertently carry over any larvae or eggs from the male plate**. Contamination of this type can quickly destroy a series of genetic crosses and if not detected can lead to erroneous conclusions. Better to first pick the males needed to a fresh plate, let them crawl around briefly, and then re-pick these "clean" males to the actual plates containing the hermaphrodites.

A few comments about picking cross-progeny:

1) It is good practice to always choose virgin hermaphrodites when picking among your candidate cross-progeny animals. For some situations this may be more critical than others. However, the idea is that you usually want to see what the self-progeny of this virgin animal will segregate and don't want to complicate matters by having additional genotypes present. The safest way to do this is to pick cross-progeny hermaphrodites at the L4 stage. Whether or not an animal was a virgin can also be determined later by looking for the presence of males in the progeny. If present, the animal was obviously not a virgin, and you may want to discard such a plate in favor of one that displays the desired phenotypes but does not contain male animals.

2) When given a choice, pick cross-progeny animals from multiple plates where the mating has appeared to go well. For some situations, not every male will carry the chromosome that we desire them to contribute to the cross-progeny. When looking at cross-progeny on the plate it is impossible to tell if they happen to be the spawn of one (lucky) male or many. However, the odds that we will pick cross-progeny that include the desired genotype end up in our favor if we pick from multiple plates. This is a further reason to set up multiple mating plates and to have a generous number of males on each mating plate. Things get chancy if we have to put all our eggs in one basket.

3) Do not carry over contaminating larvae or eggs with your picked cross-progeny (see above).

4) Pick more candidate cross-progeny animals than you think are necessary. If you expect 25% of the cross-progeny animals to be of the correct genotype, pick at least 20-40 animals anyway. Some may not be true cross-progeny. Some will crawl up the side of the plate and desiccate. Some may be damaged by picking. Odds may defy you. We have all had experiences where we pick 50 animals, expect to get at least 12 of the correct genotype, and wind up getting one! In this case we are glad we picked 50! **Picking a few more animals takes little time. Setting up the whole set of crosses again takes much time.**

II. Two-point mapping

Two-point mapping is principally used to assign mutations to a specific chromosome. It can also give at least a rough indication as to the distance between the mutation and the markers used. On the surface, the practice of two-point mapping to determine chromosomal linkage is relatively straightforward. However, it can be the source of some confusion when one begins to try to interpret the data and determine actual distances based on recombination frequencies. We will try to consider the simple cases first and then move on to more complex scenarios.

In carrying out two-point mapping, one can use marker chromosomes that contain either single or double mutations. Often we will use mapping chromosomes with two markers as the strains generated can potentially be used later for three-point mapping. The two most basic scenarios are shown in Figure 4. In scheme #1, the chromosomal configuration if the mutation happens to lie on the same/homologous chromosome as the markers. In this case it is flanked by the markers and is essentially balanced by them. The genotypes of the progeny are indicated along with the ratios (or fractions) of their occurrences.

In scheme #1, three genotypes are generated (***m*/*a b*, *m*/*m*, *a b*/*a b***) with three corresponding phenotypes (**wild type, M, A B**). In this situation, however, we essentially never see the appearance of the triple mutant phenotype **MAB**. Furthermore, if we were to pick animals of phenotype **M** and examine their self-progeny, we would never see **M A B** animals. Likewise, **A B** animals will also fail to segregate **M A B** progeny. Finally, wild-type animals will always throw both **M** and **A B** animals along with wild type. Seeing segregation patterns of this type tells us that ***m*** and ***a b*** reside on the same chromosome and that ***m*** resides close to or in between the markers ***a*** and ***b***. the pattern also tells us that we have a reasonably well-balanced strain that can be used to maintain our mutation. By isolating wild-type segregants, we have a good chance that they are balanced heterozygotes. In addition this strain can be used for three-point mapping (see below).

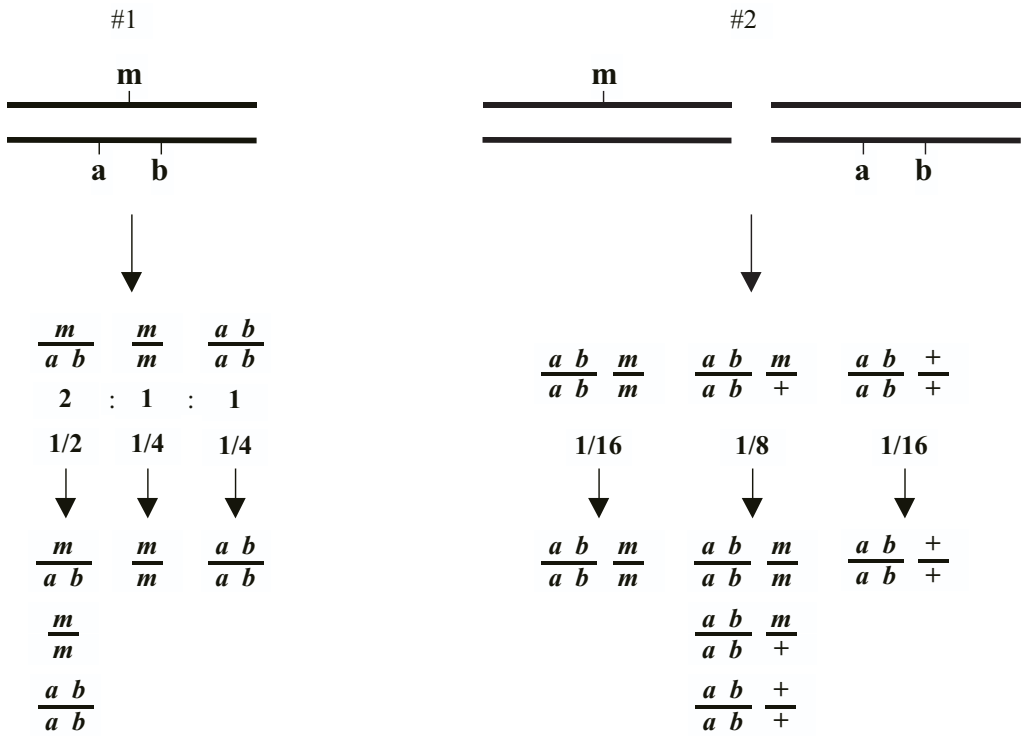


Figure 4.

In contrast, the situation depicted in scheme #2 shows ***m*** and ***a b*** on distinct chromosomes. In the first generation we therefore already expect to see one-sixteenth of the progeny displaying the triple mutant phenotype **M A B**. In addition, if we pick **A B** animals of this generation, two-thirds will throw **M A B** progeny. If necessary, draw out all the possible genotypes and corresponding phenotypes to convince yourself that these numbers are correct. Observing these kinds of segregation patterns indicates that the mutation and the markers are on different chromosomes. Another possibility is that the mutation resides on one of the ends of the chromosome (see below). If necessary, these two possibilities can usually be resolved by scoring more animals. **In general, basing linkage designation on a small number of data points (<20) should be avoided.**

The genetic patterns described above are for the ideal situation where there is no ambiguity in the determination of chromosomal location. But what happens when the mutation lies to one side of the markers, perhaps at some distance?

As shown in Figure 5, if the mutation lies to one side, a crossover may occur that will lead to the creation of the two recombinant chromosomes shown. One now contains all three mutations while the other is completely wild type. Also shown are the genotypes occurring when such a recombinant chromosome is paired with one of the parental chromosomes. Now we have a situation where an animal of phenotype **M** or **AB** can throw **MAB** animals. In addition, a wild-type animal can now fail to throw both **M** and **AB** animals.

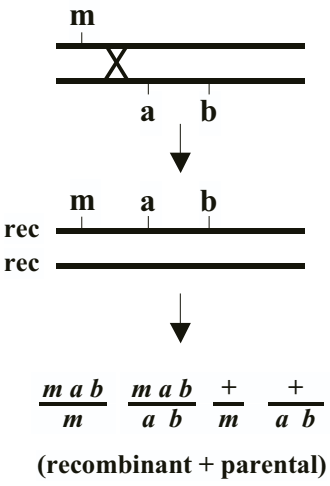


Figure 5

In these situations we must be careful since on the surface one might conclude that the presence of such genotypes would indicate that ***m*** and ***a b*** are on separate chromosomes. However it turns out to be a matter of frequency. For example, if ***m*** and ***a b*** are 10.0 map units apart, this means that such a recombination event will occur 10% of the time. To put it another way, 10% of the gametes will be recombinant for this region. Worms are of course diploid and progeny therefore have a chance to receive such a recombinant chromosome from either the sperm or the oocyte. Given this distance, the frequency with which progeny will inherit two non-recombinant (also called parental) chromosomes is 0.9 × 0.9 = 0.81 or 81%. The chance of progeny receiving two recombinant chromosomes will be quite small, in this case 0.1 × 0.1 = 1%. However the frequency of progeny receiving one recombinant and one non-recombinant chromosome is 100 - 81 - 1 = 18% . A significant fraction!

How then do we determine if a mutation is really on the same chromosome as the markers, and if so, what is the distance? This depends in part on how we are doing the mapping. Let us consider one specific example of mapping a sterile (***ste***) mutation relative to an ***unc*** mutation. In the example given in Figure 6, the ***dpy*** and ***unc*** mutations are 10.0 map units apart. Again, this means that 90% of the gamete chromosomes will be of the parental type and 10% will be recombinant. As just stated, the chance of a progeny receiving two non-recombinant chromosomes will be 81%, two recombinant chromosomes will be 1%, and one recombinant plus one non-recombinant chromosome will be 18%.

Of the recombinant chromosomes, one-half (5%) will be ***ste unc*** and one-half (5%) will be wild type (+ +). Each recombinant chromosome has an equal chance of pairing with either of the two parental chromosomes. Therefore, for the animals that contain one recombinant and one non-recombinant chromosome, one-fourth will be ***ste unc/unc***, one-fourth ***ste unc/ste***, one-fourth ***+/unc***, and one-fourth ***+/ste***. These genotypes will therefore be present at a frequency of 0.25 × 0.18 = 0.045 or 4.5% each.

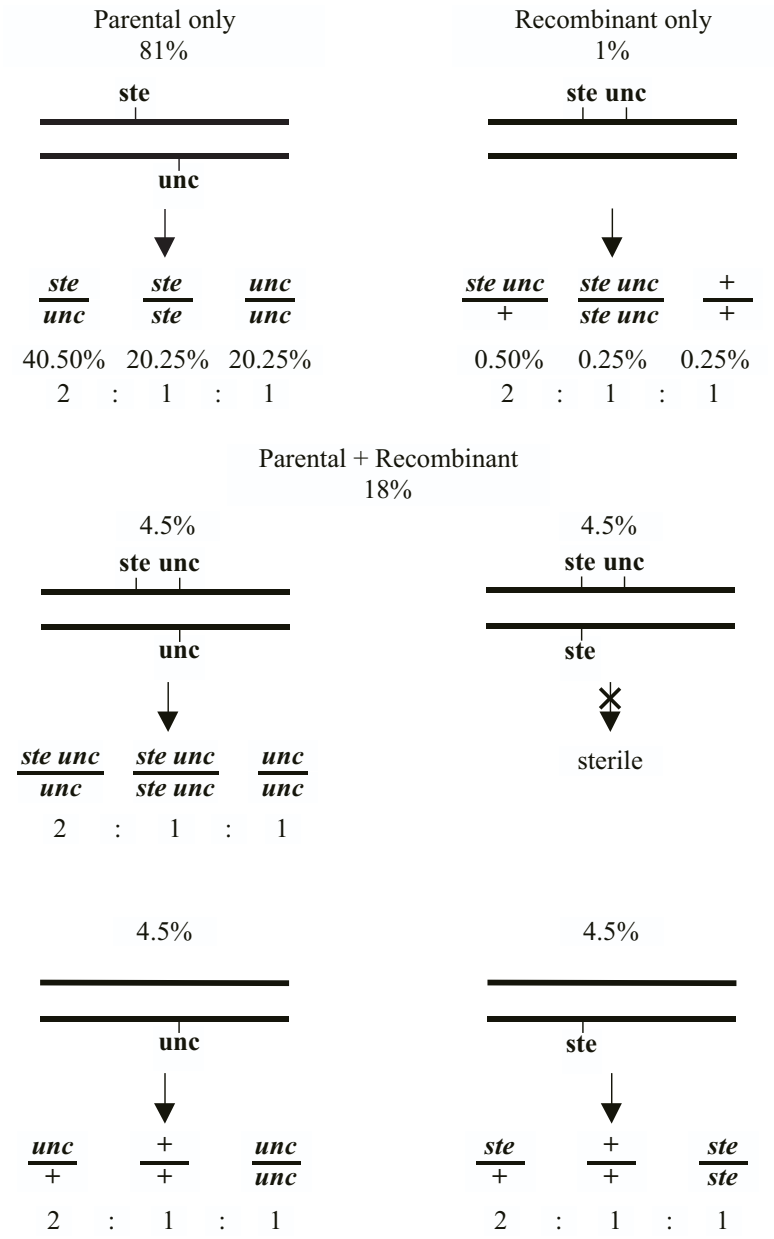


Figure 6.

