A Genetic Linkage Map for the Zebrafish


To facilitate molecular genetic analysis of vertebrate development, haploid genetics was used to construct a recombinant map for the zebrafish Danio (Brachydanio) rerio. The map consists of 401 random amplified polymorphic DNA (RAPD) markers and 13 simple sequence repeat spaced at an average interval of 5.8 centimorgans. Strategies that exploit the advantages of haploid genetics and RAPD markers were developed that quickly mapped lethal and visible mutations and that placed cloned genes on the map. This map is useful for the position-based cloning of mutant genes, the characterization of chromosome rearrangements, and the investigation of evolution in vertebrate genomes.

The zebrafish Danio rerio (formerly Brachydanio rerio) (1) is emerging as a model organism for the investigation of the genetic mechanisms of vertebrate development (2). Its short 3-month life cycle and the ease of making both haploid embryos and parthenogenetic diploid fish facilitate the identification and analysis of mutations (3). Saturating the genome with mutations that affect various aspects of the early development of zebrafish seems to be an attainable goal (2). The ability to make stable lines of transgenic zebrafish (4), as well as the ease of making genetic mosaics (5, 6), facilitates the study of gene interactions and gene function.

In spite of the demonstrated potential of zebrafish genetics in understanding vertebrate development, both the genetic and molecular investigation of zebrafish are in their infancy. For example, only two genes identified by mutation have been cloned from zebrafish (7), and only recently have any molecular genetic markers been reported (8, 9). Furthermore, no two genetic markers in zebrafish have been shown to be linked. Because a linkage map is necessary to facilitate molecular genetic analysis of zebrafish development, we constructed such a map with anonymous DNA polymorphisms and then developed ways to place genes known only by sequence or only by mutation on the map with the use of the haploid genetics available in zebrafish.

For a map to be useful, it should contain genetic markers that are abundant, evenly distributed, highly polymorphic, and readily detected in many laboratories. A map of random amplified polymorphic DNA (RAPD) markers hits these criteria (10, 11). A polymerase chain reaction (PCR) with zebrafish genomic DNA as a template and a single, 10-nucleotide-long primer of arbitrary sequence generally amplifies 6 to 12 DNA fragments (9). Different primers amplify distinct DNA fragments. Different strains of zebrafish differ somewhat in their fragment patterns, presumably because of differences in nucleotide sequence at the primer binding sites or deletions or insertions between the primer binding sites. These phenotypes are inherited as Mendelian genetic markers. A dominant allele permits the amplification of a DNA fragment with a specific primer, whereas a recessive allele results in the absence of that fragment (9, 10).

We identified RAPD markers suitable for mapping by performing PCRs with 134 different decamer primers (12) that each amplified several easily scored DNA fragments specific to the DAR (Darjeeling) or AB parental strains (neither of which is completely inbred) (9). We followed these genetic markers segregating among the 94 haploid offspring (13) of a single DAR/AB female fish (called here the linkage map cross). In haploids, recessive alleles are not obscured by their dominant alternatives, and thus the genotype of a haploid can be inferred directly from its phenotype. Of 401 RAPD markers, 203 (50.6%) were from the AB parent, 164 (40.9%) were from the DAR parent, and 34 (8.5%) behaved as codominant markers. Computer-assisted linkage analysis (14) showed that the RAPD markers fell into 29 linkage groups (LGs) that consist of 2 to 25 loci spaced at an average interval of 5.8 ± 5.9 cM (centimorgans) (Fig. 1). Only four scorable RAPD loci (15AC.850, 2AG.440, 17P.710, and 15V.400) remained unlinked to any other marker in our sample.

In addition to RAPD markers, we also mapped 13 of the 16 published simple sequence repeats (SSRs) by implementing a PCR on DNA from haploid embryos of the linkage map cross using published primers (8). Of the three unmapped SSRs, SSR 22 was unlinked to any other marker, SSR 27 did not segregate in the linkage map cross, and SSR 27 did not reliably amplify a product.

