

The Zon Laboratory Guide to Positional Cloning in Zebrafish

Division of Hematology/Oncology, Children's Hospital Boston, Boston, MA 02115
Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, 02115

I. Introduction

11. Mapping Strains

111. Families and Genetic Markers

IV. Crosses for Line Maintenance and Mapping

A. Choosing Grandparents and Parents for Better High-Resolution Mapping

B. Microsatellite Marker---Agarose Scorable

V. Preparation of the DNA

VI. Mapping Genes

A. Low-Resolution Mapping

B. Intermediate-Resolution Mapping

C. Fish Husbandry

D. Tempo

E. High-Resolution Mapping

F. Three-allele Systems Versus Four-Allele Systems

G. Collecting Mutant Embryos

H. Chromosomal Walking

I. Screening the PAC or BAC Libraries

J. BAC Libraries

K. Care of the BAC and PAC Libraries

VII. Overgo Strategy for Rapidly Doing Chromosomal **Walks** and Positional Cloning

VIII. Protocol for Overgo Probing of High-Density Filter

A. Overgo Labeling

B. Preparation of OLB (-A, -C, -N6) Solutions

C. Hybridization Procedure

D. Washing

E. Autoradiography

IX. General Flow of Information from the Radiation Hybrid Panel Maps, the Sanger Institute Sequencing Project, and Fingerprinting the BACs

A. PCR Screening of the Pooled PAC and BAC Libraries

B. Walking and Establishing Contigs by Sequencing PACs

X. Synteny Between Human, Zebrafish, *Fugu* and *Tetraodon* Genomes

XI. Proving a Candidate Gene Is Responsible for the Mutant Phenotype

XII. Morpholinos

References

I. Introduction

The process of positional cloning involves unique issues for all organisms. Success is usually based on experience. Because the Leonard Zon laboratory at the Children's Hospital of Boston has used positional cloning to isolate more than 10 zebrafish genes, the laboratory team has accumulated significant experience in the field. We have summarized this experience to develop this guide that will be helpful for the community of researchers who study and work with zebrafish.

II. Mapping Strains

Many of the problematic issues concerning positional cloning in zebrafish (*Danio rerio*) arise from the

genetic polymorphisms in the individual strains of zebrafish. In the Zon Laboratory, we have typically used the AB or TU (Tuebingen) strains for mutagenesis screens. A mutant should be maintained on the laboratory strain that was used to create mutants. These and all other widely available zebrafish strains are not entirely inbred. Genetic polymorphisms may be present in a given family of fish. One cannot assume that a family of zebrafish is as isogenic as inbred strains of mice. Therefore, it is important to track polymorphisms in a mutant family by obtaining grandparent and parent DNA samples, usually from tail clips. This tracking will enable the examination of all possible polymorphisms in progeny that subsequently become important in selection of polymorphic markers to genetically locate a mutant gene.

WIK and SJD, two zebrafish strains, are polymorphic with respect to the AB and TU strains and can therefore be used for genetic mapping of mutants on AB and TU backgrounds. Strains used in mutagenesis and mapping can be interchanged; that is, mutations can be created on WIK or SJD, and mapping can be done using AB or TU strains. For our mutations on the AB background, a heterozygote carrier on the AB strain is mated to WIK, and a mapping family is generated.

III. Families and Genetic Markers

Once map crosses have been created and heterozygote F1 mapping pairs have been identified from these mapping families, it is important to tail clip and store the DNA from these mapping pairs and from the parents of these fish [e.g., the AB mutant heterozygote and WIK wild type that created the map cross (the grandparents) as well as the ABIWIK heterozygotes (the parents)]. Parent DNA and grandparent DNA are helpful for analyzing usable polymorphisms in subsequent mapping. Once flanking markers have been found in the initial mapping stage, it is important to test the parents and the grandparents to determine which mapping families and/or parent pairs have usable polymorphism and to segregate the flanking markers in an easily interpretable manner. In other words, it is best to collect embryos with suitable allele systems for high-resolution mapping. Once polymorphism has been determined, the mapping heterozygotes that are polymorphic can be selectively used to create mapping panels that will also be polymorphic for the flanking markers. Very frequently, if a mutant does not map easily because of lacking good polymorphic markers, examining another family from a different mapping cross is very helpful. It is recommended to create multiple mapping families using different grandparent pairs at early stage of cloning to save time.

IV. Crosses for Line Maintenance and Mapping

For the purposes of this discussion, we should note that mutagenesis is performed with AB, and WIK is the polymorphic strain used for mapping. In addition, the reader should assume that the mutation of interest is homozygous embryonic lethal, and lines must be maintained as heterozygotes.

Definitions

Incross: Sibling cross

Outcross: AB (mutant)/AB heterozygote x AB wild type

Mapcross: AB (mutant)/AB heterozygote x WIK wild type

Backcross: AB (mutant)/WIK heterozygote x WIK wild type

For long-term line maintenance, we keep the mutation on the same strain in which mutagenesis was originally performed (in this case AB) so that we do not jeopardize future mapping efforts. One could perform an AB (mutant)/AB heterozygote incross or an outcross. Outcrosses are generally preferable because they help dilute out other recessive mutations acquired during the ENU mutagenesis that are not linked to the phenotype of interest. Using mapcrosses or backcrosses for line maintenance is problematic because recombination can result in a polymorphism loss of distant linked markers that could be critical for future low-resolution mapping.

For mapping, polymorphic hybrid strains must be created through a mapcross. The offspring are raised, half of which will be heterozygotes. Heterozygotes are identified by multiple random incrosses, and identified heterozygote pairs are set aside in isolation tanks for ongoing incross embryo collection. If we

cannot collect enough embryos in that generation and the close flanking markers are not available, we turn to the newest generation of the heterozygotes on original strain background [AB (mutant)/AB] and perform another mapcross to repeat the process. We do not raise embryos from an AB (mutant)/WIK incross because recombination in this generation renders the next generation of embryos useless for mapping. If, however, close markers flanking the mutation are available, a backcross can be performed **to generate a large number of parents with identical polymorphisms for fast collection of a large mapping panel**. If the markers are agarose scorable, the potential exists to identify large numbers of parent heterozygotes with a good allele system by tail clipping instead of mating.

Using a high-throughput PCR format, we have identified up to 50 heterozygote pairs in a week. It is critical to tail clip the parents used in the backcross as well as tail clip the grandparents used for the mapcross so the allele system can be accurately followed in the next generation. To facilitate this process, it is useful for a laboratory to have isolated several (about 10) tail-clipped wild-type fish (in this example, WIK is used). Before the backcross is performed, the tail-clipped WIK wild types are screened with close flanking markers along with tail-clip DNA from the AB (mutant)/WIK heterozygote to establish which wild type has the best allele system for following the mutation. Those WIK wild types are then used for the backcross.

When identifying heterozygotes by tail-clip DNA, it is important to remember that there will be a defined error rate due to recombination. The magnitude of the error depends on the distance between the marker and the mutation. The recombination rate in males is about tenfold lower than in females; therefore, if flanking markers are still somewhat far from the mutation, one should consider doing a backcross in which the AB(mutant)/WIK is male and the WIK wild type is female.

A. Choosing Grandparents and Parents for Better High-Resolution Mapping

Usually after low-resolution mapping, our laboratory segregates individual WIK fish and then genotypes them. The goal is to define the allele in the system that represents the best advantage for our high-resolution mapping purposes.

B. Microsatellite Marker---Agarose Scorable

We have recently undertaken a large-scale approach to evaluate microsatellites (available from the Fishman laboratory) for their ability to be scorable on an agarose gel (Massachusetts General Hospital, 2001). (<http://zebrafish.mgh.harvard.edu/>).

Through this analysis, we realized that SJD fishes were mostly isogenic; however, the strain has some regions with polymorphisms. SJD allows for easier mapping than WIK, but the strain is very difficult to use and cannot be propagated in our laboratory. SJD males can be obtained and used to fertilize eggs *in vitro* from heterozygote females for creating mapping families. The WIK strain works well, but it is very polymorphic between individual fish (not inbred). We use WIK in most of our mapping. We have put together a new mapping panel (288 CAs) for genome-wide scan. A list of these markers can be downloaded here ([288 Low Res Scan CAs May 2006.htm](#))

V. Preparation of the DNA

If the embryos have not already hatched, our laboratory dechorionates them before freezing and DNA preparation. We place the embryos individually into wells of a 96-well plate. We then remove excess buffer, store them dry or in methanol and make sure they are kept at -20 °C.