Now consider mapping in the following way. From plates where the parent is ***ste/unc***, we clone **Unc** progeny. We want to determine the frequency by which such **Unc** animals throw **Ste Unc** versus **Unc** only progeny. We therefore look for the presence of **Ste Unc** animals in the next generation. We know that there will be two genotypic possibilities for animals with an **Unc** phenotype, ***unc/unc***, where both chromosomes are parental, and ***ste unc/unc***, where we have one of each. The percentage of animals with the ***unc/unc*** genotype is 0.81 × 0.25 = 0.2025 (20.25%) since 81% will have only parental chromosomes and of these, one-fourth will receive two ***unc*** chromosomes. The percentage with a ***ste unc/unc*** genotype will be 0.18 × 0.25 = 0.045 (4.5%) since 18% of progeny will have one recombinant and one parental chromosome and there is a 25% chance of receiving both the ***ste unc*** and the ***unc*** chromosome (0.5 × 0.5 = 0.25). The overall

percentage of animals with an **Unc** phenotype will therefore be 4.5 + 20.25 = 24.75%. Finally, the percentage of **Unc** animals with a ***ste unc/unc*** genotype will be 4.5/24.75 = 18.2%.

The above determination tells us that if our mutation and marker(s) are 10.0 map units apart, we should expect to see about 18% of the **Uncs** cloned throwing **Ste Unc** progeny. Similar calculations can be carried out for various distances. If the marker and mutation are 1.0 map unit apart, we will see **Ste Unc** animals appearing from ~ 2% of the cloned **Uncs**. At 5.0 map units apart it will be ~ 9.5%; at 25.0 map units, ~40%. **When the mutation is very close to the marker, the frequency of animals containing the recombinant chromosome will be double that of the map distance between the marker and the mutation. As the distance between the mutation and marker increases, this factor decreases.**

By the time we get to 50.0 map units, 67% or two-thirds of **Unc** animals will throw **Ste Unc** progeny. This latter number should sound familiar. It's the same percentage you would get if the ***ste*** and ***unc*** mutations were on separate chromosomes. **In fact, at 50.0 map units or greater, two mutations will appear to be unlinked.** This usually is not an issue since we tend to carry out two-point mapping with markers at the chromosome center, guaranteeing distances no greater than about 25.0 map units.

There are often multiple ways to carry out two-point mapping using the same set of markers. For example, in the previously described cross we could have picked wild-type rather than **Unc** animals and looked for the absence of either **Unc** or **Ste** animals in their progeny, signifying a + + or wild-type recombinant chromosome. If the marker and mutation are 10.0 map units apart, we will predict to have 0.81 × 0.5 = 40.5% of animals with an ***unc/ste*** genotype. We will also have 0.18 × 0.5 = 9% of animals with either an ***unc/+*** or ***ste/+*** genotype (4.5% each). Thus we predict that 9.0/49.5 = 18.2% of the wild-type animals we pick will fail to segregate either **Unc** or **Ste** progeny. These numbers are identical to those previously calculated for picking **Unc** progeny and looking for **Ste Unc** in the next generation.

Consider, however, this final case. Imagine you are trying to map an embryonic lethal mutation (***emb***) relative to a known ***unc***. The easiest way to do this would be to pick wild-type animals from an ***emb/unc*** parent and then look for the absence of **Unc** animals in the progeny (embryonic lethals are usually difficult or impossible to score directly by their plate phenotype). If the ***unc*** and ***emb*** are on the same chromosome and close, very few phenotypically wild-type animals will fail to throw **Unc** (as well as **Emb**) progeny. To calculate the map distance, however, we must realize that we did not count ***unc/+*** animals as "recombinants". For example, if the distance between the ***unc*** and ***emb*** is 10.0 map units, we predict to have 0.81 × 0.5 = 40.5% animals of genotype ***emb/unc***. We will also have 0.18 × 0.25 = 4.5% of animals with an ***emb/+*** genotype and 4.5% with an ***unc/+*** genotype. Therefore, when picking among the phenotypically wild-type animals, the frequency of ***emb/+*** animals will be 4.5/(40.5+4.5+4.5) =

9% and not 18.2% of the total. **Being aware of these factors and, as always, drawing out the cross carefully will prevent interpretive errors.**

A question of strategy: To map all at once or sequentially? This may depend on several factors such as time constraints and competitive pressures. Everything being equal, mapping sequentially is the most efficient allocation of time since once one has positively identified a chromosomal location, one need not check all the other chromosomes. In practice though, we often want to map our mutant as quickly as possible and will test multiple chromosomes at once. In addition, the presence of clear negative data can strengthen conclusions when the mutation lies at some distance from the markers.

III. Three-point mapping

Once you have assigned your mutation to a chromosome, it is time to begin three-point mapping. Three-point mapping is pretty much the backbone of worm genetics and is essential for us to identify our mutant genes. Even SNP (single nucleotide polymorphism) mapping (see below) is really just a high-tech variation on classical three-point mapping. The basic idea is that we cross our mutant strain into a strain with two markers, ***a*** and ***b***. We follow two classes of progeny: those with phenotype **A** and those with phenotype **B**. By seeing which of these two classes also produce the mutant phenotype (**M**) in the subsequent generation, we can determine whether our mutation lies to the left, right, or in between our set of markers. In the case where the mutation lies in between, we may then determine the approximate distance from each marker.

Figure 7 depicts the outcome of a recombination between markers ***a*** and ***b*** when ***m*** lies either to the left or right of the markers. As can be seen, when ***m*** lies to the left essentially all **B-non-A** recombinant animals will throw **B M** progeny (as well as **B** and **A B**), while **A-non-B** recombinant animals will only throw **A** and **A B** progeny. When we see this kind of pattern, we can conclude that ***m*** lies to the left of ***a*** or perhaps to the right of ***a*** but very close. The reason for this is that if ***m*** were very close to ***a***, but between ***a*** and ***b***, the frequency of generating the ***a m*** recombinant chromosome would be very low (see below). Thus while ***m*** is most likely to the left of ***a***, we often have this caveat. Greater numbers of recombinants can help to diminish this possibility, if not rule it out completely. The situation for ***m*** lying to the right is simply the reverse.

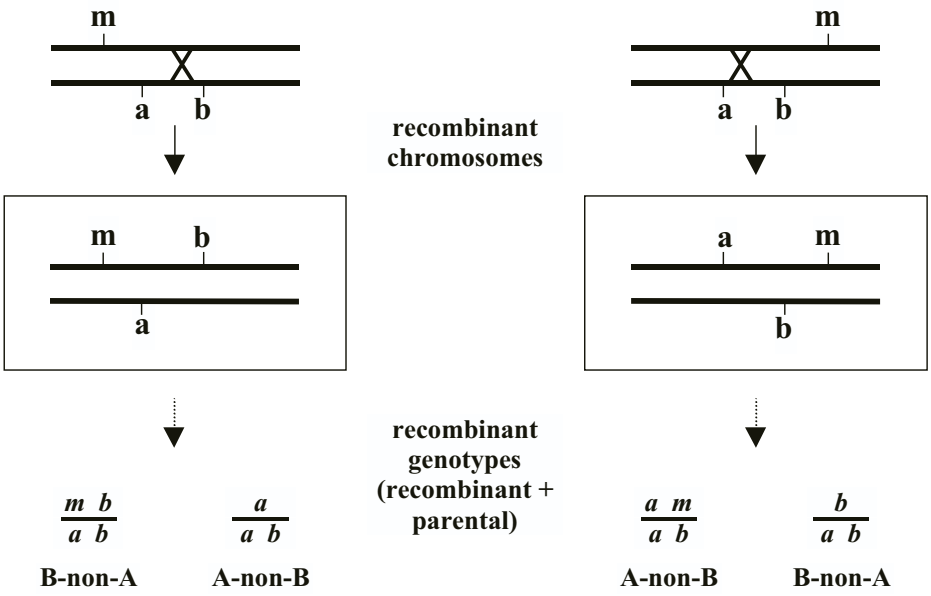


Figure 7.

The mapping described above, though useful, only tells us that ***m*** is likely to be left or right of our given markers. It doesn't provide any information about how far from these markers ***m*** might reside. To determine this we need to use markers that flank ***m*** as shown in Figure 8.

Here we see that depending on the site of the cross-over, **A-non-B** recombinant animals can in some cases acquire ***m*** (#1) and in other cases not (#2). The same is true for **B-non-A** animals. In three-point mapping we seek to determine the ratio of recombinant animals that pick up the mutation versus those that do not. This ratio provides us with a direct chromosomal position for the mutation as illustrated in Figure 9.

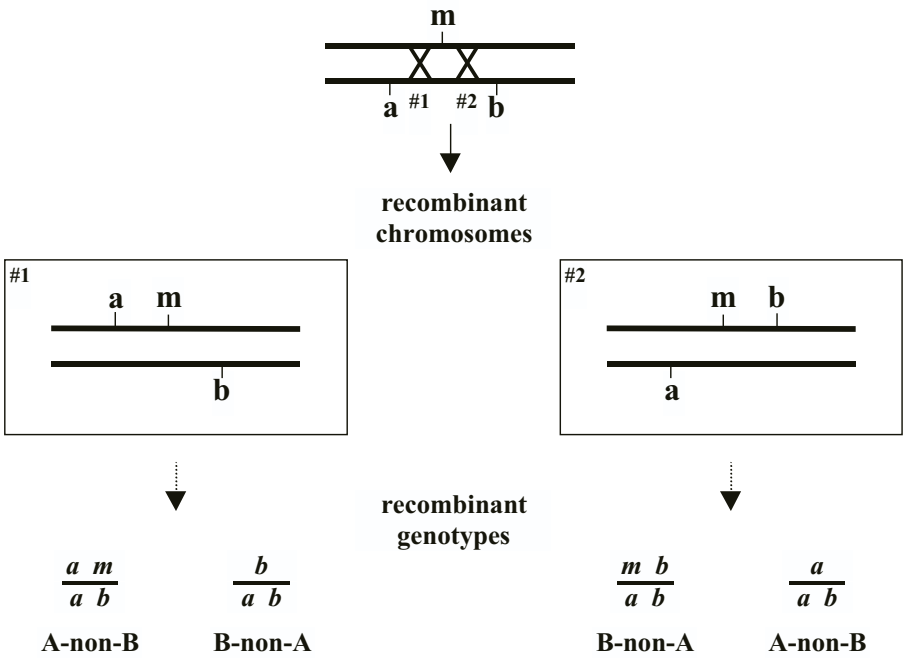


Figure 8.

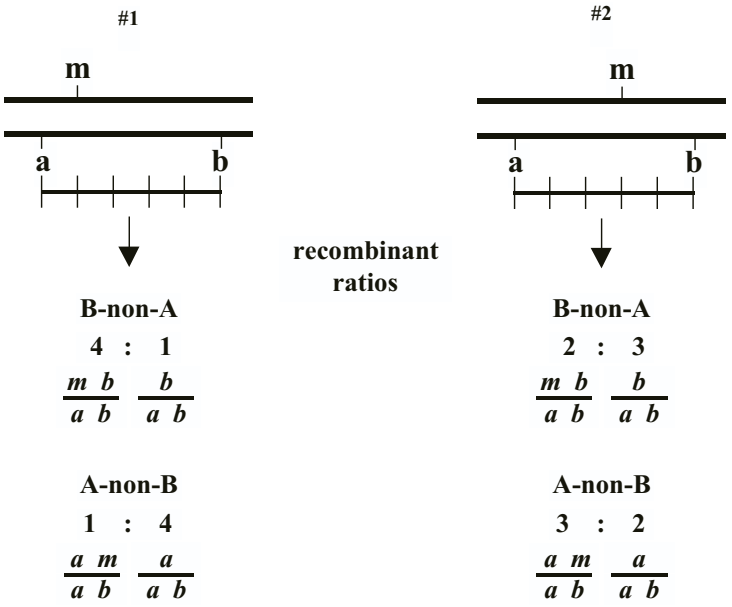


Figure 9.