Because the mapping of cloned genes facilitates the identification of candidate genes for mutations, we developed a strategy to map genes known by DNA sequence. In this approach, DNA from each haploid of the linkage map cross serves as a template for a PCR with the use of a pair of primers that amplify a fragment containing either the target gene's 3' untranslated region (UTR) or one of its introns. In some cases, this results in a fragment size polymorphism. For example, primers spanning intron 1 in the major histocompatibility complex (MHC) class II β chain locus (15) amplified fragments of 520 and 580 base pairs (bp) that segregated as codominant alleles in the linkage map cross and provided a location on LG VIII (Fig. 1). Likewise, snail1 (16) was mapped to LG XI on the basis of codominant length variants of 210 and 250 bp arising from the CA repeat contained in its 3' UTR. Although primers amplifying the 3' UTR of mssB (17) yielded fragments of uniform size from all 94 haploids of the linkage map cross, the restriction enzyme Msp I cut the DAR allele but not the AB allele, allowing mssB to be localized to LG 1 near lep (Fig. 1).

The utility of a genetic map is related to its degree of completeness. The current map has four more LGs than the number of haploid chromosomes (18), and of the 405 RAPD markers and 1 of the 14 SSR markers remain unlinked to any other marker; thus, a minimum of nine gaps remain to be filled. Assuming that markers are randomly distributed, about 99% of loci are estimated to be located within 20 cM of a marker on the map (19). The fact that only 5 of the 419 DNA polymorphisms studied remain unlinked to any other marker supports this estimate.

A parameter to consider when planning map-based cloning of mutated loci is the relationship of physical to genetic distance. To estimate the average length of 1 cM in kilobase pairs (kb) of DNA, one must compare the size of the zebrafish physical map to the length of its genetic map. The size of the linkage map shown is 2317 cM (Fig. 1). To estimate the size of the complete map, one must account for the distance from the end markers on each chromosome to their adjacent telomeres and any unfilled gaps. This would add 145 cM to account for the telomeres (25 chromo-
somes × 2 telomeres per chromosome × an average of 5.8/2 cM from end marker to telomere), and a minimum of 261 cM to account for at least nine unfilled gaps (5 unlinked markers plus 4 too many linkage groups multiplied by 29 cM, the maximum distance accepted in the map), to give a minimum estimate of the complete map size of 2720 cM. The length of the map can also be estimated from considerations of marker density (20), which yield a value of about 2550 cM. Because the haploid genome of zebrafish contains about $1.7 \times 10^9$ bp (21), this estimate gives an average of about 625 kbp cM$^{-1}$, although the relation of physical and genetic distance may vary from one region of the genome to another (22). In contrast, the mouse genome has on average 1800 kbp cM$^{-1}$ (23).

A strategy that exploits the haploid genomics of zebrafish and the advantages of RAPD markers allowed us to quickly identify loci closely linked to a mutation. This in turn allowed us to place the mutation on the map, thus serving as a prelude to molecular isolation of the gene (24). This strategy was illustrated by the mapping of brass, a pigment pattern mutation (25). Heterozygous brass females were mated to DAR males. From one of the resulting F1 heterozygous females, we collected 38 brass and 38 wild-type haploid offspring. We collected DNA from each embryo individually and then made two DNA pools—one from 20 brass haploids and one from 20 wild-type haploids. We screened this pair of DNA pools with 48 different decamer primers and compared their patterns of amplified fragments. Genetic markers unlinked to from the same 94 haploid siblings was scored for SSRs as described (8), except that 4% agarose gels were used to resolve the DNA fragments. The MHC class II β chain locus was mapped in the linkage map cross with the primer pair Tu385 and Tu360 (15). The eight mutations visible in haploid embryos were placed on the map by means of the protocol illustrated in Fig. 2. Because leopard is not distinguishable in haploids (35), it was mapped in a diploid backcross. Numbers on the left of each LG indicate genetic distance in centimorgans.
brass should have been present in both pools, but those closely linked to the brass locus should have been present predominantly in one pool or the other. The use of haploids instead of diploids in this method [called bulked segregant analysis (26)] eliminates the need to remove heterozygous individuals from phenotypic pools in order to identify markers associated in coupling with either the mutant or the wild-type allele of the gene under study. The analysis identified several fragments that were amplified predominantly or exclusively in one phenotypic pool but not in the other (Fig. 2A). These fragments became candidates for genetic markers linked to brass.