When working with embryos frozen or methanol preserved, plates should be kept on ice unless otherwise noted. To prepare the embryos, remove all methanol from the wells. All of the following incubation steps can be carried out in a PCR machine:

1. Add 50 μ l of lysis buffer (composition follows) to each well and incubate at 98 °C for 10 min to lyse cells. Quench on ice or 4°C in the PCR machine.
2. Add 5 μ l of Proteinase K (10 mg/mL stock) to degrade proteins.
3. Incubate at 55 °C for at least 2 h. We recommend at least one mixing during the PK incubation to increase consistency of the DNA preparations. This incubation can also be left to run overnight. The longer the incubation time, the cleaner the DNA tends to be.
4. Incubate at 98°C for 10 min after 55°C incubation to destroy Proteinase K. Quench on ice (or 4°C sink in PCR machine or 10°C sink in PCR machine that is better for PCR machines especially in an overnight procedure).
5. Spin down lysed embryo debris at 4000 rpm for 10 min.
6. Draw off supernatant into a clean 96-well plate.
7. Dilute as necessary.

Embryo Lysis Buffer

Solution: 1 x PCR buffer made to 0.3% Tween 20 (10% stock) and 0.3% NP40 (10% stock)

For 10 ml of buffer, use the following:

10 ml PCR buffer (see following composition)

300 μ l NP40, 10% stock

300 μ l Tween 20, 10% stock

PCR Buffer

Solution: 10 mM Tris-HCl, pH 8.3, and 50 mM KCl

For 50 ml of buffer, use the following:

500 μ l 1 M Tris, pH 8.3 (autoclaved)

2.5 ml 1 M KCl

47 ml sterile ddH₂O

VI. Mapping Genes

A. Low-Resolution Mapping---Bulk Segregant Linkage Analysis

Two preferred methods exist for low-resolution mapping of a gene to a particular chromosome. The first method makes use of scanning microsatellite markers throughout the genome (Knapik *et al.*, 1998; Shimoda *et al.*, 1999). The second method, called half-tetrad analysis (Kane *et al.*, 1999), makes use of early pressure-treated embryos to evaluate the mutated chromosome.

1. Genome wide scanning---Low-Resolution Scanning

The first step in mapping a recessive mutation to a chromosome is the generation of mapping hybrids (AB(mutant)/WIK). To make this mapping cross, a heterozygous AB carrier mutant /AB) is mated to a wild-type WIK fish, and the resulting F1 generation is raised. Practically, we generate two to four mapcross families with different WIK founders to ensure an informative allele system is obtained (because the WIK strain is not completely inbred). To identify heterozygous F1 individuals, these AB/WIK hybrids are mated to each other (incrossed), and their clutches scored for the mutant phenotype and the mutant/wildtype ratio. Once a pair of heterozygous hybrids is found, they are mated, and their wild-type and mutant progeny are collected.

In addition, we tail clip the mapping heterozygotes to obtain DNA for analysis of the alleles carried in a family. For the initial low-resolution mapping, we use 40 mutants and 40 wild types. Two mutant pools of 20 and two wild-type pools of 20 are made from these stocks. To make the pools, we take 8 μ l of the individual stock DNA from each of the 20 individuals (mutant embryos for the mutant pool or wild-type embryos for the wild-type pool) to give 160 μ l, and we increase the volume to 1.6 ml with water and use 4

μl for each PCR amplification. Last, we use the agarose-scorable microsatellite markers to scan the genome for linkage to the phenotype in bulk segregant analysis.

Bulk segregant analysis uses 288 agarose-scorable microsatellites that are typed on a set of DNA samples from wild-type and mutant embryos. Each set contains two pools of twenty embryos from wild types and mutants. -PCR products are run on 3% agarose gels at 200 V for 2 h to separate bands. Thus far, most polymorphisms encountered have been subtle; hence, running the gels longer than necessary is always better. Linkage is assumed when a band present in the wild-type pools is absent in the mutant pools. The mutant band(s) may also have a size shift when compared with the wild types. This observation may also indicate linkage. It is best that individuals for this stage as well as for the intermediate mapping stage (see next section) come from the same family. Occasionally, our laboratory is not able to map a mutation in one family, and we have to test another family. It is critical that the 80 embryos come from the same family. Introducing individuals with a different set of alleles might lead to false positives. One might assume linkage to a particular chromosome based on the pattern obtained with the additional family. Very frequently, we find linkage to about three chromosomes, but only one of these linkages is real. To determine which microsatellite is truly linked, each positive marker must be tested on individuals that are used to create the pools. By testing individuals, chromosomal linkage is confirmed, and the distance between markers can be evaluated.

2. Half-Tetrad Analysis

Although the zebrafish is a diploid organism, haploids can live for several days. Maternally homozygous diploid fish can be produced by applying early pressure (EP) to inhibit the second meiotic division after fertilization with UV-inactivated sperm. Analogous to the creation of the maternal diploids, diploid androgenotes can be obtained by UV inactivation of the egg, subsequent fertilization by normal sperm, and application of EP to inhibit the second meiotic division. The ability to create gynogenetic diploids allows the rapid assignment of a gene to a particular chromosome while obtaining information about its distance from the centromere. To genetically localize a mutation, female offspring of an AB mutant × AB (wild type) outcross are squeezed and fertilized with sperm from a genetically unrelated male (also wild type). The F1 heterozygous females are identified by random matings between the F1 offspring. The F1 heterozygous females are subsequently squeezed, and gynogenetically diploid embryos are derived. The distance between the mutation and the centromere can be approximately calculated using the following equation:

$$\text{Distance [cM]} = 50[1 - (2 \times \text{Mutant Number} / \text{Total Number of Embryos})]$$

Because the second meiotic division was inhibited in creating gynogenetic diploids, the region of the chromosome between the centromere and the mutation could not recombine in mutant embryos. Hence, markers proximal to the mutant (so-called centromeric markers), when polymorphic between the background and wild-type strain, are homozygous for the background strain allele in mutants and have the wild-type allele in unaffected embryos. Because centromeric markers have been defined for all 25 zebrafish chromosomes, chromosomal localization and distance from the centromere can rapidly be assigned (Kane *et al.*, 1999).

In our laboratory's early work with zebrafish, we used half-tetrad analysis. As we developed robotics in the laboratory, it became easier for us to develop the scanning method described in the above paragraphs. Recently, with the advent of many polymorphic markers that are agarose-scorable, we favor the above scanning method.

B. Intermediate-Resolution Mapping

The purpose of intermediate-resolution mapping is to position the gene between close flanking markers that are scorable on an agarose gel. This technique allows us to do high-resolution mapping with 1500 embryos with relative ease. Mapping with this number of embryos is not always possible, but it is a goal worth striving for.

We collect 8 wild types and 88 mutants for intermediate-resolution mapping. To clone a mutant gene, flanking markers should ideally be less than 10 cM apart. We scan microsatellites on the chromosome by ordering roughly six microsatellites on the linked chromosome arm. If these microsatellites are not polymorphic, we test another six markers until the mutation is linked to two flanking polymorphic microsatellites on the chromosome arm. Based on this recombination mapping strategy, it should be possible to define the flanking microsatellite markers. Markers that are far away from the mutation should yield more recombinants than markers that are close to the mutation. Markers that are on opposite sides of the mutation should give different sets of recombinants, and markers that are on the same side of the mutation should share recombinants. We can narrow down the region by studying more microsatellites until we have markers that are 10 cM (or less) apart. When we are able to define microsatellite markers that are polymorphic on our agarose gel and that are close enough to use as flanking markers, we set up mapping crosses with zebrafish that have this allele system. These new stocks of fish and validated flanking markers are then used for the high-resolution mapping.

C. Fish Husbandry

The number of tanks needed to map a mutant varies based on zebrafish sex ratio in the tanks and on the ease of scoring the phenotype. We will typically generate eight mapcross tanks or backcross tanks for a genetic mapping. Ultimately, we will sacrifice most of these fish, but the goal is to have at least five pairs of fish with an advantageous allele system so that genetic mapping can be conducted very quickly.