Markers ***a*** and ***b*** are in cis and located 5.0 map units apart, while our mutation, ***m***, is in trans to ***a*** and ***b***. In the situation on the left, were we to pick **B-**

non-A recombinant animals, four-fifths or 80% will now have ***m*** in cis to ***b***. **A-non-B** recombinants, on the other hand, would acquire ***m*** only one-fifth or 20% of the time. On the right, **B-non-A** animals will acquire ***m*** only 40% of the time, while **A-non-B** animals will acquire it 60% of the time. Obviously, when picking recombinants from both sides, the numbers should converge on a single location, i.e., the frequencies should add up to 100%. These numbers can be used to specifically assign a genetic location. For example, in the left diagram, if ***a*** were at genetic position 0.0 on the chromosome and ***b*** at 5.0, having 20% of **A-non-B** recombinants acquire ***m*** would lead to a map position assignment of 1.0. Obviously, the greater the number of recombinants scored, the greater the certainty of the assignment.

Always save recombinants; they often prove very useful for subsequent mapping, not to mention genetic studies where having a linked marker may prove indispensable. Figure 10 shows an example of how to use the recombinant chromosome for further mapping (also see Deficiency Mapping).

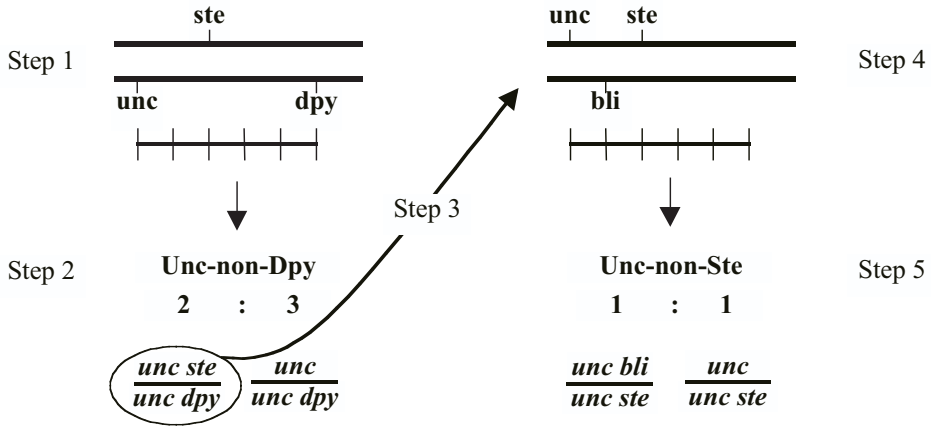


Figure 10.

Imagine we are mapping a ***ste*** mutation and have placed it between ***unc*** and ***dpy*** markers that are separated by 5.0 map units (step 1). The ratios place the ***ste*** mutation closer to the ***unc*** marker (10 out of 25 **Unc-non-Dpy** recombinant animals threw **Unc Ste** progeny; step 2). We save the ***unc ste/unc dpy*** strain and cross it to a strain that is homozygous for a ***bli*** mutation (step 3). We obtain the strain shown in step 4 and then screen for **Unc-non-Ste** animals (step 5). In this case, 50% of the **Unc-non-Ste** recombinants acquired the ***bli*** mutation, placing ***ste*** and ***unc*** mutations at an equal distance (but on opposite sides) from ***bli***.

In this way, we continue to refine the map position of our mutation. Usually the data from different mapping schemes will tend to agree, though not always. This may be due to a number of factors. **However, the farther apart the markers are, the less precise the mapping generally tends to be.** Thus we put more weight on data acquired using nearby markers than those that are at some distance. **In addition, it is highly advisable to map using markers that have**

already been cloned. This provides a precise chromosomal location and allows one to directly compare the genetic and physical maps. If you have no choice but to use a non-cloned mutant for mapping purposes, check the database or journals for information regarding how this gene was mapped to its present location.

Double recombinations, where two recombination events have occurred in the same region of the chromosome are relatively rare events. However, when using markers far apart, this may become an issue. For this reason, it is generally wise to stick to markers that are <5.0 map units apart when doing three-point mapping. Even doing so, however, does not guarantee some low frequency of double recombinants. This is especially true if large numbers of recombinants are scored, thereby increasing the likelihood of such an event. Always be aware of this possibility and refine your interpretations if necessary.

A word of caution. What happens if you initially map your mutation to the wrong chromosome and then try to carry out three-point mapping? Essentially, your mutation segregates independently of the recombinant chromosome and will be picked up two-thirds of the time. Thus, if for example, 67% of your **Dpy-non-Unc** and **Unc-non-Dpy** animals throw your mutation, you may want to consider redoing the two-point mapping.

Finding and picking recombinants. At the most basic level, two things should be anticipated in advance of picking recombinants for mapping: 1) the expected frequency of recombinants; and 2) the plate phenotype(s) of the recombinant animals. The first concern is relatively easy to calculate. Since you should know the distance between the two genetic markers, the frequency of recombination events between these markers can be directly determined. For example, if the two markers ***a*** and ***b*** are 2.0 map units apart, there will be a 2% chance of a recombination event between ***a*** and ***b***. Since hermaphrodite worms are diploid for all chromosomes, this effectively doubles the chance of acquiring a recombinant chromosome in the progeny. However, to detect the recombinant, it must be over the 'correct' parental chromosome, which will occur only 50% of the time.

For example, if the parental animal has the phenotype ***m/a b***, where ***a*** and ***b*** are markers and ***m*** is the mutation to be mapped, then recombinants that give ***a*** alone or ***b*** alone (with or without ***m***) would need to be over the parental ***a b*** chromosome in order to actually see the **A-non-B** or **B-non-A** phenotypes. The end result is that if one is looking specifically for **A-non-B** recombinants, and ***a*** and ***b*** are 2.0 map units apart, then an animal with an **A-non-B** phenotype will occur on average about 1% of the time. Likewise, **B-non-A** animals will occur 1% of the time. Obviously, if the mapping allows picking of either **A-non-B** or **B-non-A** non-recombinants, this will effectively double the total number of recombinant animals that can be obtained from a given number of plates.

As with all genetics, it is wiser to pick more worms than is anticipated to be necessary. The rate-limiting step for all genetics is growth of the animals and

not the time required to transfer a few more to plates. Still, there is a limit to how much information can be gleaned from any one cross.

The next step is to recognize and pick the recombinant animals. But first it is important before picking from any plate to ask the question: **Do the animals on this plate display the expected phenotypes?** In effect, you are thereby asking: **Did the parental animal have the correct genotype?** This is exceedingly important to determine before picking any recombinants. The reason is that recombination events may have occurred in the previous generation such that the cloned parental animal may not have had the correct genotype. For example, you have picked phenotypically wild-type animals from a plate where the parental animal was of genotype ***m*/*a b***. Given that self-progeny with the genotype ***m*/*a b*** will be wild type, you might imagine that you are safe in assuming that all wild-type progeny will therefore have genotype ***m*/*a b***. But imagine the following two scenarios depicted in Figure 11.

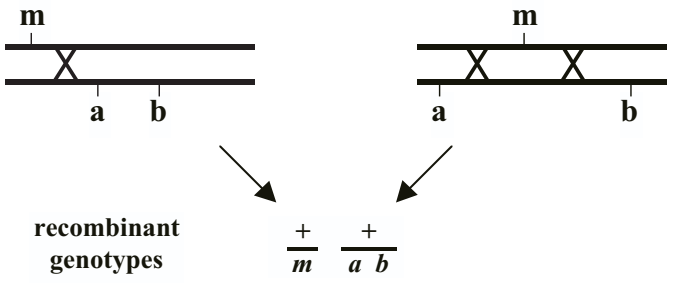


Figure 11.

In the scenario on the left, ***m*** lies to one side of the markers ***a*** and ***b***. A recombination event between the markers and ***m*** can result in the creation of a wild-type chromosome (+) as well as a triple mutant chromosome (not shown). Therefore, when the recombinant + chromosome is paired with one of the parental chromosomes, phenotypically wild-type animals would be generated with the genotype ***m*/+** or **+/*a b*** (and not the expected ***m*/*a b*)**. The same thing can happen if ***m*** is between markers ***a*** and ***b*** (as shown on the right). In this case, a double recombination must occur to generate the wild-type chromosome, which will be a relatively rare event. In addition, recombination could occur to give ***m*/*a*** or ***m*/*b*** animals, which are also phenotypically wild-type (see below).

Clearly, one does not want to pick recombinants from plates where the parental animal had the incorrect genotype. This will wreak havoc on one's mapping and lead to incorrect conclusions. The solution is simple: **Make sure the phenotypes observed on the plate correspond to the correct parental genotype.** Generally this is quite simple in practice. For example, if the parental animal has the expected genotype ***m*/*a b***, then one should see wild-type animals (***m*/*a b***), **M** animals (***m*/*m***), and **AB** animals (***a b*/*a b***). In addition, it should be possible to find occasional recombinant animals (**A-non-B** and **B-non-A**), which is exactly what you are looking for.

While simple in practice, it is easy to make errors. For example, consider using ***dpy*** and ***unc*** markers. Some strains of ***dpy*** animals may appear somewhat **Unc**, while some strains of ***unc*** animals may appear somewhat **Dpy.** Thus, if one is not careful, it could be possible to lose the ***dpy*** or ***unc*** marker without immediate knowledge of its loss. In the end, strict diligence is the only weapon against such mistakes. Other markers such as ***let*** and ***egl*** may require even greater care to maintain. **Bottom line: Do whatever you consider necessary to insure that recombinants are obtained only from plates with the correct parental genotype.**

How many recombinants should one pick from any given plate? This may depend on several factors. However, as a rule, **be very cautious of plates where you seem to have hit a "gold mine"!** ("Wow, I can get all 20 recombinants off of one plate!" NOT.) The simplest explanation when encountering such a plate is that a recombination event must have occurred in the previous generation to affect the parent. This is precisely the situation that was described above. Looking at such plates it will likely be clear that the parent animal did not have the correct genotype. In this case it is permissible to pick a single recombinant animal, since this does represent one legitimate recombination event.

Even in cases where most animals correspond to the non-recombinant phenotypes (indicating that a parental recombination event did not occur), **it is still advisable to pick only 2-3 recombinant progeny from any one plate.** The worry is that a rare mitotic recombination event may have occurred in the distal (mitotic) region of the gonad to generate a clone of identical recombinant oocytes.

Often when looking for recombinants to pick, one will examine the same set of plates for several days in a row. It is a common experience that recombinants that are "invisible" one day will jump out at you the next. Certainly for some types of mutants such as ***ste*** or ***egl***, the recombinant phenotype may only be obvious once animals are well into adulthood. When scanning the same set of plates over several days, **keep whatever notes necessary to insure that you don't keep picking your recombinants off the same plate** without knowing it. Proper note taking and labeling of plates will prevent this from happening.

Recognizing the recombinants that you want may not be trivial! Or it may be, depending on the nature of the mutant phenotypes and your level of experience. For example, you acquire a ***dpy unc*** strain for mapping purposes. The double-mutant animals indeed look both **Dpy** and **Unc**, but what will the **Dpy-non-Unc** or the **Unc-non-Dpy** recombinant animals actually look like? Often one does not have either the ***dpy*** or ***unc*** mutation alone for comparison. The problem is that ***dpy*** mutants may be somewhat **Unc** and ***unc*** mutants may be somewhat **Dpy** or **Sma** (small). In the absence of having the single-mutant strains available for comparison, the best approach is to read up on the descriptions of the mutant phenotypes in the back of *C. elegans II* and to ask others in the lab who may have worked with these mutations for advice. Once you have isolated a few true recombinants, finding new ones will suddenly get much easier.

What if you think you have picked a recombinant animal and it turns out to be of a non-recombinant genotype? This turns out not to be a problem as it will be obvious when looking at such a plate that a recombinant was not picked. For example, if you believe you have picked a **Dpy-non-Unc** animal and notice several days later that the "recombinant" worm has failed to throw appreciable numbers of **Dpy-non-Unc** animals, or is perhaps throwing phenotypically wild-type animals, obviously the parental animal was not a true recombinant. Chuck the plate and move on. **It is better to pick some false recombinants (and eliminate them later) than to miss picking any true recombinants.**

A note of caution: **Make sure that when picking recombinants, you do not carry over contaminating eggs or larvae!** This is surprisingly easy to do and will usually ruin your ability to score that particular recombinant since the plate will be contaminated with animals of non-recombinant phenotypes. If the plate is crowded, move the recombinant animals to a less populated region of the plate in order to "clean" the recombinant animal of larvae or eggs that may have stuck to its side. Sometimes it may even be necessary to transfer the recombinant to a "clean-up" plate before cloning to its own plate. As a second line of defense, always watch the recombinant animal after transferring it to its own plate and destroy any contaminating eggs or larvae that may come off. Such procedures become second nature very quickly.