To determine whether any of the candidate markers were indeed closely linked to brass, we scored the 38 brass and 38 wild-type haploid siblings individually for these markers. These experiments confirmed that two of the 48 original primers (13A1 and 16A1) amplified DNA fragments that identify genetic markers linked to brass (Fig. 2B). The 16A1.1190 DNA fragment amplified from all of the 10 wild-type haploids shown but from only 1 of the 10 brass haploids; thus, there was 1 recombinant among the 20 haploids shown. In all, there were 6 recombinants between brass and 16A1.1190 out of 76 haploids, indicating a genetic distance of about 8 cM. Because RAPD marker 16A1.1190 had already been localized to LG XIII (Fig. 1), brass must also be localized to LG XIII. Another DNA fragment amplified by the same primer (fragment 16A1.900) (Fig. 2B) was unlinked to brass (39 recombinants out of 76 haploids) and maps to LG XX (Fig. 1). Primer 13A1 amplified marker 13A1.800, which had no recombinants with brass out of 76 haploids. Because 13A1.800 did not segregate in the original linkage map cross (Fig. 1), however, it was not informative for LG assignment. To verify the location of brass and to make a more complete local map (Fig. 2C), we scored the 76 haploid siblings segregating for brass for other nearby markers in LG XIII. All LG XIII markers that we tested and that were segregating in the cross were found to be linked to brass. The combination of the original linkage map and the local map gave a composite map (Fig. 2C).

With the use of this general strategy, we mapped a total of nine loci identified by mutation. These include the visible mutations leoprd (LG I) (27), brass (LG III) (25), sparx (LG II) (25), and albino (LG XXII), as well as the embryonic lethal mutations floating head (LG XIII) (28), throbless (LG VIII), no tail (LG XIX) (7), silent heart (LG XXVII), and cyclops (LG XI) (5). Linkage analysis revealed that the cyclops mutation is associated with a chromosomal rearrangement that obscures the precise location of cyclops with respect to the other markers shown. The map showed that the homologs of two genes that are linked 15 cM apart on chromosome 17 in mice [MHC class II gene H-2Ab (15) and Brachyury] (23) are unlinked in zebrafish [MHC class II locus (DAB) (15) and no tail (7) (Fig. 1).

Our results indicate that zebrafish embryonic lethal mutations segregating in various genetic backgrounds can be integrated quickly into the linkage map by simply mating a mutant stock to either the DAR or AB lines (or both) and performing haploid-based bulked segregant analysis as described here. A screen with RAPD primers of a number of zebrafish stocks, including the Singapore, Hong Kong, and German lines, showed that most lines had many bands in common with line AB (9), which suggests that most of the loci described here will segregate in typical crosses. The mapping of most of the published SSRs (Fig. 1) further demonstrates that this type of molecular marker is also easily integrated into the current linkage map.