D. Tempo

While working out the flanking polymorphisms and genotyping, it is important to continue collecting mutant embryos. We find that between 1500 and 2000 embryos are required for positional cloning. Assuming an interval of 600 kb/cM and a meiotic recombination frequency of one per embryo (this is true for haploid individuals; diploids have an average of 1.3 to 1.5 meioses/individual---one from the mother and 0.3 to 0.5 from the father) will give a resolution of close to 30 kb per meiosis event. This resolution allows the positioning of the mutant gene on a BAC or PAC clone. The following discussions are based on the RH map. These are still applicable to cases when the genome assembly or other genome maps are used. The goal is to identify closer polymorphic markers between the existing flanking markers on the mutant mapping panel. Ultimately, the mutant locus can be located between two closer flanking markers that are less than 1 cM away from the mutant locus. Ideally, the mutant locus is located on a genomic clone (BAC or PAC) or within a small region on the genome assembly. These discussions are even more applicable when the mutant region is in a gap or mis-assembled region on the genome sequence assembly. Please read section IX in this guide for detail information.

The characterization of markers between the flanking polymorphic markers is important. The number of recombinants obtained with each of the flanking markers [which should be placed on the radiation hybrid (RH) map if this information is not available], from intermediate-resolution mapping is considered a guide for estimating the position of the mutated gene. The RH map contains a multitude of expressed sequence tags (ESTs) and BAC ends, which can be used as markers on the walk toward the mutated gene. Typically, our laboratory will pick three to four ESTs in proximity of the estimated gene location and first check if the 3'-UTRs of these ESTs are polymorphic in our mapcross. Frequently, the primers used to map the ESTs are on the Washington University in St. Louis Web site created for EST projects (Washington University in St. Louis, 2004a) or our RH map website at <http://zfrhmaps.tch.harvard.edu/ZonRHmapper/>. If this is not the case, single-stranded conformation polymorphism (SSCP) primers are obtained from a noncoding DNA sequence in the region. In our experience, one out of four ESTs is polymorphic in a given mapcross. If a polymorphism with the chosen ESTs is not found, more ESTs can be tested for polymorphism, or EST primers can be used to isolate BACs or PACs. Sequencing the ends of BACs or PACs will increase the likelihood of identifying polymorphic markers. Recently, we have found that BLAST analysis of the Sanger Institute genome sequence (The Sanger Institute, 2004a) with an EST will give more extensive sequence with introns and UTRs. These sequences can be searched for polymorphisms. Introns have a higher rate of polymorphism than 3'-UTRs. In addition, CA repeats or contigs can be used to search for

polymorphic markers. Once polymorphic markers are identified, the panel of recombinants identified at each of the flanking markers is tested.

Example

Problem: A total of 30 recombinants were identified with the left flanking marker and 35 with the right flanking marker. If 1500 mutant embryos have been collected as part of a the mapping panel, the distance to the mutant gene is estimated to be 2 cM from the left flanking marker and 2.3 cM from the right flanking marker (both are too far to initiate a walk). Two ESTs are identified in the estimated region of the mutated gene. Both are polymorphic. Test the panels of recombinants from either side with these markers as shown in Fig. 1.

Solution: The mutated gene is situated between EST 1 and 2. The estimated distance from the closer EST 1 is 0.26 cM. This distance is sufficiently small to initiate a chromosome walk using PAC or BAC clones: To solve this problem, one can also use the Sanger Institute genome sequence. A comparison is made of contigs that contain ESTs from the RH panels. If concordance is found, this improves the confidence of the genetic interval.

	Flanking Marker	EST1		EST2	Flanking market	
Left						Right
Recombinants	30	4		0	0	
	0	0		16	35	Recombinants
		<i>GENE</i>				

Figure. 1. Mapping of mutant genes.

E. High-Resolution Mapping

We have traditionally collected between 1500 and 2000 mutant embryos in an effort to clone these positionally. In high-resolution mapping, these mutant embryos, arrayed in a 96-well format, are tested with the flanking markers found in the low and intermediate stages of mapping. It is critical that every recombination event is scored in this step, assuming a two-allele system. If a three- or four-allele system is used for mapping, some recombination events will he missed. Therefore, it is recommended that the families used for collecting the embryos be chosen according to the most useful allele system as well as the correct polymorphic mapping strain. Furthermore, it is advantageous to limit the number of families used to collect the mutant embryos. Although collecting embryos from only two or three pairs of fish may lengthen the time needed to reach the target number of embryos, this step will simplify allele systems of markers in downstream steps in positional cloning. We recently moved our high-resolution mapping to the ABI 3730 for quick scanning through entire mapping panel using the closest flanking markers.

Allele-Specific Oligonucleotide Hybridization

Alternatively, allele-specific oligonucleotide hybridization can be used instead of SSCP to quickly screen through the mapping panel to assemble a recombinant mutant panel for positional cloning. Polymorphism between mapping strains for markers recovered from intermediate mapping, such as ends of EST, BAC, PAC, and YAC clones, can be converted into an allele-specific oligonucleotide (ASO) hybridization assay for a meiotic map. The ASO assay (Farr *et al.*, 1988; Wood *et al.*, 1985) is capable of detecting single nucleotide polymorphic differences between mapping strains. The basis for the assay is PCR amplification from genomic DNA for the corresponding BAC, PAC, or YAC end, dotting the PCR product on nylon membrane, differential hybridization with 5'-labeled 19-mers containing the single nucleotide difference present in the mapping strains, and autoradiography of the washed nylon filter. All the hybridizations of the

5'-labeled ASO primers are performed in tetramethylammonium (TMA) buffer (Farr *et al.*, 1988) at 45°C, and they are washed at 55" to 56°C in TMA for any oligo length of 19-mers, irrespective of GC-content.

To use this assay, you need to know the sequences for the corresponding BAC, PAC, or YAC end from your two mapping strains to design ASO primers. You can use the PCR primers to amplify the fragment from your homozygous mutants and homozygous wild-type embryos (these have been ascertained by other microsatellite markers). The PCR fragments are then directly sequenced or subcloned to determine their single nucleotide differences, which are used to design your ASO 19-mers. The PCR fragments are immobilized onto nylon membrane in duplicate. For one set of membrane, you would hybridize with an ASO primer for one strain, and then you would hybridize the duplicate membrane with the second ASO in a separate container. After hybridization at 42°C for minimum of 4 h, wash the filters in 2 x SSC 0.1% SDS at room temperature for 20 min, and then wash them in a TMA wash buffer at 55" to 56 "C for specificity. Then you can expose the filters for autoradiography.

Once the ASO has been shown to work in the mapping cross, genotyping of large numbers of embryos can happen in a high-throughput manner. Because the dot blot manifold can accommodate 96 samples, you can array 96 samples of PCR products from individual embryos/filters. Thus, using this assay, you could genotype approximately 1000 embryos with 10 to 11 filters in only one experiment.

Note the supplies needed for the ASO assay:

*ASO: You will need to design an oligonucleotide of 19 nucleotides in length. This oligonucleotide has a single nucleotide difference from your mapping strains located at the center of the 19-mers to maximize their T_m differences.

*Dot/slot blot manifold (Schleicher and Schnell, YEAR; Biorad, YEAR): You will need this manifold that facilitates application of PCR products onto a nylon membrane.

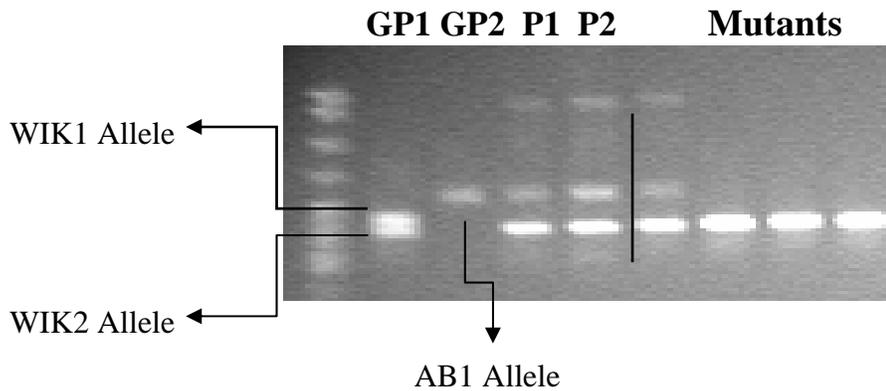
*Tetramethylammonium (TMA) chloride solution: This solution is the key component to the ASO assay because it equalizes the T_m differences based on GC content of the 19-mers. As a result, the T_m stability is a function of length that exactly matches to the target DNA. A premade 5-M stock solution is available from Sigma (Cat. #T3411).