Some other considerations to keep in mind: The **Bli** (blister) phenotype is often masked (suppressed) by ***dpy*** and ***rol*** mutations; ***unc*** mutations may mask the **Rol** phenotype; ***dpy*** mutations will usually mask a **Lon** (long) phenotype; certain ***dpy*** and ***unc*** mutations may sometimes appear **Egl**, etc. Obviously, there may be a lot to consider and going into the mapping well informed is the best weapon. Surprisingly, one can sometimes map with mutations that would seem unlikely. For example, it may be possible to identify certain **Unc*X*-non-UncY** mutants, depending on the nature of the two **Unc** phenotypes.

Obviously, the **best type of three-point mapping would allow for picking of recombinants from "both" directions.** For example, one can pick **Unc-non-Dpy** and **Dpy-non-Unc** recombinant animals from a strain with a ***dpy unc*** chromosome. The benefit of this setup is that it effectively doubles the number of recombinants obtained from a given number of plates, and provides independent mapping information from both types of recombinants (which will hopefully correspond!). However, it is not uncommon that the markers will dictate that only one of two possible recombinants is picked. For example, when looking for recombinants between ***unc*** and ***let*** (lethal) mutations, it will only be possible to identify and pick **Unc-non-Lets** for obvious reasons. The same thing occurs when picking recombinants with sterile mutations, in situations where one of the mutations normally masks the other, or where penetrance is an issue (see above).

Reconciling the physical and genetic maps. By the end of three-point mapping you will hopefully have data that will allow you to undertake cosmid-

rescue injections. At this point you will want to find ground zero on the physical map, the stretch of DNA where your mutation is predicted to reside. One simple way to do this is to first construct a complete graphic of the physical map for the region, a process that may require some actual cutting and pasting of a printout of the map. One then takes actual physical measurements of the distances between the markers used (e.g., 135 mm) and, based on the obtained mapping data, finds the point on the chromosome that has been implicated. Note: When generating this map, do not use AceDB, as this program does not currently draw to scale regions containing cosmid gaps. The SNP database currently provides a good graphic for this purpose.

Alternatively, one can avoid the graphic map entirely by calculating the predicted site of the mutation based on the numerical chromosomal locations of the markers used (e.g., 9,279,450). This is quite straightforward and eliminates potential errors associated with any graphic representation. Detailed information for this purpose can be found on the Wormbase Web site. Note that while it may be satisfying to point to a single base pair on the chromosome as being the most likely site of one's mutation, this prediction is only as good as the mapping data. Furthermore, regional variations in recombination frequency along the chromosome will lead to discrepancies between the actual locations and those predicted using these methods. Still, this will allow you to compile a list of likely rescuing cosmids and prioritize your order of injections.

IV. Mapping with deficiencies and duplications

Deficiency (***Df***) mapping works great...when it works. The problem in mapping with ***Dfs*** is that while positive results are generally unambiguous, negative results can be more difficult to nail down. Deficiencies refer to specific deleted regions within chromosomes. The sizes of ***Dfs*** vary greatly from just a few cosmids wide to the absence of a large portion of the chromosome. The end-points of the ***Df*** may have been determined precisely using molecular techniques or may be rough guesses based on genetic tests with various mutations. Homozygous ***Df*** animals are almost always embryonic lethals as removal of multiple genes usually includes some that are necessary during early development. The basics of ***Df*** mapping strategy are shown in Figure 12.

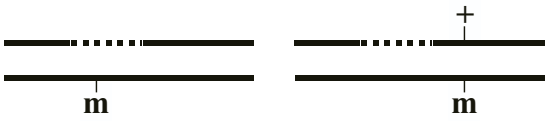


Figure 12.

In the case on the left, the mutation lies within the deficiency (dashed line) and is therefore not rescued by the corresponding wild-type gene on the opposite chromosome. Animals with such a genetic configuration will generally show the mutant (**M**) phenotype. The exception to this is when the mutation is a hypomorph (partial loss-of-function) and a 50% reduction in the gene dosage leads to a new "**M**" phenotype that is more severe than the original phenotype displayed by ***m/m*** animals. In the case on the right, the mutation is outside the ***Df*** and the **M** phenotype will not be displayed. If the breakpoint of the ***Df*** in this example is

near ***m***, one has a balanced heterozygote that could be useful, but keep in mind that recombination can occur between the right breakpoint and ***m***, thereby destroying the balanced stock.

The way most deficiency mapping is done is as follows: You set up a situation where you are looking for a mutant phenotype in the F1 generation of the cross. For example, as shown in Figure 13, your mutation-which is linked to an ***unc*** or some other visible marker in cis-is crossed to male animals (step 2) that were created by mating N2 males into a balanced deficiency strain (step 1; provided ***Df*** is not on X). In this case the ***unc*** mutation linked to your mutation is known to be outside this particular ***Df***.

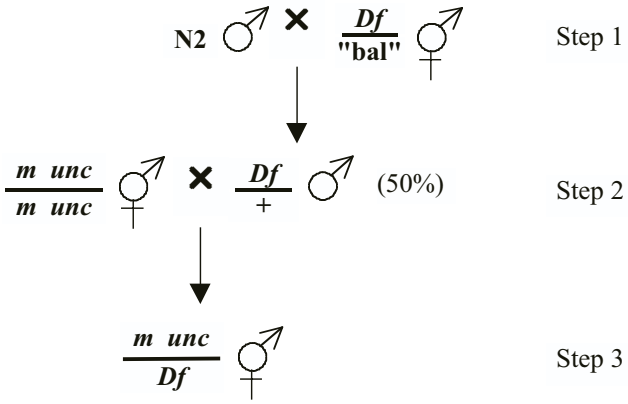


Figure 13.

In this scenario, if the mutation is within the ***Df***, you will observe **non-Unc** animals displaying the **M** phenotype in the F1 generation (step 3). The presence of the linked ***unc*** marker is necessary here to clearly identify cross-progeny. For ***Df*** mapping, it is important to set up as many mating plates as possible (10-15 is a reasonable number) in order to guarantee generation and detection of the ***m***/***Df*** genotype.

Failure to observe the **M** phenotype would indicate that ***m*** is outside the ***Df***, provided that on most of your mating plates you observe good numbers of **non-Unc** cross-progeny. As a test, though, it is best to clone a good number of supposed cross-progeny and to make certain some of them throw progeny that include both the **M** and **Df** (usually embryonic lethal) phenotypes.

As an alternative approach, shown in Figure 14, you can also mate your mutation via the male into the ***Df*** strain (step 2) and look for your **M** phenotype in the first generation of cross-progeny (step 3).

In this case it is best to have the linked ***unc*** mutation (or other cis marker) inside the **Df** so that cross-progeny of the desired phenotype can be easily identified via the **Unc** phenotype. This approach can also be used without any linked marker where one just looks directly for the appearance of the **M**

phenotype in the F1 generation. As with the previous example, if the result appears negative (***m*** is outside the ***Df***), it is important to try and verify this by cloning supposed F1 cross-progeny and making sure that some of the phenotypically wild-type F1s throw progeny of both **M** and **Df** phenotypes.

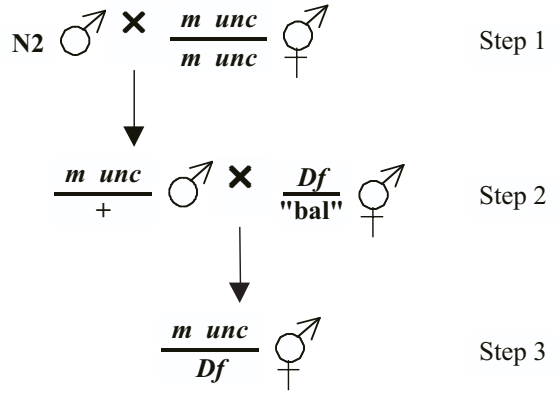


Figure 14.

Mapping with Duplications (***Dps***) is done less frequently than ***Dfs*** and is probably of less utility. Free duplications are autonomous pieces of DNA derived from normal chromosomes. They are usually relatively small compared to full-length chromosomes and exhibit segregation properties that are independent of other chromosomes, including the chromosome from which they were derived. In many ways they most resemble extra-chromosomal arrays and, like arrays, tend to be significantly less stable (especially meiotically) than normal chromosomes. ***Dps*** will vary significantly in their genetic stability, and some published information exists describing the properties of various ***Dps***. Animals that contain a ***Dp*** will effectively be triploid for the genes that lie within the ***Dp***. ***Dps*** are often used to balance a homozygous lethal mutation.

The idea in mapping with free ***Dps*** is to determine whether or not the mutation of interest lies within the duplicated region. If it does, then one would observe "rescue" of the mutant phenotype. To do this, one will set up crosses ultimately leading to the isolation of the genotype shown in Figure 15.

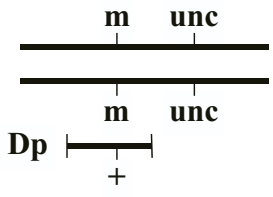


Figure 15.

This can in theory be accomplished using several approaches but usually takes a number of steps as one must re-homozygose animals for ***m***. Having a marker linked to ***m*** but outside the region covered by the ***Dp*** can be useful for identifying candidate (***m***/***m***) animals. As with ***Dfs***, it will generally be quite obvious when one's mutation lies within the ***Dp*** but rather more difficult to prove that it definitely lies outside. Creating a chromosome where the mutation is flanked by two visible markers (one outside and one inside the ***Dp***) can help to clarify this issue.

V. Single nucleotide polymorphism (SNP) mapping (Dan Starr)

Often times, all of the above approaches will fail despite the great advice of Dr. Fay!!! It may be due to a lack of good markers in your region, or you may get unlucky and fail to rescue your mutation by cosmid or YAC injections

despite all the "perfect" mapping data in the world. At this point you might want to think about SNP mapping. **In fact you might want to think about SNP mapping much earlier in the mapping process.**

Typically, SNP mapping is undertaken after your mutation has been mapped to a chromosomal subregion and physically linked to nearby markers. New techniques have recently been developed that utilize SNP mapping at a much earlier stage in the mapping process. Although SNP mapping is a little more work than traditional three-point mapping, due to the large number of recombinants you need to pick, it is a very powerful technique, which leads to a much more precise location of your mutant. The conscientious worm breeder should have no problem mapping his or her mutation from a 2-map-unit region down to a 100-kb region (or even down to a <10-kb region in some cases) with absolute boundaries in about a month.

What is a SNP? How do you find one?

SNPs are **s**ingle **n**ucleotide **p**olymorphisms. The polymorphisms can be single nucleotide changes (for example from an A to a G) or very small deletions/insertions between two divergent populations of "wild-type" *C. elegans* worms. They are usually found in non-coding regions of the genome.

The N2 wild-type strain and all the mutant strains we use are derived from a single hermaphrodite found in the dirt in Bristol England back in the early 1970's. A "wild-type" worm found in your backyard might not have a common ancestor with the Bristol N2 worm within the past million years or more. During this long period of isolation between these two populations, many mutations have accumulated to make these populations genetically different. Some mutations were harmful and were likely selected against. Rare mutations were beneficial and made the worm population better suited for its local environment thousands of miles away from England; these mutations were easily selected for. However, most mutations are in non-coding regions and are "silent" so that they do not change the fitness of the worm. The frequency of a silent mutation within a population can then randomly drift to fixation within a population. Now when you sequence a short stretch of genomic DNA from an N2 worm and compare it to a worm from your backyard, you can detect SNPs that are fixed within each population but differ between the two populations.

Previously, the most time-consuming aspect of SNP mapping was identifying SNPs. However, in the post-genomic era, this is a breeze. In fact for the most part it has already been done for you. The Genome Project has an excellent Web site (http://genome.wustl.edu/gsc/C_elegans/SNP/index.html) detailing the location of many SNPs. They have compiled the data form about 19,000 shotgun reads of the divergent "wild-type" strain CB4856 from Hawaii and compared these sequences to the N2 wild-type strain used for the whole-genome sequencing project. They have found thousands of SNPs spread throughout the genome; by checking their site it is simple to identify a SNP in the genome region you are interested in. They are finding SNPs at a rate of about 1/1000. This means that if you need to find your own SNP in a particular region,

simply sequence a kilobase or so of non-coding genomic DNA from the Hawaiian strain and you are likely to find a SNP. (I have found this approach to be successful in finding a SNP about 80% of the time.)