Haploid-based bulked segregant analysis

---

**Fig 1.** (Continued).
with RAPD markers allows the rapid identification of DNA polymorphisms linked to either the mutant or wild-type allele of any gene in zebrafish. In mapping the nine mutations mentioned above, about 1% of the primers screened (18 out of 1778) identified a RAPD marker located within about 5 cM of any given mutant locus; thus, only about 1000 decaamer primers should be needed on average to find a marker about 0.5 cM from any mutation. Some of the mutant genes studied were very close to their nearest RAPD marker—for example, brass failed to recombine with its nearest marker among 76 haploids and floating head failed to recombine with its nearest marker among 1332 haploids (29). A marker 0.5 cM away from a mutation should be on average only about 300 kb from the mutant locus, a distance that corresponds to about four P1 clones and is shorter than a single, large-insert yeast artificial chromosome (30). Once closely linked markers are identified, the PCR products can be readily cloned and sequenced. We have already cloned more than a third of the codominant markers plus a number of dominant markers closely linked to various mutants. Some of the cloned dominant markers segregate as codominants in other crosses or can be converted to codominant markers by restriction enzyme digests (31). The cloned markers can also be used as probes in the initiation of chromosome walks (29).

The genetic map (Fig. 1) permits the analysis of chromosome rearrangements, which can expedite cloning experiments. For example, Talbot (29) has shown that the cyclops 2213 mutation is a translocation between LG II and LG XII. Mapping of noncomplementing mutations that produce related but distinct phenotypes can also help test hypotheses regarding their genetic basis. In addition, analysis of epistatic gene interactions requires distinguishing double and triple mutants from single mutant embryos. Ambiguity can be resolved by the extraction of DNA from pieces of the caudal fin of breeding adults or from parts of experimental embryos and then by assay for RAPD markers closely linked and flanking the mutation.

Such genomic analysis of fish may prove useful in investigations of mammalian genomes. Genetic linkage relations in mammals have been locally conserved over distances averaging up to 8 cM during 100 million years of mammalian divergence (23). Linkage analysis of fish and amphibian genomes has lagged behind, but studies have revealed linkage relations that have survived 400 million years of vertebrate evolution (32), which suggests that the primitive vertebrate gene arrangement may have been largely preserved during fish evolution. Brenner and co-workers (33) have suggested that the ordering of genes in a fish may prove useful in the analysis of the human genome. Because one can systematically collect embryonic lethal mutations in zebrafish (2) [some of which are phenotypically similar to mutations in the homologous genes in mammals (7)] and because the methods reported here will facilitate map-based molecular isolation of these genes, studies of zebrafish embryos are likely to further our understanding of developmental genetic mechanisms conserved among all vertebrates, including humans.

**Fig. 2.** Mapping the brass locus. We collected haploid embryos from a heterozygous brass/DAR female, scored their pigment patterns, and extracted their DNA individually. (A) DNA samples from 20 phenotypically brass (b) haploids were pooled, as was DNA from 20 of their wild-type siblings (+). DNA samples from the two pools were amplified with 48 different RAPD primers, 10 of which are shown (lanes 1 to 4, 6 to 13, and 15 to 22; lanes 5 and 14 are 100-bp size standards; the bright, third band from the bottom of the gel is 600 bp). Most fragments appeared in both pools, but fragments 13Al.800 and 16Al.1190 and a few others seemed to be present only in one phenotypic pool or the other, as expected for DNA fragments from closely linked genetic markers (26). (B) To determine if marker 16Al.1190 was closely linked to brass, we scored the DNA of 38 brass (brs+) and 38 wild-type haploid siblings (brs-) from a brass/DAR mother for the ability to amplify the 16Al.1190 fragment (upper arrowhead); the results from 20 of these haploids are shown here (lanes 2 to 21). The phenotype of the embryo in lane 20 reflects a recombination event between the 16Al.1190 marker and the brass locus. Marker 16Al.900 (lower arrowhead) assorts independently of brass. Similar experiments were conducted with the other candidates identified in the bulk segregant analysis experiments. (C) To integrate brass on the linkage map, we scored in haploid offspring of the brass/DAR mother other markers previously shown (Fig. 1) to be linked to 16Al.1190. Marker 1V.800 showed 6 recombinants with brass out of 60 haploids. Marker 5Y.450 did not segregate in the cross. The local map (DAR x brass) constructed from these data was incorporated into the larger linkage map with markers segregating in both crosses.