*[γ - ³²P] ATP of 3000-6000 Ci/mmol specific activity:

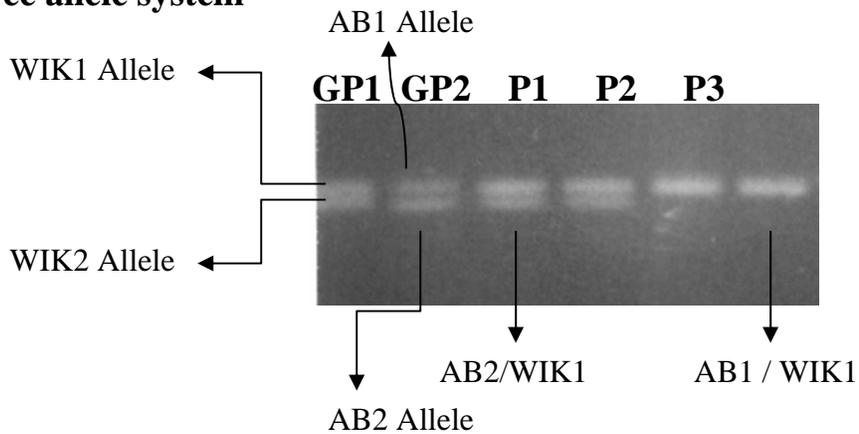
F. Three-Allele Systems versus Four-Allele Systems

Typing mapcross parents is essential to evade the "allele traps" encountered by this type of bulk segregant analysis. Once flanking markers have been found, mapcross parents should be tested for the proper allele segregation. Three-allele systems are quite common in the AB/WIK crosses used in our laboratory. We often see that the flanking markers do not segregate similarly between different families. This becomes a problem with the high-resolution scan that includes individuals from all mapcrosses. One of the wild-type alleles sometimes has the tendency to migrate the same as the mutant alleles in the agarose gel, so when a high-resolution scan is performed, recombinants are missed because they look like a mutant embryo. A simple way to avoid these had-allele systems is to type all the parents and grandparents before collecting all 1500 embryos. While the low and intermediate stages of mapping are being performed, embryos can be collected. One should remove those crosses that have bad allele migration from the mapping collection. Four-allele systems can be just as confusing if not fully investigated before the high-resolution phase. Generally, four alleles can be tracked with ease. Problems, however, can arise when an AB allele segregates with the mutant WIK allele. Heterozygotes would be counted as homozygous mutants because similar allele size of both AB and WIK and not recombinants. The same situation is seen with three-allele systems (Fig. 2).

Good-three allele system



Bad-three allele system



Good-four allele system

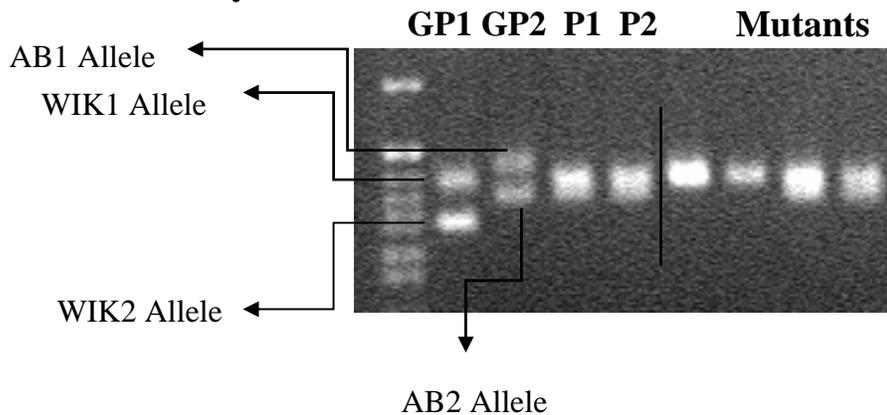


Figure 2. Allele systems in zebrafish

G. Collecting Mutant Embryos

With the conclusion of high-resolution mapping, a number of recombinants have been identified. These recombinants are rearranged on two new plates to create the recombinant panel. During this process, it is advisable to continue collecting mutant embryos from the mapping strain beyond the initial collection number of 1500 to 2000 because more embryos may be required later in the process. The recombinant

panel is now used in positional cloning, with the number of recombination events lessening as the mutation is neared.

H. Chromosomal Walking

After having the recombinant panel made from 1500 to 2000 collected mutant embryo mapping panel and the closest flanking marker (s) within 1 cM distance, a chromosomal walking can be tried (Fig. 3). The flanking markers that are polymorphic by SSLP or SSCP and linked to your gene are used to find the next an internal marker between two flanking polymorphic markers. Based on the flanking marker sequence, an overgo probe (a double-stranded 40-mers) can be designed. Typically, marker sequences are derived from 3'-UTRs of an ESTs or ends of a P1-derived artificial chromosome (PAC) or a bacterial artificial chromosome (BAC) clones screened with one of the close linked markers. Ends of isolated genomic clones are sequenced if their sequences are not available. Polymorphisms of the end regions were tested on the mapping panel and new polymorphism markers are identified. Using these internal polymorphic markers, orientation of the walk relative to the gene is established by studying meiotic recombinants in F2 embryos. By taking the marker that is closest to the gene, a large-insert genomic libraries (PAC or BAC) are screened again and the new genomic DNA clones are isolated and sequenced from both ends. Polymorphism of these ends is tested on the mapping panel again and the orientation of the walk is again established. These end sequences are used to generate new overgo probes. These overgoes are used to screen the large-insert genomic libraries, and, therefore, a "walk" towards the mutant gene has been established. When starting a chromosomal walk, keep in mind that the narrower the genetic interval, the faster the positional cloning will go.

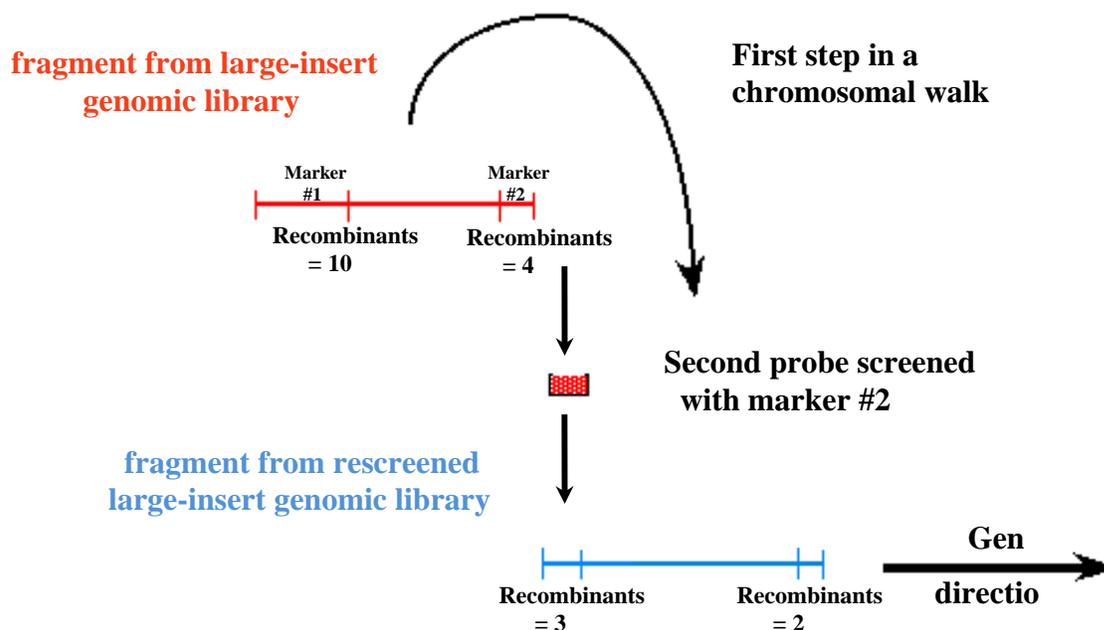


Figure 3. Chromosome walking.