In addition, they have done *in silico* restriction digests and identified RFLPs (restriction fragment length polymorphisms; check out your intro genetics textbook) for most SNPs (also known as "snip SNPs"). **I highly recommend that you chose a SNP with a RFLP as it is much easier, faster, and cheaper to detect a RFLP in a recombinant worm than it is to sequence the worm's DNA.**

One major caveat for SNPs described on the Web site is that **many of them may not be real!** The nature of shotgun sequencing means that such SNPs have not been confirmed through repeated reads and may be due to sequencing errors or mistakes with the interpretation of the data. In any case, a probability number is listed for each SNP. Probabilities <50% are extremely dubious, but all SNP predictions should be treated with suspicion. *Whatever you do, don't waste precious samples until you have confirmed with controls that the predicted SNPs are in fact real and work in your hands!*

Using SNPs for traditional three-point mapping.

Mapping with SNPs uses the same theory as described above for traditional three-point mapping. However, you are always using the SNP as the third middle marker because it obviously doesn't have a plate phenotype. Of course the greatest advantages of SNPs are that they occur so frequently and their precise physical locations are known.

You can use SNP mapping to map your mutant from either side, or in some cases from both sides at the same time (see examples at end of section). If you are planning to map from only one side, or one side at a time, you need a strain with your mutation linked to an obvious marker that is easy to work with (such as a ***dpy*** or an ***unc***; preferably less than 1-2 map units away). Hopefully, you have frozen away such a strain in the course of traditional three-point mapping as described above. To map from both sides simultaneously, your mutation needs to be linked on both sides to good markers. If your mutation doesn't have an obvious plate phenotype, mapping from both sides simultaneously is preferable. I describe here the relatively simple case of SNP mapping from one side by picking recombinants between a ***dpy*** and your mutant (***m***) as shown in Figure 16. In theory, mapping from the other side or from both sides simultaneously is analogous.

In this example, your ***dpy m*** strain (in the N2 genomic background-represented by thin lines in the figure) is crossed to the divergent wild-type strain CB4856 from Hawaii (represented by bold lines in the figure) to make the starting F1 heterozygotes. These heterozygotes are then allowed to self-fertilize. About 75% of the resulting F2 progeny will appear wild type, and about 25% will be **Dpy M**; pick the rare **Dpy-non-M** recombinant (or **M-non-Dpy** if it's easy to pick). You will need to pick a lot of these rare recombinants, perhaps 100 is a good starting point (see examples below). Therefore, you will probably want

to plate out at least 100 plates with single F1 heterozygotes. Be careful of "jackpot" plates as described previously. Also, if you need more recombinant lines, keep making fresh F1 heterozygotes by repeating the original mating so as not to accidentally pick an extremely rare double recombinant. After picking **Dpy-non-M** recombinants you need to homozygose the recombinant chromosome (see above). Once you homozygose a recombinant chromosome, you are ready to look at SNPs to determine where the recombination event took place.

Since you picked **Dpy-non-M** recombinants, by definition, the recombination occurred between ***dpy*** and ***m***. By using SNP markers in the region, you can narrowly define where the recombination event occurred. In the example in Figure 16, SNPs are used to map the recombination area into one of four regions.

Using the first SNP. The first SNP (in this particular example, a deletion of an A in CB4856) is used to greatly narrow down the location of ***m***. (Note that while Figure 16 shows all the SNPs in this example, one should choose SNPs sequentially based on mapping information from the previous SNP.) It will tell you if the recombination event occurred in region 1 versus regions 2 through 4. For each of your **Dpy-non-M** recombinants, check to see if the first SNP is N2 (GAT) or CB4856 (G-T). If it's CB4856 (G-T), then the recombination event occurred in region 1 between ***dpy*** and the first SNP. This recombinant strain is no longer useful and can be tossed. (However, if all the recombinants are CB4856 for the first SNP, it is likely that the first SNP is past ***m***, and you will need to analyze all the recombinant lines using a SNP much closer to ***dpy***.) If the first SNP is N2, then the recombination event must have occurred to the right of the SNP, in regions 2-4, between the SNP and ***m***. **This is your informative class** and will be saved to check further SNPs to refine the position of ***m***.

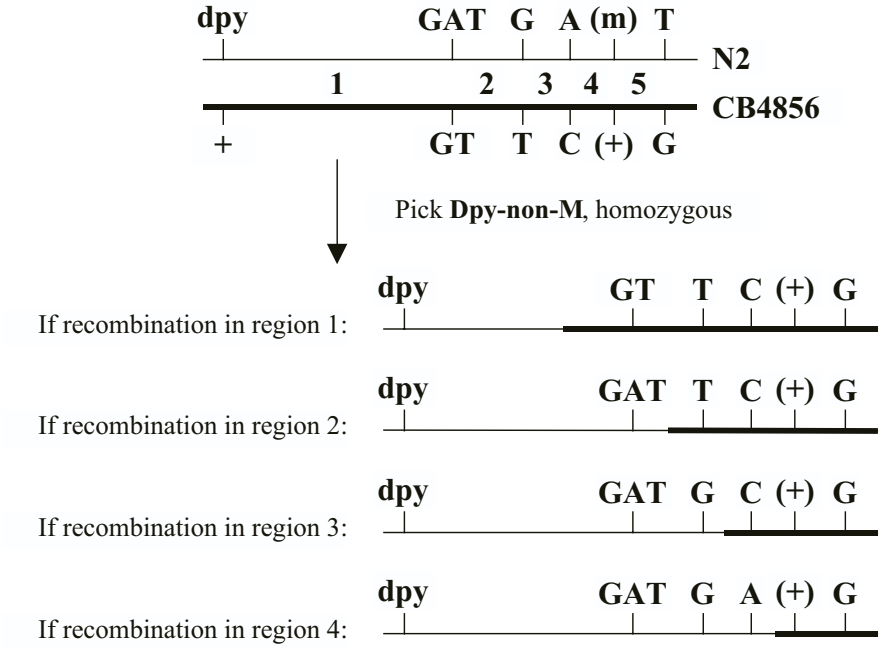


Figure 16.

You should choose the region of the first SNP based on the quality of your previous mapping data. You want a first SNP that is significantly closer to ***m*** than it is to ***dpy***. This will allow you to eliminate a large proportion of your recombinant lines in the first step, making subsequent steps much less labor intensive. Of course, be careful that you don't pick a SNP past ***m***.

Note that these data can be used to get an approximate physical position of ***m***. The ratio of the recombination events between ***dpy*** and the SNP over the number of events between the SNP and ***m*** should give you a good idea as to the location of ***m*** and help in picking the next SNP to examine.

When picking the first SNP, I highly suggest that you pick a SNP that can be detected by a RFLP. This will save both time and sequencing costs. Additionally, you won't have to maintain uninformative recombinant strains any longer than necessary. Along these lines, **test the ability to detect your SNP in both N2 and CB4856 animals before starting to collect recombinants.** This point can't be stressed enough. Maintaining upwards of 100 recombinant lines is not trivial, and you will want to narrow this number down as quickly as possible. You do not want to maintain these stocks while working out PCR problems. Additionally, testing the detection of the SNP before starting ensures that you are using the proper CB4856 strain, since on the plate CB4856 animals appear identical to N2 animals.

Using subsequent SNPs. After narrowing down the location of ***m*** using the first SNP, it is analogous to further narrow down the location with additional SNPs. Simply take all the lines that were N2 for the first SNP and check the second SNP. If it is CB4856, the recombination occurred between the two SNPs, and it isn't informative (again assuming the second SNP isn't past ***m***). If the SNP is N2, the recombination occurred to the right of the SNP, between the SNP and ***m***. These are the informative lines. Take them to check with a third SNP, and so on, to narrow the region of ***m***. Eventually, you will find a SNP for which all the recombinant lines are CB4856. This SNP is very close or past ***m*** (in the example above, the right-most SNP). How close to ***m*** you get depends on the number of recombinant lines you collect and, to some extent, how lucky you are.

This procedure keeps giving you an absolute boundary for the left-most position of ***m***. To get a boundary from the other side, you will need to SNP map from that side or map from both sides simultaneously as suggested above.

How many recombinants should I pick? Two case studies.

As with any genetic mapping, there is no set rule to answer this question; however, it can be simply put that more recombinants will lead to better mapping of your mutation. The following examples should give you some idea.

I used SNP mapping to map ***anc-1*** to an absolute region between 2 SNPs that are about 100 kb apart. Furthermore, I still had four recombinants left within

this region, so I could have mapped it down to an even smaller physical region (probably to about 40 kb). In this case I mapped from both sides simultaneously. I had a triple mutant (***unc-73, anc-1, dpy5***) and crossed that with the Hawaiian strain CB4856 to make heterozygotes, and from their progeny I picked **Dpy-non-Unc** and **Unc-non-Dpy** recombinants. ***unc-73*** and ***dpy-5*** are 1.85 map units apart, and I picked and homozygosed 89 recombinants.

The first report of SNP mapping in *C. elegans* was used to map an allele of ***cdf-1*** between two SNPs a mere 9.6 kb apart (Jakubowski and Kornfeld, 1999, *Genetics*, 153:743-752). These authors also mapped from both sides by crossing ***lon-2, cdf-1, unc-6*** to a different divergent wild-type strain (RC301) and picking 201 **Lon-non-Unc** recombinants. ***lon-2*** and ***unc-6*** are 4.1 map units apart, but they already knew ***cdf-1*** was much closer to ***unc-6***. In this case, 201 recombinants might sound like a bit of an overkill, but they eliminated 140 of the recombinants after looking at the first SNP, so what they really were looking at was 61 recombinants in an approximately 1-map-unit region used to map the allele within a 9.6-kb physical region (which may or may not be overkill depending on the specifics of your gene; they had to sequence the whole 9.6 kb region).

VI. Mapping dominant mutations (Andy Spencer)

Dominant genes were first described by Mendel to account for the patterns he observed with respect to flower color. For example, red flowers (encoded by **R**) were said to be "dominant to" white flowers because a single copy of the red gene (genotype **RR**, **Rr**, or **rR**) resulted in red flowers. In contrast, white flowers were observed only when the red alleles were **rr** (i.e., white was "recessive to" red). As the study of genetics has matured since Mendel's time, the common definition of a *dominant allele* has come to mean one whose mutant phenotype is observed when a single copy of the mutant allele is present in an animal.

A good example of a dominant allele in *C. elegans* is the ***rol-6(su1006)*** allele, which causes a "roller" (**Rol**) phenotype. ***rol-6(su1006)*** animals exhibit the **Rol** phenotype when they are of the following genotypes: ***rol-6***/***rol-6***; ***rol-6***/+; or +/***rol-6***.

Because the **Rol** phenotype is observed when a single mutant copy of ***rol-6(su1006)*** is present, the ***rol-6(su1006)*** allele is said to be dominant. Keep in mind that not all alleles of a particular gene will be dominant; there are several ***rol-6*** alleles that exhibit recessive phenotypes. Dominance or recessivity are allele-specific properties. They are not gene-specific properties.

Isolating dominant alleles. Depending on the particular developmental question in which you are interested, the systematic isolation of dominant alleles may be desirable. If you decide this is the case, the isolation of dominant alleles is straightforward. While the typical genetic screen in *C. elegans* often aims to isolate recessive, loss-of-function alleles, as shown in Figure 17, the isolation of a

dominant mutation requires one to simply screen the F1 generation (i.e., the self-progeny of mutagenized P0 worms), as shown in Figure 18.

This is trivial. One can also isolate dominant alleles in the first screen, because dominant mutations will exhibit the mutant phenotype in the F2 generation as well as the F1 generation. In a non-clonal screen, where F2 worms on a plate will derive from several different P0 animals, one may not notice that a particular allele is dominant until outcrossing fails to eliminate the mutant phenotype in cross-progeny. When outcrossing any newly isolated mutation, one should carefully observe the genetic behavior of an allele to determine whether it is dominant or recessive.

Mapping a dominant mutation. Whether your dominant mutation was isolated on purpose or by chance, the next step will be to map it to a chromosome. You'll recall from earlier sections that recessive alleles are crossed into and then scored for segregation of the mutations and known genetic markers. However, if an allele is dominant, it is necessary to change our thinking slightly when scoring the segregation of phenotypes in mapping strains. Since we cannot determine whether an animal is heterozygous or homozygous for the dominant allele by simple observation, we use the alternative strategy of mapping the absence of our dominant allele. Once one has thought about it carefully, it can often be easier to map true dominant mutations than recessive mutations.