**REFERENCES AND NOTES**

12. Primers and their sequences can be obtained from Operon Technologies (Alameda, CA). Marker names use the inverse of Operon's primer label (to avoid confusion with zebrafish allele designations) followed by the apparent size of the amplified fragment [for example, genetic marker 20Y.670 (A) results in a 670-bp DNA fragment amplified from the AB haplotype by Operon primer Y.20; a (D) suffix indicates marker origin in the DAR stock].
13. Haploid embryos were produced by in vitro fertilization. Eggs obtained by gently pressing the abdomens of anesthetized animals were mixed with sperm with similar genotypes and treated with ultraviolet light. Such sperm activate embryonic development but do not make a genetic contribution to the offspring. Haploids developed until hatching at 3 days, when their DNA was collected. Several hundred PCR assays can be performed on the DNA from a single haploid embryo.
14. Linker (a program written by M. Hutchinson) compares all possible pairs of markers and computes the frequency of recombinants between pairs, whether or not in a coupling or repulsion phase. Linker ignores DNA fragments that were not present in 50% of the samples (range, 40 to 60%, chi square test; P = 0.02). Linker analyzed the data set pairwise and showed significant linkages (chi square test; 0.001 > P > 0). The data set was filtered to include only marker pairs with less than a 26% frequency of recombinants, which converts
Abnormal Development of Peripheral Lymphoid Organs in Mice Deficient in Lymphotixin

Piero De Togni,* Joseph Goellner, Nancy H. Ruddle, Philip R. Streeter, Andrea Fick, Sanjeev Mariathasan, Stacy C. Smith, Rebecca Carlson, Laurie P. Shornick, Jenna Strauss-Schoenberger, John H. Russell, Robert Karr, David D. Chaplin†

Mice rendered deficient in lymphotixin (LT) by gene targeting in embryonic stem cells have no morphologically detectable lymph node nodes or Peyer's patches, although development of the thymus appears normal. Within the white pulp of the spleen, there is failure of normal segregation of B and T cells. Spleen and peripheral blood contain CD4+CD8+ and CD4-D8+ T cells in a normal ratio, and both T cell subsets have an apparently normal lytic function. Lymphocytes positive for immunoglobulin M are present in increased numbers in both the spleen and peripheral blood. These data suggest an essential role for LT in the normal development of peripheral lymphoid organs.

Lymphotixin (LT, also designated TNF-β) is a soluble product of activated lymphocytes that was first defined by its cytotoxic activity against fibroblasts (1-3). LT is now recognized to be produced by activated CD4+ T helper cell type 1 (T1H1) lymphocytes, CD8+ lymphocytes, and certain B lymphoblastoid and monocytic cell lines (4-6). The gene encoding murine LT is located 1100 base pairs (bp) upstream of the evolutionarily related gene encoding tumor necrosis factor-α (TNF-α) within the major histocompatibility complex (MHC) (7, 8). The LT and TNF-α proteins are structurally related and show similar activities in vitro and when given to experimental animals (9). In solution, LT is a homotrimer with a conformation similar to that of TNF-α. A membrane-associated form of LT has been described, consisting of a heterotrimeric complex containing two LT monomers together with a 33-kD transmembrane protein designated LT-β (10). The gene encoding LT-β is located immediately centromeric to the gene encoding TNF-α. The biological effects of LT and TNF-α are mediated by two receptors, designated p55 and p75 (11).

In vitro, LT and TNF-α can modulate many immune and inflammatory functions. As implied by their names, both cytokines are cytotoxic for a variety of transformed and normal cell types (12-14). In order to be killed, target cells must express LT- TNF-α receptors, with the p55 receptor appearing to mediate the cytotoxic response (15). LT and TNF-α can augment the proliferation of activated thymocytes (16).