I. Screening the PAC or BAC Libraries

There are various methods for screening the PAC and BAC libraries. We use a hybridization strategy, using a prediction program to make overgoes (Washington University at St. Louis, 2004b). The overgoes are individually hybridized onto filters. The availability of "double-positives" (positives with two independent overgoes designed against the same sequence fragment) is an advantage for isolating true positives. We typically isolate between 5 and 15 clones per hybridization. The first PAC library for zebrafish was originally isolated from AB zebrafish red cells (Arnemiya and Zon, 1999). The PAC library

has an insert size of roughly 100 kb that encompasses 5x coverage or 250 384-well plates. The first BAC library was made from a single AB fish and has clones that are only 82 kb in size on average. The BAC libraries constructed for the genome sequencing project were made from Tu fish and clones are ~150 kb in size on average. Both types of libraries are assembled in a similar format. These libraries are in-house and readily available to us for our everyday work. These BAC and PAC libraries are also available at the Deutsches Ressourcenzentrum für Genomforschung (RZPD), translated as the Resources Center of the Human Genome Project in Berlin, Germany.

J. BAC Libraries

BAC libraries were prepared by Chris Amemiya's, Peter De Jong's, and Ron Plasterk's laboratory. The later two libraries were created by isolation of DNA from blood pooled from 10 male and 10 female Tuebingen fish, for a total of 20 fish. The BAC vector backbone, pBACe3.6, has the advantage of being far smaller than the PAC vector. This is advantageous when shotgun sequencing a single large-insert genomic clone (more clones will represent your insert and not vector). The average insert size is approximately 150 kb, with a large number of 200 kb clones represented in the library. Researchers P. de Jong and R. Plasterk have also constructed two additional libraries, CHORI 211 and *Danio* Key Zebrafish BAC. These two libraries are the backbone of the Sanger Institute's *Danio rerio* Sequencing Project and are available at the Children's Hospital Oakland Research Institute (BACPAC Resources Center, 2004) (<http://bacpac.chori.org>) and the Deutsches Ressourcenzentrum für Genomforschung (RZPD, 2004).

K. Care of the BAC and PAC Libraries

Condensation and cross-contamination are sources of problems for maintenance of the glycerol stocks of the PAC and BAC libraries. To reduce cross contamination and maintain viability of the cultures, the 384-well plates are handled carefully to limit defrosting. Plates are removed from -80°C storage and are allowed to warm at room temperature for approximately 5 min. A sterile pipette tip is used to remove a tiny chip of the frozen culture, and subsequent streaking is performed on an agar plate or inoculation of a broth culture. The PAC clones are kanamycin resistant (25 µg/ml), and the BAC clones are chloramphenicol resistant (34 µg/ml). The inner sides of the lids are wiped with ethanol if necessary, and the plates are replaced in storage.

VII. Overgo Strategy for Rapidly Doing Chromosomal Walks and Positional Cloning

Overgoes are extremely useful for designing hybridizations. We have found that the strength of the hybridization signal is equal to that of cDNAs and certainly much better than that of end-labeled individual oligonucleotides. Based on this excellent hybridization, our approach to positional cloning is to isolate ESTs or other markers that are close to a region and isolate clones using the hybridization strategy of overgoes.

Overgoes are generated to two separate regions on a marker, and double positive BAC clones are thus picked for further analysis. We use these clones to evaluate the fingerprinted BAC contigs on the Sanger Institute Web site, and then we see if the walk is extended. If it is not, individual clones are sequenced from both ends. Polymorphisms are tested and relative orientation to the mutant gene is determined. New overgoes are designed to all the clones. In a mass hybridization of all the overgoes, we repeat the process, often with four to six overgoes at a time. Before overgoes are designed, we use Repeat Masker and BLAST sequences to see if there is a repetitive sequence presented by a newly designed overgo. In addition, we may test the overgoes on mock filters from the BAC library to see if there is a low copy number repeat. Checking this number helps prevent blackening of filters and allows the walk to move forward.

Hybridization of multiple overgoes produces many clones, perhaps 40 clones at a time. These are individually picked and gridded out on a filter. We then hybridize each of the overgoes separately to define which of the ends are represented on which clones. This hybridization strategy is extremely efficient

because it builds the contig of these BACs. The strategy is also useful for verifying contigs from the Sanger Institute if there is one in your region. Once we determine the direction of walk based on polymorphic markers, we repeat the process again and then assemble the contigs.

VIII. Protocol for Overgo Probing of High-Density Filters

The overgo should be two 24-mers with an 8-hp overlap in the middle; the final product size is 40 bp. Note that there exists online software called Overgo Maker to help you design the overgo that can be found on the Washington University Genome Sequencing Center Web site (Washington University in St. Louis, 20Mc). <http://genome.wustl.edu/gsc/overgo/overgo.html>.

A. Overgo Labeling

The following list is the procedure used for overgo labeling.

1. Heat the stock solution (10 pmol/ μ l) of mixed oligos at 80 °C for 5 min, then at 37 °C for 10 min, and store on ice.
2. Use the labeling reaction (10 μ l): oligo + H₂O = 5.5 μ l. Note that doubling the amount of probe (2x this latter reaction, which makes 20 μ l) results in more consistent high-quality hybridizations. Incubate at room temperature for a minimum of 1 h.

Oligonucleotides	10 pmol each (1 μ l in this case)
BSA (2 mg/ml)	0.5 μ l
OLB (-A, -C, -N6)	2.0 μ l
³² P-dATP ^a	0.5 μ l
³² P-dCTP ^a	0.5 μ l
H ₂ O	To 10 μ l (4.5 μ l in this case)
Klenow fragment	1 μ l (2 U/ μ l)
^a Concentration 3000 Ci/mmol	

3. Remove unincorporated nucleotides using Sephadex G50 columns. Denature probes before adding to blots.

B. Preparation of OLB (-A, -C, -N6) Solutions

OLB (-A, -C, -N6) = A:B:C (1:2.5:1.5)

Solution O

1.25 M Tris-HCl, pH 8.0

Solution A

1 ml solution O
 18 μ l 2-mercaptoethanol
 5 μ l 0.1M dTTP
 5 μ l 0.1M dGTP

Solution B

2M HEPES-NaOH, pH 6.6

Solution C

3 mM Tris-HCl, pH 7.4
 0.2 mM EDTA
 Aliquot and store at -20°C.

C. Hybridization Procedure

It is important to note that the Zon Laboratory team performs hybridizations in a hybridization oven at 58 °C. This temperature works well for overgoes with a GC content between 40 and 60%. Using overgoes that are AT-rich may require lowering the hybridization temperature.

1. A warmed 20-ml hybridized solution is added to a 30 cm x 4 cm hybridized bottle. Filters are first wet down with warmed 2 x SSC, are sandwiched between mesh, and then are inserted into the bottle. With the cap on, the bottle is rotated to allow the filter to unroll slowly and prevent the trapping of trap air bubbles. A maximum of eight filters are placed in the same bottle. Filters are prehybridized for 4 h the first time they are used and 1 to 2 h thereafter. Sheared salmon testes DNA is not used.
2. After prehybridization of the filters, the labeled oligos are denatured at 90°C for 10 min. They are then added to each bottle. Probes are typically allowed to hybridize overnight; however, a hybridization of 2 days gives somewhat stronger signals. This longer procedure time may be useful when working with older filters.

Hybridization Solution	
Composition	For 1000 ml ^a
1% BSA (Fraction V, Sigma)	10 g
mM EDTA	2ml of 0.5 M EDTA @H 8.0)
7% SDS (use 99.9% pure SDS)	70 g
0.5 M sodium phosphate ^b	500ml of 1 M sodium phosphate ^b

^aTo make 1000 ml of solution, use autoclaved deionized H₂O.

^bAn amount of 1 M sodium phosphate must be made as follows (this is really 1 M with respect to sodium but follows the referenced nomenclature): use 134g of Na₂HPO₄·7H₂O, and add 4ml of 85% H₃PO₄, make to 1000ml.