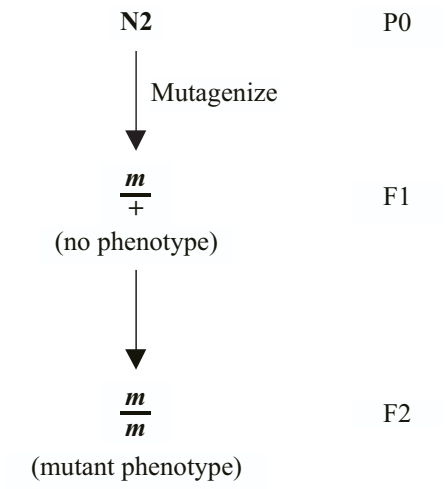
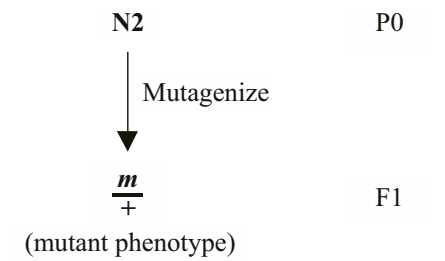


Figure 17.

The phenotype of a recessive mutation disappears when crossed into a mapping strain. Consider ***lin-1***, which causes a multivulva (**Muv**) phenotype when crossed into a strain as shown in Figure 19. We can then score the co-segregation of ***lin-1*** and ***dpy-17,unc-32*** in the normal manner by picking **Muv-non-DpyUncs** and noting how often the **DpyUnc** phenotypes co-segregate with ***lin-1*** **Muv**.

Now consider a dominant mutation that, in contrast to ***lin-1*** above, exhibits the mutant phenotype when the allele is present as a heterozygote. When we cross into our chromosomal mapping strains, all the heterozygous cross-progeny will exhibit the mutant phenotype. Let's consider an imaginary dominant mutation, ***dom-1***, as shown in Figure 20, which we'll say causes a "spiked head" phenotype. These ***dom-1***/+ heterozygotes will display the spiked head phenotype and be



indistinguishable from ***dom-1*** homozygotes. Thus, we will be unable to score the segregation of ***dom-1*** with ***dpy-17***, ***unc-32*** by following spiked-head animals because we won't know whether the animals are ***dom-1***/+ or ***dom-1***/***dom-1.***

The trick to mapping such true dominant mutations is to follow the animals that **do not** display the dominant phenotype. In other words, ignore the *dom-1* phenotype, and look only for those animals that are wild type. The reason for this is as follows: since we can always tell when ***dom-1*** is present due to the dominant spiked-head phenotype, we follow the absence of ***dom-1*** and note how often the markers segregate with **non-Spiked-Head** animals.

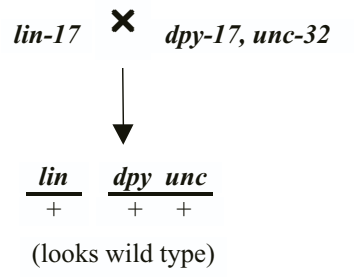


Figure 19.

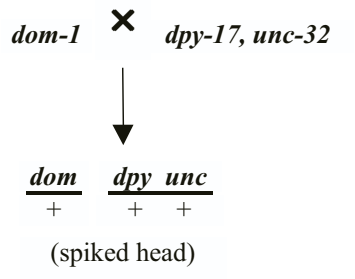


Figure 20.

The results from chromosomal mapping of dominant mutations are apparent in the F2 generation. Since you are following the absence of ***dom-1***, at this stage you are looking for **non-Spiked-Head** animals. If ***dom-1*** is on the same chromosome as and relatively close to your markers, as shown in Figure 21, case #1, then the only animals with normal heads will be the **DpyUncs,** because in the F1 generation ***dom-1*** will be over the markers. If, however, ***dom-1*** is on a different chromosome from the markers, as shown in Figure 21, case #2, then it will segregate independently from the markers and there will be both **non-DpyUnc** and **DpyUnc** animals with spiked heads. Again, all the F1s will have the spiked-head phenotype. Whether or not the **dom-1** mutation lies on the same chromosome as **dpy-17, unc-32** will be apparent in the next generation when we go to pick **non-Spiked-Head** animals. As you can see from examining the case on the left in Figure 21 where **dom-1** is on the same chromosome as **dpy-17**, **unc-32**, only **DpyUnc** animals will have normal heads. In contrast, if **dom-1** lies on a different chromosome, one-third of the animals with normal heads will be **non-DpyUnc**.

Three-point mapping of dominant mutations. Once we have our dominant mutation mapped to a chromosome, it is similarly easy to collect data for three-point mapping. We'll start again with our balanced strain that is heterozygous for both ***dom-1*** and our markers. This strain will have the spiked-head phenotype. In this example, we'll assume ***dom-1*** lies between ***dpy-17*** and ***unc-32***, as shown in Figure 22.

When looking for recombinant **Unc-non-Dpy** or **Dpy-non-Unc** animals, we will know immediately whether or not the recombinant picked up the mutant ***dom-1*** allele due to its dominance. Recombinants that pick up the ***dom-1*** allele

will have a spiked head, and recombinants that don't pick it up will have wild-type heads. Quick and easy.

Different types of dominant mutations. Why do some mutations act in a dominant fashion? Below we examine some different mechanisms through which a mutation can cause a dominant phenotype. In certain situations, different dominant alleles may require different mapping strategies. These situations must be managed on a case by case basis. In each example below, we will consider the fictional ***dom-1*** gene and imagine different situations that could give rise to various types of dominant alleles in ***dom-1***.

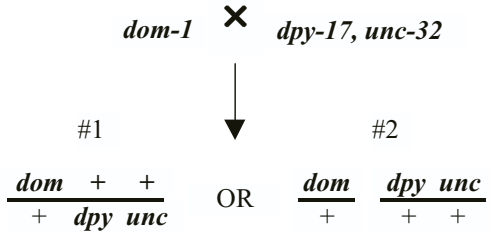


Figure 21.

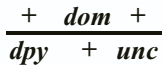


Figure 22.

Haploinsufficiency. This describes a situation in which one copy (*haplo*) of a wild-type gene is not enough to provide wild-type function when the other copy is compromised. This can only occur for loss-of-function alleles. Consider again our fictional dominant mutation, ***dom-1***. Let's assume that a certain threshold of ***dom-1*** activity is required to avoid the abnormal spiked-head phenotype, two copies of the wild-type gene are required to achieve that threshold, and any drop below that threshold allows the mutant spiked head to form. Mutations in ***dom-1*** that reduce or eliminate its activity would therefore behave dominantly because in heterozygous animals, the single remaining wild-type copy of the ***dom-1*** gene would be *insufficient* to provide the wild-type levels of gene activity. Thus, the loss-of-function ***dom-1*** mutant allele gives a similar phenotype whether present in one or two copies and behaves in a dominant fashion.

Dominant-negative alleles. These typically occur when the mutant allele does not function normally *and* the mutant protein inhibits the activity of the wild-type protein. Such a situation can result in the loss-of-function of the wild-type gene, but differs markedly from haploinsufficiency. Consider an animal that is heterozygous for a dominant-negative allele of ***dom-1***. In this case, we'll also imagine that the single wild-type copy of ***dom-1*** would normally provide enough ***dom-1*** activity to avoid the spiked-head phenotype. However, since a dominant-negative version of ***dom-1*** would actually *interfere* with the function of wild-type *dom-1*, its activity is further reduced and a mutant phenotype results.

A well-known example of a gene that can incur dominant-negative mutations is the small GTPase Ras. These dominant-negative alleles of Ras are not functional themselves because they preferentially bind GDP and stay locked in the inactive state. In addition, they also prevent the Ras exchange factor

(which binds Ras-GDP and catalyzes GDP/GTP exchange and subsequent Ras activation) from acting on wild-type Ras, essentially killing all Ras activity.

Dominant gain-of-function (gf) alleles. These can occur when a mutation results in an inappropriate level of gene activity. It is possible to imagine numerous scenarios in which the normal constraints on a protein's activity are removed. For example, a mutation in the promoter region could lead to over expression of the gene and the saturation of negative regulatory pathways. Alternatively, point mutations in a region of a gene important for its regulation could lead to inappropriate activity and mutant phenotypes. Let's revisit ***dom-1*** and imagine it is an enzyme whose activity promotes head development. Assume that normal levels of ***dom-1*** activity result in normal head development and any ***dom-1*** activity above normal levels results in a spiked head. Also assume that a negative regulatory phosphate group is added to an N-terminal serine when ***dom-1*** activity gets to the threshold required for normal development. A point mutation that makes this serine phosphorylation impossible could remove the negative regulation of ***dom-1***, allow its activity to proceed unchecked, and lead to the spiked-head phenotype. In short, too much of a good thing can lead to developmental abnormalities.

Semi-dominance. Some alleles behave in a partially dominant fashion. Alleles are designated semi-dominant when the homozygous mutant phenotype (-/-) is still observable when the allele is present as a heterozygote, but to a lesser extent. For ***dom-1***, this would be the case if ***dom-1***/***dom-1*** animals were 100% spiked head and ***dom-1***/+ animals were 60% spiked head.

I hope you enjoyed this discussion of dominant alleles. Now get back to work, dammit.

VII. Mapping suppressor and enhancer mutations (Wade Johnson)

A suppressor/enhancer screen is a classical way to uncover more information about a known mutation. Very simply, a suppressor screen starts with a known mutation and then identifies second site mutations that either suppress or enhance the mutant phenotype. One classic example is the ***let-60*** (gf) suppressor screen carried out in recent years by the Han lab. In this screen, ***let-60*** worms were mutagenized by EMS and scored for suppression of the ***let-60*** multivulval (**Muv**) mutant phenotype. Suppression of this phenotype leads to either a normal vulva or loss of the vulva, vulvalless (**Vul**). From this screen a variety of genes were isolated and characterized. Many of these genes posses silent phenotypes on their own and hence would only be isolated by such a screen. This is one of the greatest strengths of a suppressor screen. Also, all of the genes so far characterized from this screen have led to a greater understanding of the ras pathway not only in worms but in other organisms as well.

What you need to know before you get started

Before starting a suppressor screen, you obviously must have a mutation to suppress. The better characterized the mutation, the better able you will be to

design an effective screen. The type of mutation you work with will affect the type of suppressor you can hope to isolate. Suppressors fall into two classes: informational suppressors and functional suppressors. The former class includes gene products that will suppress your mutation through a generic mechanism, such as the suppression of a stop codon. Other informational suppressors include mRNA degradation mutants (smg genes), protein-degradation mutants (ubiquitination enzymes). For the most part this class of suppressors is less interesting but often cannot be avoided. Knowing the molecular lesion within your mutant of interest will allow you to determine what types of informational suppressors you might expect to uncover within your screen. The more relevant class of functional suppressors act through mechanisms that will hopefully shed light on the process or gene of interest.

The question then becomes how can one determine which mutations are informational and which are functional? By setting up a series of experiments designed to test a variety of informational suppressors specific to the mutation of interest, one can easily avoid spending too much time on these non-specific mutations. A variety of experiments could be designed to answer this question; a few are listed below. These are not necessarily the only experiments that could be done, but they are a place to start. Each mutant being suppressed will have unique experiments that can be carried out.

First: cross suppressors to multiple mutant alleles of the gene of interest. This a good test for overall suppression; however, mutations that do not suppress other alleles may still be interesting (i.e., functional) due to their allele specificity. This must be measured against the other tests to determine the validity of the suppressors.

Second: use RNAi to disrupt any residual message from the mutant of interest in the mutant:suppressor background. If the suppressors are functional and not informational, you should *not* see a reversion back to the 100% mutant phenotype. However, the same stipulation stated above for alternative alleles also holds true for RNAi. Moreover, certain informational suppressors may act by affecting mRNA degradation, and thus could compromise the RNAi pathway. This potentially can complicate interpretation of tests using RNAi.

Third: in the case of a mutation of interest containing a premature stop codon, cross the suppressor into a strain containing another mutant gene with the same premature stop codon (e.g., an opal stop in both cases). Suppression of this unrelated mutated gene suggests you have isolated an informational suppressor.

Mapping Strains

Another key consideration in suppressor mapping is the creation of mapping strains. Depending upon the suppressors isolated, many suppressors will be silent (i.e., they will have no observable phenotype on their own). Therefore, in order to map these silent suppressors the original suppressed mutation must be included in all your mapping strains. Another consideration is

mating difficulties: some mutations may affect male fertility in the homozygous state. If this is the case, then it is always best to mate males into the mapping strain rather than into the suppressed strain to obtain heterozygote fertile males. The reason for this is to reduce the risk of losing the suppressor. Figure 23 shows a typical mating between a specific suppressor and a mapping strain.