Using the information in the table entitled Hybridization Solution, follow these steps:

1. Heat 100 ml H₂O for 15 sec in microwave on high. Add 10 g BSA, and then stir to dissolve.
2. To 500 ml 1 M sodium phosphate, add 200 ml of H₂O, 2 ml of 0.5 M EDTA, and 70 g of SDS. Stir until the SDS is dissolved (approximately 1 h). Keep in mind that the 1 M solution must be made as follows: use 134 g of Na₂HPO₄·7H₂O, and add 4 ml of 85% H₃PO₄; make to 1000 ml.
3. Add the dissolved BSA, and make volume to 1000 ml.
4. Filter the hybridization solution, and store at 37°C, preventing the SDS from precipitating.

D. Washing

1. Remove the hybridized solution, and fill the bottle to the 213 line with room temperature 2x SSC, 0.1% SDS. Return the bottle to the oven, and rotate the bottle for about 30 min.

2. Transfer the filters to a larger tub on a rotary platform, and wash them as follows:

2 L 1.5 x SSC, 0.1% SDS 58°C 30 min
 2 L 0.5 x SSC, 0.1% SDS 58°C 30 min

E. Autoradiography

1. Seal the filters in plastic wrap, and expose them using XAR5 film at -70°C.
2. Expose the filters overnight, which is usually all the time that is required. This time period can be increased as the filters age (typically a 3-day exposure).

IX. General Flow of Information from the Radiation Hybrid Panel Maps, the Sanger Institute Sequencing Project, and Fingerprinting the BACs

The Sanger Institute has released version 7 of the zebrafish genome assembly, called Zv6. This version includes the whole genome shotgun sequence and is interfaced with 65% sequence from the BAC libraries. The BAC sequence is generated from end sequencing and also from fingerprinted BAC shotgun sequencing projects that are ongoing at the Sanger Institute. Entire BAC contigs are being sequenced and interfaced with the whole genome shotgun sequence. These large contigs provide an incredibly useful point to start looking at your interval. Our laboratory traditionally examines markers from the radiation hybrid panel maps that are in the critical genetic interval, and we BLAST those to obtain the Sanger contigs. Very often, we will find that the same contig contains two individual ESTs on the RH maps in the region. This process establishes that the contig does in fact represent a sequence between the flanking markers.

The Zon Laboratory has also developed two independent programs, available on the laboratory Web site, that evaluate the contigs. The first program involves BLASTing a human gene query to the assembly contigs (Zon Laboratory, 2004a). This procedure will allow any human gene in an interval to be BLASTed in order to find a contig that represents that human gene. Using BLAST in this manner is very useful for examining conserved synteny relationships in zebrafish and human genomes because other genes in the interval may be present. In addition, for our second program, we are using a reverse BLAST in which we have examined the Sanger Institute Zv6 assemblies BLASTed to known human proteins (Zon Laboratory, 2004b). The same can be done between zebrafish and other fish species such as Japanese pufferfish (*Fugu rubripes*) and Spotted green pufferfish (*Tetraodon nigroviridis*) genomes as well as between human and other fish genomes. Reverse BLASTing allows an investigator to see if other genes are present on the contig and can be used again to make a synteny story. In addition, the Ensembl database at the Sanger Institute and the UCSC genome browser has gene annotation in an individual contig, which is a tremendously helpful resource (Sanger Institute, 2004h). We find that in our laboratory, we need to use both Ensembl and our own databases because even though the Ensembl database provides genes, the Web site is difficult to use for directly finding human proteins. Our Web site compares sequences by a BLAST statistical number and allows this type of analysis to occur very rapidly by providing information that is preBLASTed for the entire assembly against the human protein database. From this information, we have determined that a contig exists within our critical interval.

Our laboratory studies contigs that are very close to our gene based on meiotic recombination. We use the assembly sequence, find simple repeats at <http://danio.mgh.harvard.edu/markers/ssr.html> of greater than 12 nucleotides, and use flanking primers in single-stranded length polymorphism (SSLP) analysis to determine whether linkage is available. This process is much better than looking for SSCPs that are present within the interval (although in important areas, we do use SSCPs). By finding CA repeats on a contig in the region, we can quickly study the critical interval and then genotype more embryos that will narrow this interval. Candidate genes on contigs can be tested by rescue assays or by morpholino analysis.

A. PCR Screening of the Pooled PAC and BAC Libraries

In addition to screening the PAC and BAC libraries by hybridizing to filters, we have found it useful to screen a pooled library by PCR (Amemiya and Zon, 1999). It is often helpful to do a PCR reaction using two separate sets of 20-mers oligos (to screen sequences). This process confirms true positives and reduces the number of false positives. The PAC library is used as an example here. Similar pooling can be done using a different library. There are 270 plates (384-well) in the PAC library. All of the wells from one plate are combined to make 270 plate pools. The 33 superpools are created by pooling either eight (superpools 1-27) or nine (superpools 28-33) plate pools.

The first step is to PCR screen the 33 superpools. It is always useful to include a positive control (such as zebrafish genomic DNA) and a negative control (either random plasmid or water). The positive control should have strong bands that preferably amplify with more than one primer pair. Once a superpool positive has been found, the next step is to screen the plate pools and row/column DNA. There are five 384-well plates that are divided into 33 sections corresponding to the 33 superpools. Each section contains the eight or nine plate pools that correspond to the superpool and row/column DNA. There are 16 rows of DNA pools (for rows A*) and 24 columns of DNA pools. For example, a well may contain all the A rows from the eight or nine plates that are in the superpool. The 48 or 49 wells in the section that correspond to the superpool positive should be screened by PCR. This section of our preparation should be diluted before use (-1:30 in TE). This PCR should yield three positive wells: one that corresponds to a plate pool (giving you the plate number), one that corresponds to a row pool (giving you the row letter), and one that corresponds to the column number (giving you the column number).

An example is shown in Fig. 4.

	1	2	3	4	5	6
A	plate 1	pp1 row A	pp1 row I	column 1	column 9	column 17
B	plate 2	pp1 row B	pp1 row J	column 2	column 10	column 18
C	plate 3	pp1 row C	pp1 row K	column 3	column 11	column 19
D	plate 4	pp1 row D	pp1 row L	column 4	column 12	column 20
E	plate 5	pp1 row E	pp1 row M	column 5	column 13	column 21
F	plate 6	pp1 row F	pp1 row N	column 6	column 14	column 22
G	plate 7	pp1 row G	pp1 row O	column 7	column 15	column 23
H	plate 8	pp1 row H	pp1 row P	column 8	column 16	column 24

This shows the upper left section of a 384 well plate containing the plate pools and row/column pools corresponding to superpool 1.

Figure 4. Plate orientation

These "positive" clones can then be retrieved from the PAC library and streaked out on LB/Kan plates. One should confirm that this is an overlapping clone by direct PCR of this clone or by sequencing with a sequence-specific primer.

B. Walking and Establishing Contigs by Sequencing PACs and BACs if a mutant is in a gap region

A situation might occur in which taking the next step in a walk becomes impossible because of a lack of SSCP or SLP markers. Assume that you isolated PAC 1, sequenced both ends, found polymorphic SSCP markers on 60th ends, and oriented the walk by identifying recombinants on both ends (Fig. 5a). According to the data you have acquired, you would make the correct decision to walk to the left (T7 end) because this end has fewer recombinants and is therefore closer to the gene than the Sp6 end is.

1. You now isolate the PACs with the primers from the T7 end. You obtain PACs 2, 3, 4, 5, and 6 (Fig. 5b).
2. You then sequence the ends of all five PACs (although in reality, you may have obtained fewer than five PACs, or you did not acquire reliable sequence from all ends). You may now find that none of the primers have yielded reliable SSCP polymorphisms (although this is quite rare if sequence is available from all 10 ends).