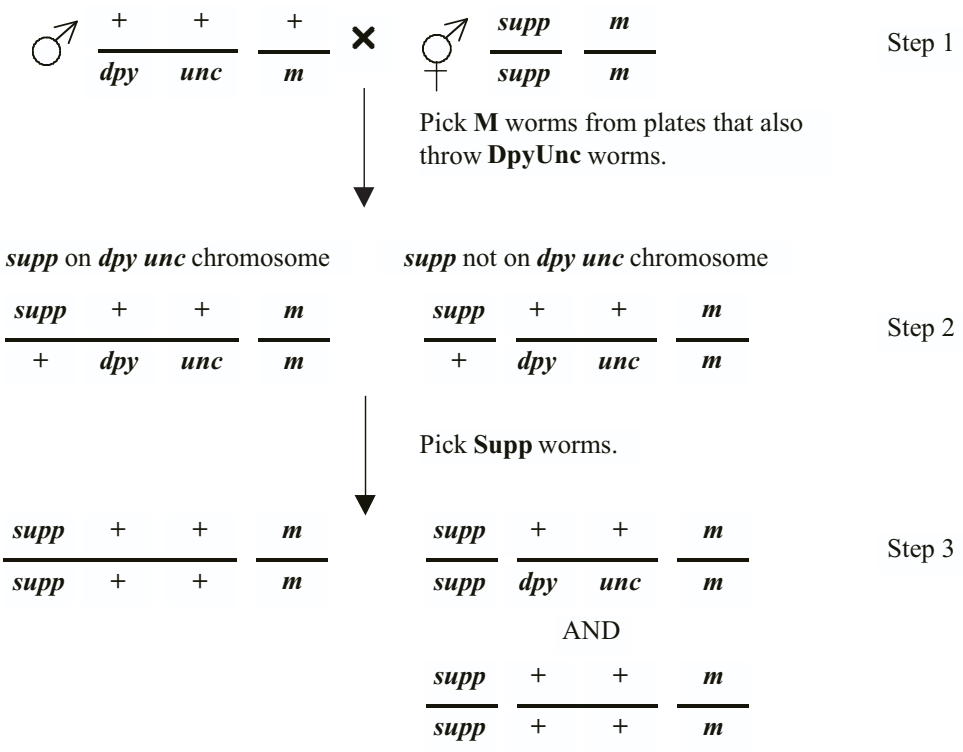


Figure 23.

The worms isolated in step 3 are then scored for the **Dpy** and **Unc** phenotypes in their progeny. If the suppressor is on the same chromosome as the markers, then 100% of the worms will *not* throw the markers (i.e., their progeny will not express these phenotypes). If the suppressor is not on the same chromosome, then two-thirds of the worms should throw the markers (see two-point mapping for a better explanation). These numbers, of course, are to be expected if the world were perfect, which it is not. Recombination makes this a little more difficult and a little more telling.

Recombination gives vital information for suppressors that are on the same chromosome as the markers. Due to recombination, a few of the progeny from step 3 may be either **Dpy** or **Unc**. The frequency with which these recombination events arise allows you to map the suppressor not only to a specific chromosome, but also provides information about the suppressor's direction and distance from the known markers. Three-point mapping is then carried out with other markers that should bookend the mutant. For a further and more thorough explanation, please see the two- and three-point mapping sections.

VIII. Mapping synthetic mutations

With the inevitable saturation of the genome for mutations that cause obvious plate phenotypes, the field will increasingly rely on the identification of mutations that act synthetically. Such mutations may have little or no functional consequence on their own, but when combined produce a strong phenotype. Mapping a synthetic mutation requires building marker strains that contain one of the mutations (presumably the previously cloned one) in the background. If the synthetic phenotype is a viable one, then mapping a synthetic mutation is relatively straightforward. If, however, the synthetic phenotype is lethal or sterile then a more involved approach will be necessary. It is this latter class that are addressed in this section. The key to the methods described below is that while labor intensive, unambiguous results can be obtained that will steadily move the mapping forward.

Making the mapping strains. Here is the problem. You have to put a synthetic mutation in the background of a marker strain. But it has no phenotype! How do you even know its there? How can you follow its presence or absence? The way around this is to make use of the opposite chromosome as shown in Figure 24.

A visible marker is chosen that maps close to the synthetic mutation, ***synA***, in this case an ***unc***. Following mating to N2 males (step 1), the ***unc***-het (heterozygous) male is next mated to the desired markers (step 2) to generate double-het males, which are then mated into the ***synA*** homozygous strain (step 3). We now identify cross-progeny animals that throw both **Unc** and **A B** progeny (step 4). By identifying an animal that has lost the ***unc*** mutation in the next generation (step 5), we have effectively selected for the ***synA*** homozygous strain. In step 6 the ***a b*** mutations are also homozygosed. In the construction of a ***dpy unc*** mapping strain, ***dpy*** or ***unc*** counter markers may be used, though it is preferable that they have phenotypes distinguishable from those of the ***dpy*** or ***unc*** markers.

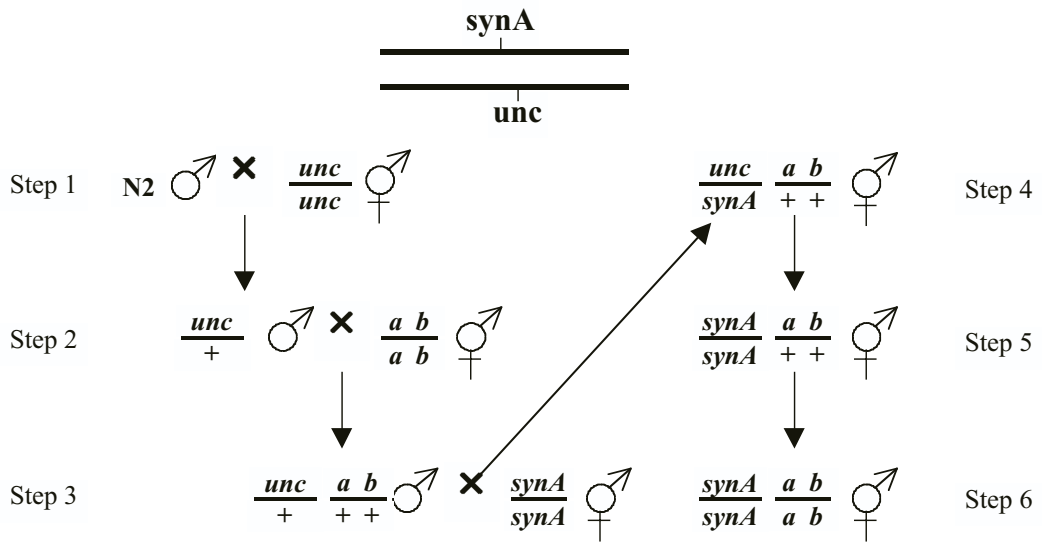


Figure 24.

While this method requires a fair amount of picking to guarantee selection of the ***unc*/*synA*; *ab*/+** animal, it is mostly fool proof assuming the counter-marker (in this case ***unc***) is close to the synthetic mutation (***synA***). Nevertheless, it is wise to generate at least two independent mapping strains to insure that the correct strain (***synA*** homozygous) is obtained. If known synthetic interactors of ***synA*** already exist, these mutations (or *RNAi*) can be used to test for the presence of ***synA*** in the mapping strain.

Two-point mapping. This method is for mapping synthetic mutations isolated using an extra-chromosomal array containing both a **GFP** marker and **syn A** rescuing sequences. The method is outlined in Figure 25.

In step 1, homozygous ***synA*** males are crossed into the mapping strain to generate trans-het males, which are mated to the double-mutant ***syn*** strain containing the extra-chromosomal array (step 2). Depending on whether or not the ***synX*** mutation is on the same chromosome as ***a b***, we have two scenarios. In #1, they are on different chromosomes. Therefore when we identify progeny where ***synX*** is once again homozygous (step 4), 67% of these will throw **A B** progeny. If ***syn X*** is on the same chromosome and close to the markers, rehomozygosed ***synX*** animals will fail to throw appreciable **A B** progeny. All the basic rules of two-point mapping apply here. In this case the frequency of recombinants will be about twice that of the actual map distance (see 2-point mapping for further details). Generally speaking one will want to pick about 100 animals for each chromosome since only 1/4 will be rehomozygosed to score. **Note that we will pick only Ex+ animals, even though animals that have lost the array are now viable.**

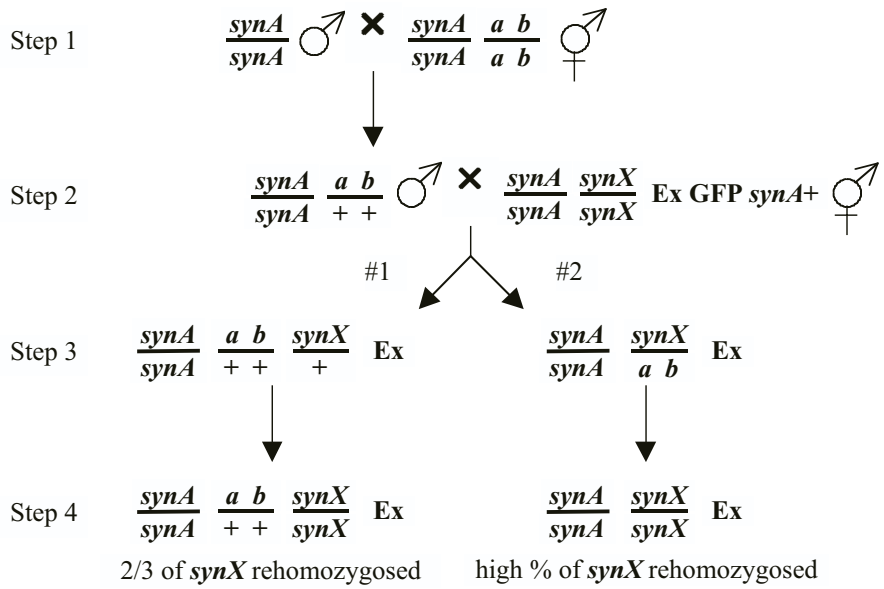


Figure 25.

A second source of information comes from the ***synX*** **non-rehomozygosed** plates. When the ***synX*** lies on same chromosome and close to the markers ***a b***, ***synX*** is essentially balanced. Thus nearly all non-rehomozygosed animals for ***synX*** will throw **A B** animals. (In this case the percentage of recombinants will directly equal the map distance.) In contrast, if ***synX*** and ***a b*** are on separate chromosomes, only 2/3 will throw **A B** progeny.

Three-point mapping. Once assigned to a chromosome, three-point mapping can be undertaken. The approach is reminiscent of the method used for two-point mapping in that we will seek to rehomozygose the ***synX*** mutation after picking recombinants. **An important point in mapping a mutation with no phenotype on its own is that we have to be absolutely sure that the *synX* mutation is actually present in the generation of animals from which we will pick the recombinants.** There are essentially two ways to ensure this, as shown in Figure 26.