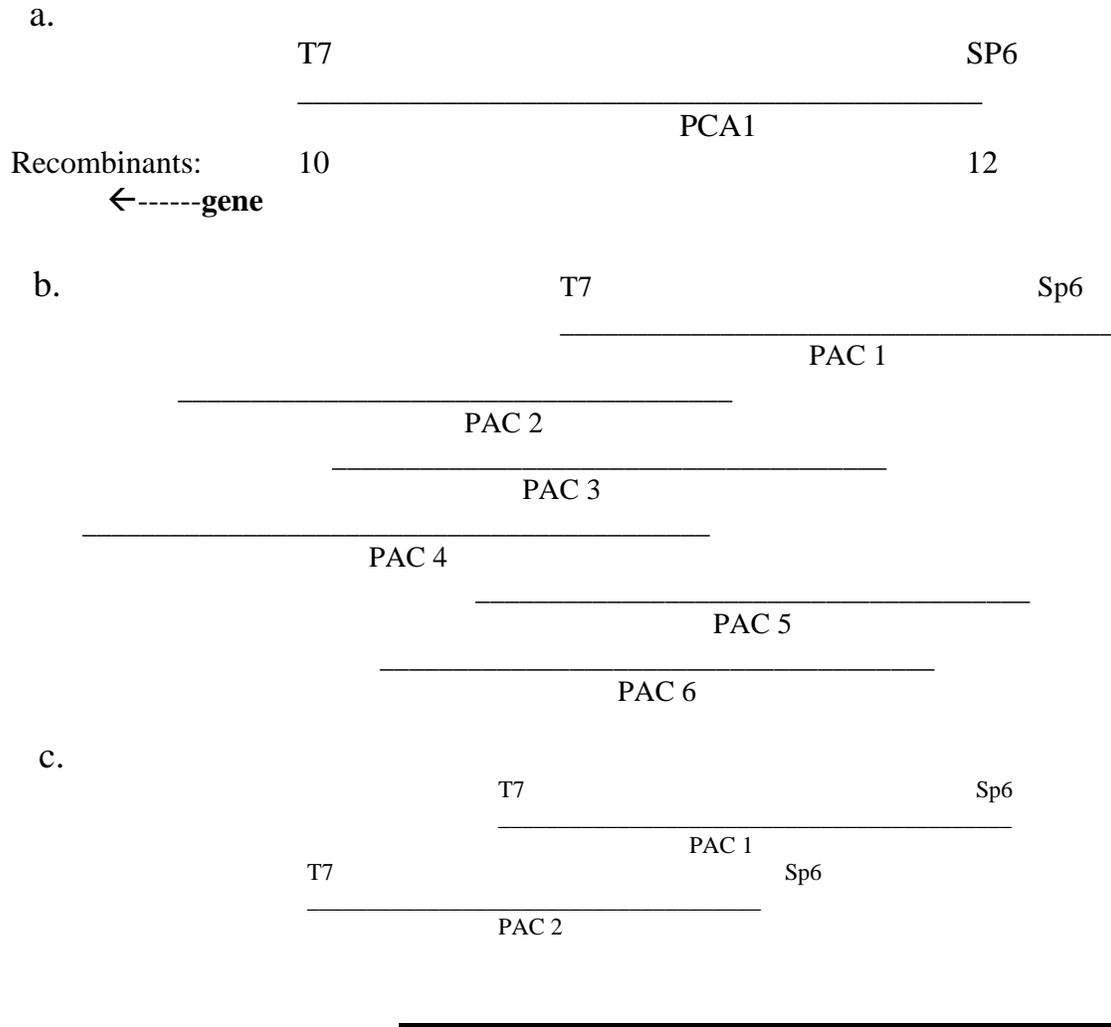


Figure 5. Chromosomal walking using genomic libraries.

3. To address this poor yield, you can take the forward (or reverse) primers from the T7 and Sp6 ends of PAC 2 (or 3,4,5, or 6), 9-end radiolabel them, and then use them as priming oligos to sequence PAC 1. Only one reaction will work (e.g., the Sp6 end of PAC 2 sequenced off PAC 1). You should not walk from the Sp6 end of PAC 2, however, because it points in the wrong direction (Fig. 5c). From your results, you gather that the T7 end points toward the gene of interest.

4. From this point, you can proceed in one of two ways:

a. You can sequence the remaining PACs (3, 4, 5, and 6) with the T7 end of PAC 2. Positive results here indicate that the sequenced PAC extends farther toward the gene than PAC 2 (in our example, this would be true for PAC 4). Negative results indicate that PAC 2 extends farther toward the gene than the sequenced (negative) PACs (in our example, these are PACs 3, 5, and 6). You now use PAC 4 for further walking. To determine the end of PAC 4 that is closer to the gene, use both ends and sequence off PAC 2 (or any other PAC). The end that yields a negative result is the one that is closer to the gene. You can now continue the process as described in the following possible Step 4b.

b. You can use the end that is closer to the gene to isolate another set of PACs (see step 1 described previously). Sequence the ends and attempt to generate polymorphic SSCP markers. If

this is not possible, go to step 2 described previously and continue the process from that step. Depending on your estimated distance from the gene (see recombinants on PAC 1), you can take several steps before attempting to re-estimate the distance to the gene by testing recombinants on polymorphic SSCP markers from the PAC ends.

When referring to step 4a, note that instead of sequencing PAC 1 using PAC 2 primers as described, you can attempt the PCR using the primers derived from the ends of PAC 2 and DNA from PAC 1 as a template. This works in many cases, but it requires stringent controls (positive and negative controls) because PCR is not as specific as sequencing.

In reference to step 4b, note that if you do a number of sequencing reactions simultaneously, you can feasibly only run the " G (or A or T or C) lane because knowing the exact sequence is not as important as ascertaining that the reaction worked.

Because many versions of the genome have been released, our laboratory would typically BLAST Zv6 or other assemblies with BAC or PAC end sequences. This often yields more sequences for finding polymorphisms. In addition, we examine the fingerprinted BAC and PAC contigs to order the region. Note that we still require SSCPs at some point to orient our walk.

X. Synteny Between Human, Zebrafish, Fugu, and Tetraodon Genomes

The Department of Energy (DOE) Joint Genome Institute Japanese pufferfish (*Fugu rubripes*) genome Web site has been updated in 2004 and Spotted Green Pufferfish (*Tetraodon nigroviridis*) are wonderful resources for those involved in positional cloning and comparative genomics (DOE Joint Genome Institute, 2004; and Genoscope and the Broad Institute of MIT and Harvard, 2004). The *Fugu* genome has been shotgun sequenced at a 4x the coverage level. This allows assembly of 10- to 100-kb scaffolds of genomic sequence. The *Tetraodon* genome was sequenced at 8x coverage with an average size of supercontig at 7.62 Mb. Because of the relatively small size of the *Fugu* and *Tetraodon* genome, synteny is extremely helpful for positional cloning in zebrafish. We make use of browser sites at ENSEMBL and UCSC by first finding a zebrafish EST near a mutant gene. We then look at the human genome at browser sites set up by the University of California at Santa Cruz (University of California at Santa Cruz, 2004) and ENSEMBL. We hope to establish if there is zebrafish and human synteny. We look at all the adjacent genes in the human and see if there are zebrafish orthologs. We then use TBLASTN analysis of the *Fugu* and *Tetraodon* genome using the predicted peptide sequence of either human or zebrafish genes. Synteny can usually be found on a scaffold. Each scaffold is nicely annotated on the browser sites and individual project sites, and the sequences can be downloaded and used for BLAST analysis.

Alternatively, one can use gene prediction programs, such as BLAST P and GeneScan. These programs allow investigators to probe the syntenic relationships. When one clicks on the BLAST P or GeneScan program, individual predicted exons are revealed at the bottom of the page. Clicking on the exons gives the predicted homology between *Fugu*, *Tetraodon*, and human genomes, and this result can be cross-referenced to the zebrafish. Clicking on the upper bar gives the predicted peptide of the *Fugu* and *Tetraodon* sequences. Using the predicted peptide of the *Fugu* and *Tetraodon* sequences, we use our TBLASTN server and establish if the Sanger Institute sequence has the zebrafish ortholog of the *Fugu* and *Tetraodon* gene. When putative orthologous genes are not available on the zebrafish genome assembly, they are mapped using the zebrafish radiation hybrid panels to confirm if syntenic relationships occur. We have found that for most chromosomal walks, synteny does exist. These sites are also extremely useful for isolating zebrafish orthologs of human or *Fugu* genes. The reader can browse the Zon laboratory's comparative genomics data (e. g, for human, zebrafish, and *Fugu* genomes] for more information (Zon Laboratory, 2004c).