One is to only pick recombinants immediately after obtaining the trans-heterozygous strain (scheme #1). This guarantees the presence of both ***synX*** and the ***a b*** marker. This is an effective way to initially do three-point mapping, provided the markers ***a*** and ***b*** are reasonably far apart (i.e., several map units or more). For markers that are closely spaced, it is often necessary to pick recombinants over several generations or more in order to get sufficient numbers (scheme #2). Here however we run the risk that a recombination event will lead to the loss of the ***synX*** mutation, and that this may go undetected. This can happen even though the markers have effectively balanced the silent ***synX*** mutation. To ensure that recombinants are picked only from plates where the parent is a true trans-het, we must make certain that ~25% of sibling plates are re-homozygosed. For example, we pick 40 animals (step 4)and 10 turn out to be homozygous for ***synX*** (step 5). We then have confidence to pick recombinants off the 30 non-re-homozygosed sibling plates. This strategy can be carried out

indefinitely (step 6) until sufficient recombinants are obtained. **Note that in picking recombinants as well as propagating the trans-het strain, we will only pick Ex+ animals, even though the array is not essential for viability.**

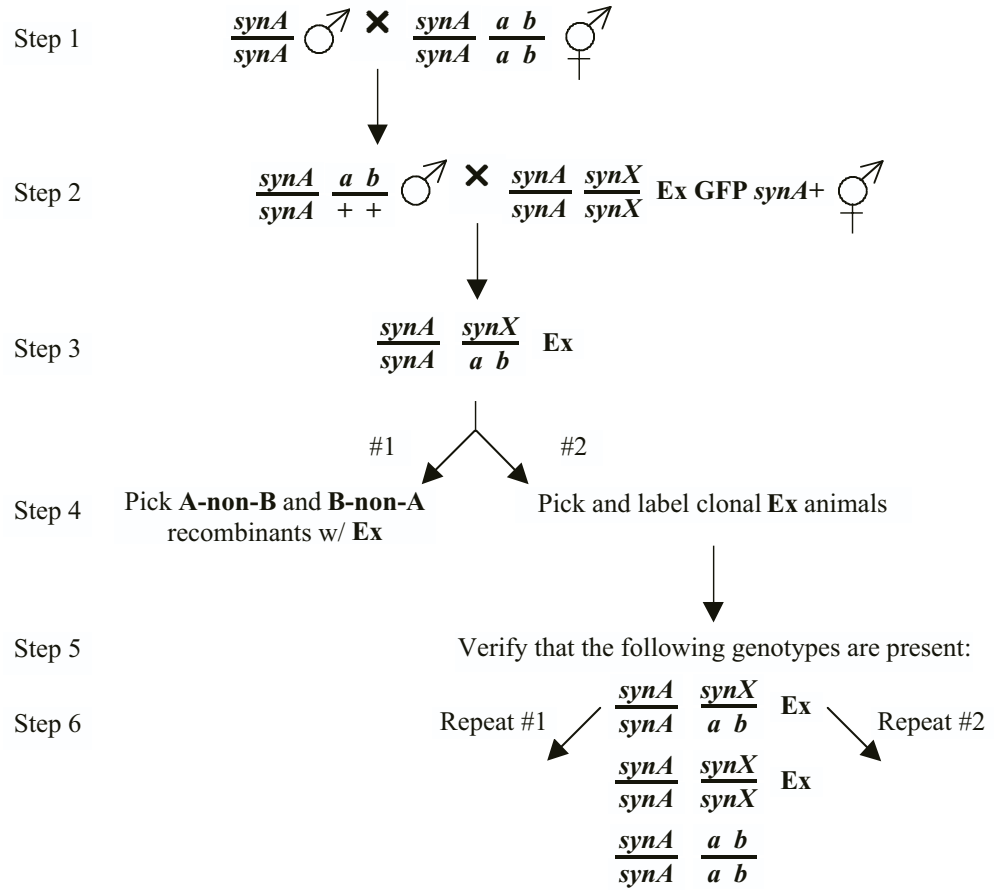


Figure 26.

Scoring recombinants. Once a recombinant has been picked we need to determine whether or not the recombinant chromosome has acquired the ***synX*** mutation, as shown in Figure 27.

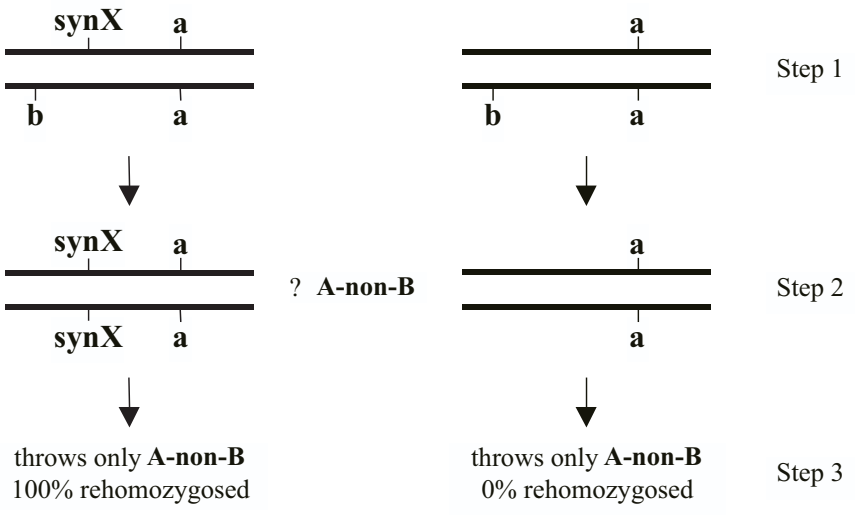


Figure 27.

This turns out to be quite straightforward. In the example on the left, an **A-non-B** recombinant has picked up the ***synX*** mutation (step 1). This animal will throw additional **A-non-B** progeny, 1/3 of which will be homozygous for the recombinant ***synX a*** chromosome (step 2). These animals will require the presence of the rescuing array and can therefore be scored positively for the presence of ***synX***. On the right, ***synX*** is not present on the recombinant chromosome and progeny **will never throw re-homozygosed *synX*** (step 3). In a typical situation we might pick 15 (Ex +) progeny from an F1 recombinant animal. If the ***synX*** mutation has been acquired, 5 animals on average will be homozygous for the recombinant chromosome. For those that are homozygous (based on the absence of ***a b*** progeny) we then ask: are these worms also homozygous for ***synX***? Though fairly laborious, this approach will generally give unambiguous data points.

It should be noted that if the mutation lies outside the markers and at some distance, the chance for a second recombination occurring where the *syn X* mutation is lost in some percentage of the progeny becomes substantial. Thus we might have a situation where five animals are clearly homozygous for the **A-non-B** chromosome, but only four are rehomozygosed with respect to ***synX*.** In this case we would count the recombinant as positive for acquiring ***synX***, and might conclude that ***synX*** is unlikely to lie between the two markers. This added complexity is not a factor when ***synX*** is not initially acquired by the recombinant since such progeny will not contain the ***synX*** mutation regardless.

Deficiency mapping. The good news is that we need not construct **Df** strains that are homozygous for the ***syn A*** mutation. The scheme for **Df** mapping is outlined in Figure 28.

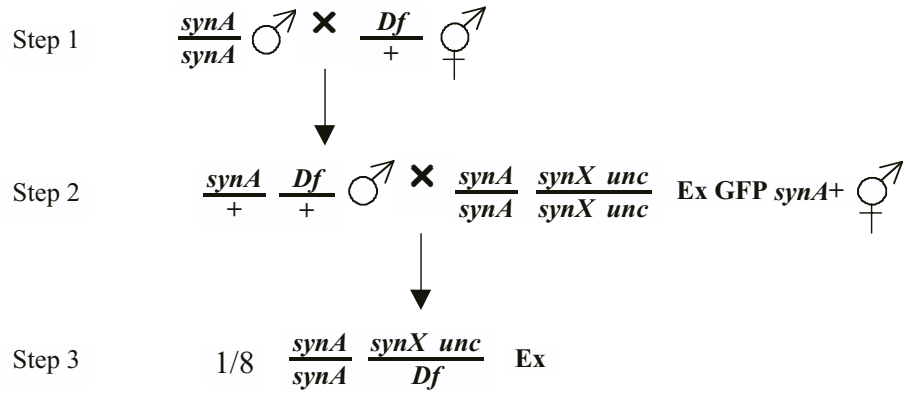


Figure 28.

Here the **Df** is crossed into the double-mutant strain containing the array (step 2). The key is to unambiguously determine cross-progeny at this stage. In the example shown above, the chromosome with ***synX*** contains in a visible marker (***unc***) in cis. Alternatively, the double-mutant strain can harbor visible markers on a separate chromosome (such strains are readily obtained during two-point mapping). In the scheme above, 1/8 of the cross-progeny will be both homozygous for ***synA*** and trans-heterozygous for **s*ynX*** and the ***Df*** **(step 3)**. Assuming that ***synX*** results from a relatively strong LOF mutation, such a strain would be predicted to show the genetics of a rehomozygosed one, i.e., only Ex+ animals will be viable. As discussed above (see ***Df*** mapping section), a positive result will be more meaningful than a negative result.

SNP mapping a synthetic mutation. Thankfully, SNP techniques can be used for mapping synthetic mutations. The first key steps is to construct a strain that harbors your mutation of interest in the CB4856 (Hawaiian) background. This can be accomplished by initially crossing your mutation of interest (***synA***) into the CB4856 strain, then successively backcrossing this strain to CB4856. You will need to positively identify backcrosses that contain your mutation of interest either by RNAi injection or sequencing. RNAi feeding will not work, because CB4856 contains several loci that make the worms much less sensitive to feeding methods.

Second, you will needs to create a strain where the mutation to be mapped, ***synX*** is flanked by two visible markers (e.g., ***dpy synX unc***). The basic scheme is shown in Figure 29**. However, under no circumstances should you attempt to score recombinants from CB4856 crosses by using the RNAi feeding of the synA. <u>The scheme shown in Figure 29 is currently faulty in this respect and will be fixed within the coming weeks!</u>**

The infinitely preferred current method is to generate the doubly marked strain (e.g., ***dpy synX unc***) in the ***synA*** background with the rescuing array. Then by crossing into the backcrossed ***synA***/CB4856 strain that was generated, carry out in the manner described for the 3-point mapping of synthetic mutants. The

presence of the mutation and rescuing array in the SNP mapping process, obviates the need for an RNAi step, which would lead to misleading results.

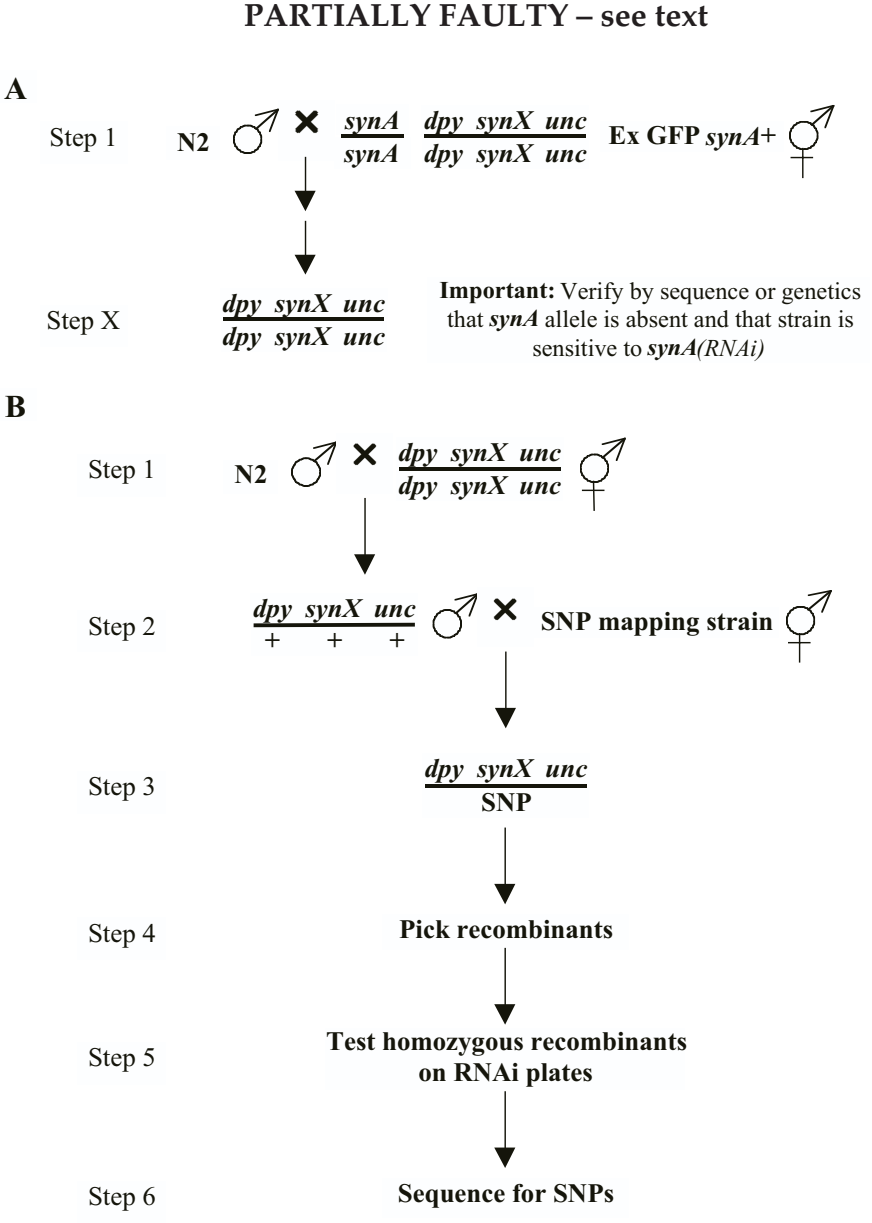


Figure 29.