The Sanger Institute, along with scientists Robert Geisler and Ron Plasterk, has recently undertaken a large-scale fingerprinting project for the CHORI 211 and *Danio* Key Zebrafish BAC libraries using the Hind III enzyme. A physical assembly of the zebrafish genome has been created using fingerprinting

software called FPC. This software is available on the Sanger Institute Web site (Sanger Institute, 2004a), and tiles of fingerprinted genomic clones can be visualized. The name of a BAC can be entered, and a contig may be available. As more BACs are fingerprinted, this physical map will eventually fill in and eliminate the need for chromosomal walking in many instances. The Sanger Institute has recently annotated these tiles with BACs that are to be end-sequenced. It will be useful for investigators to use this end-sequencing to establish polymorphic markers in the region. Using the radiation hybrid panels, these BAC contigs are being mapped on the zebrafish genome by the Zon Laboratory and the Robert Geisler Laboratory, providing an excellent resource for positional cloning (<http://zfrhmap.tch.harvard.edu/ZonRHmapper/>).

A number of sites have been created to help with positional cloning. These sites are listed on the ZFIN database at The Zebrafish Model Organism Database Web site (ZFIN, 2004). We have provided ZFIN with a short description of each of the Web sites.

In a positional cloning project, we find a close polymorphic marker after the high resolution map, and we proceed to evaluate the fingerprinted BAC contigs and available end sequences of BACs in contigs. We then derive markers from those end sequences and determine where on the chromosomal walk the markers may lie relative to the mutant gene. This process will be used increasingly in the future as chromosomal walking will become something of the past.

The Sanger Institute is willing to extend contigs for individuals. An interested individual can simply contact zfish-help@sanger.ac.uk and provide information regarding a chromosomal walk. The assembly team can lower the stringency of the FPC software and thereby extend walks over many more contigs. The fingerprinting software is very specific for matches of Hind III fragments. By relaxing the stringency (where not all Hind III fragments must match), one can extend contigs, but one can also encounter less certainty that the contigs are truly forming a physical linear assembly. Nevertheless, these BACs can be tested by polymorphisms on the mutant mapping panel or RH panel mapping of markers within these regions.

XI. Proving a Candidate Gene Is Responsible for the Mutant Phenotype

What is termed the *rescue* of the mutant phenotype is the gold standard for confirming that a candidate gene is responsible for the mutant phenotype. One can demonstrate that the wild-type gene can rescue the genetic defect (assuming that the defect has compromised the function of the gene rather than generated a gain-of-function mutation). To do this, both wild-type and mutant cDNAs should be subcloned into a vector that is suitable for the synthesis of capped mRNA (i.e., an RNA polymerase site at the 5'-end of the cDNA, a stable 3'-untranslated region containing a polyadenylation signal, and unique restriction enzyme sites for linearization of the plasmid). Commonly used vectors include pCS2+, pSP64T, and pXT7. Both wild-type and mutant cDNA plasmids should be used to generate an *in vitro* translated protein (i.e., ³⁵S-met labeled), which can be resolved on acrylamide minigels and detected by autoradiography. This assay provides an assessment of protein size, stability, and translational efficacy of the wild-type and mutant constructs.

"Run-off" synthetic mRNA transcripts can be generated from the linearized template using commercially available kits, purified, and then resuspended in either RNase-free distilled water or 1 x Danieau's solution (typically at a stock concentration of 1-3 µg/µl). Diluted mRNA (typically 5-500 pg/µl, depending on the potency and toxicity of the encoded protein) is then injected into one to four cell-stage embryos produced by parents carrying the mutation. Amelioration of the phenotype (either by morphology or biochemical assay) is then assessed at the appropriate developmental stage. Generally, the later the onset of the mutant phenotype, the more difficult it is to *rescue*, owing to the instability of the injected mRNA and protein during development. Thus, it is unreasonable to expect every mutant embryo to be rescued. In addition, many proteins, particularly transcriptional regulators when over-expressed early, can have severe effects on early embryonic patterning events, thereby making it impossible to rescue later developmental pathways. Injection of mRNA synthesized from the mutant cDNA plasmid permits an assessment of the severity of

the mutation [i.e., no *rescue* (or morphological effect on development) would suggest a complete loss of function, whereas *partial rescue* may indicate a hypomorphic mutation].

XII. Morpholinos

To evaluate candidate genes, one morpholino is used against the ATG region to prevent translation, and another morpholino is designed against a splice site. Typically, we use the splice donor because this seems to create aberrant splicing. An RT-PCR analysis of the target mRNA products can be extremely useful in showing that there is no normal splice form or a small percentage of normal splice product.

Acknowledgments

The original text for the section entitled Protocol for Overgo Probing of High-Density Filters was courtesy of John D. McPherson at Washington University in St. Louis, Missouri.

References

1. Amemiya, C. T., Alegria-Hartman, M. J., Aslanidis, C., Chen, C., Nikolic, J., Gingrich, J. C., and de Jong, P. J. (1992). A two dimensional YAC pooling strategy for library screening via STS and Alu-PCR methods. *Nucleic Acids Res.* 25; 20(10), **2559-2563**.
2. Amemiya, C. T., and Zon, L. I. (1999). Generation of a zebrafish P1 artificial chromosome library. *Genomics* 1; 58(2), 211-213.
3. BACPAC Resources Center (2004). BACPAC home page. <http://hacpac.chori.org> DOE Joint Genome Institute (2004). Japanese pufferfish *Fugu rubripes*. <http://genome.jgi-psf.org/fugu6/fugu6.home.html>
4. Farr, C. I., Saiki, R. K., Erlich, H. A., McCormick, F., and Marshall, C. J. (1988). Analysis of RAS gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. *Proc. Natl. Acad. Sci. USA* 85, 1629-1633.
5. Johnson, S. L., Africa, S., Home, S., and Postlethwait, J. H. (1995). Half-tetrad analysis in zebrafish: Mapping the *ros* mutation and the centromere of linkage group I. *Genetics* 139(4), 1727-1735.
6. Kane, Zon, and Detrich (1999). *The Zebrafish: Genetics and genomics*. Vol. 60. Academic Press, San Diego.
7. Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, I., Neuhaus, S., Shimoda, N., Driever, W., Fishman, M. C., and Jacob, H. J. (1998). A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nat. Genet.* 18(4), 338-343.
8. Massachusetts General Hospital (2001). Zebrafish microsatellite map server. http://zehralish.mgh.harvard.edu/mapping/ssr_map_index.html.
9. RZPD (2004). Deutsches Ressourcenzentrum für Genomforschung GmbH home page. <http://mvw.rzpd.de/>
10. Shimoda, N., Knapik, E. W., Ziniti, I., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., and Fishman, M. C. (1999). Zebrafish genetic map with 2000 microsatellite markers. *Genomics* 15; 58(3), 219-232.
11. The Sanger Institute (2004a). Sequencing projects BLAST search services. <http://www.sanger.ac.uk/DataSearch/>
12. The Sanger Institute (2004b). Pre! Project Ensembl. <http://pre.ensembl.org>
13. University of California at Santa Cruz (2004) UCSC Genome Bioinformatics. <http://genome.ucsc.edu/index.html>
14. Washington University in St. Louis (2004a). GSC EST projects. <http://genome.wustl.edu/est/>
15. Washington University in St. Louis (2004b). Tools, protocols, and technical information. <http://genome.wustl.edu/tools/>
16. Washington University in St. Louis (2004c). Genome sequencing center. <http://genome.wustl.edu>

17. Wood, W. I., Gitschier, I., Lasky, L. A., and Lawn, R. M. (1985). Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA* 82, 1585-1588.-
18. ZFIN (2004). The zebrafish model organism database. http://zfin.org/zfin_info/dbase/db.html
19. Zon Laboratory (2004a). Human BLAST against zebrafish assembly 3.
<http://134.174.23.160/humanblastzv3>
20. Zon Laboratory (2004b). Zebrafish assembly 3 BLAST human gene.
<http://134.174.23.160/zv3blasthuman/>
21. Zon Laboratory (2004c). Zon lab comparative genomics.
<http://zfblasta.tch.harvard.edu/CompGenomics/>