Insertional mutagenesis in zebrafish identifies two novel genes, *pescadillo* and *dead eye*, essential for embryonic development

Miguel L. Allende, Adam Amsterdam, Thomas Becker, Koichi Kawakami, Nicholas Gaiano, and Nancy Hopkins

Center for Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Recently our laboratory described an efficient method for generating retroviral provirus insertions in the zebrafish germ line, and we showed that provirus insertions induce embryonic mutations at a frequency of roughly one mutant per 70 insertions. To date we have isolated four insertional mutants and, using the proviruses as a molecular tag, have cloned the genes disrupted in three of them. The proviruses in all three mutants lie within or just 5' of the first coding exon, point in the opposite transcriptional orientation from the gene, and disrupt transcription. Here we present a molecular characterization of two genes identified by this method and describe the associated mutant phenotypes. The *pescadillo* (*pes*) gene is predicted to encode a protein of 582 amino acids with no recognizable functional motifs, which is highly conserved from yeast to humans. *pes* mRNA is expressed widely and dynamically during the first 3 days of embryogenesis. Prominent sites of expression are the eyes and optic tectum on day 1, the fin buds, liver primordium, and gut on day 2, and the branchial arches on day 3. Beginning at day 3 of embryogenesis, *pes* mutant embryos exhibit small eyes, a reduced brain and visceral skeleton, shortened fins, and a lack of expansion of the liver and gut, and then die on the sixth day of development. The *dead eye* (*dye*) gene encodes a protein of 820 amino acids that is homologous to genes of unknown function in human, mouse, and *Xenopus*, and that has weak homology with the yeast NIC96 (nucleoporin-interacting component) gene. *dye* mutants can be recognized on day 2 of embryogenesis by the presence of necrotic cells in the tectum and eyes. *dye* mutants die on day 5 of development. These results demonstrate the power of insertional mutagenesis in zebrafish for rapidly finding and characterizing novel genes essential for embryonic development. Using our current methodology, we estimate that our laboratory could screen -25,000 insertions in 2-3 years, identifying perhaps 250-350 embryonic lethal genes. Assuming that all genes are accessible to proviral insertion, the wider application of this approach could lead to the rapid identification of the majority of genes that are required for embryonic development of this vertebrate.

*Key Words: Danio rerio; mutant; teleost; organogenesis; branchial arches*

Received August 20, 1996; revised version accepted October 29, 1996.
roughly, that there are ~2400 genes with essential or visible functions in the fish embryo and that ~50% were identified in the chemical mutagenesis screens [Haffter et al. 1996].

Despite the wealth of new genetic information emerging from chemical mutagenesis screens in the zebrafish, knowledge about the molecular nature of the affected genes and their products will not be immediately forthcoming. Cloning the mutated genes will depend upon the development of reagents for positional cloning in the zebrafish. For now this technology remains laborious and expensive because of the large size (~1.6 x 10^9 bp) of the zebrafish genome.

As an alternative to chemical mutagenesis, recently we developed a method for generating insertional mutants in zebrafish using integration of retroviral proviruses into the genome [Lin et al. 1994; Gaiano et al. 1996a,b]. Although the frequency of mutagenesis is considerably lower than that of chemical mutagenesis, the molecular tag provided by the inserted retroviral provirus allows the immediate isolation of flanking genomic fragments, which are likely to include the disrupted gene. Of the four zebrafish insertional mutants isolated thus far, we quickly cloned genes disrupted in three of them. We believe these disrupted genes are likely to be responsible for the phenotypes of the respective mutants. Cloning was extremely rapid because in all three cases the provirus that is linked genetically to the mutant phenotype integrated close to coding sequences of the gene it disrupted, because so many gene sequences are now available in the database, and because the disrupted genes are highly conserved evolutionarily (this study; Gaiano et al. 1996b).

Here we describe the molecular characterization of the genes mutated in two insertional mutants pescadillo (pes) and dead eye (dye) and present a preliminary characterization of the mutant phenotypes. The pes gene encodes a novel protein of unknown function that is very highly conserved across species: Homologous sequences are present in human, mouse, Caenorhabditis elegans, and yeast. The dye gene encodes a protein homologous to predicted proteins identified in human, mouse, and Xenopus [Nagase et al. 1995; Hudson et al. 1996]. Vertebrate dye proteins share limited similarity with the product of a yeast gene NIC96 (nucleoporin-interacting component of 96 kD), which has been shown to be essential for viability [Grandi et al. 1993].

These results provide strong support for the prediction that genetic analysis in zebrafish will identify many novel genes essential for vertebrate development. If the retroviral mutagen we have used integrates at random into the fish genome, given the efficiency of mutagenesis we have observed to date, it should be possible within a few years for a number of fish laboratories together to identify and clone the majority of the genes essential for the early development of this vertebrate species.

Results
In an ongoing insertional mutagenesis screen, visible or lethal phenotypes were sought among the progeny of crosses between pairs of F2 fish heterozygous for a single identical proviral insertion [Gaiano et al. 1996b]. We screened by observing at least 25 live embryos under a dissecting microscope on days 1, 2, and 5 after fertilization, and scored for consistent abnormalities visible in 25% of the embryos as described by Haffter et al. [1996]. Four recessive lethal mutations tightly linked to proviral insertions have been identified. The no arches (nar) mutation and gene are described elsewhere [Gaiano et al. 1996b]. The 80A mutation has not been studied further as a disrupted gene has not yet been identified for this mutant. Here we describe the genes and phenotypes associated with proviral insertions 67D and 404, which have been named pescadillo (allele designation pes^{hi2}, referred to hereafter as pes) and dead eye (allele designation dye^{hi2}, referred to hereafter as dye), respectively.

Evidence that dye is an insertional mutant and preliminary characterization of the mutant phenotype

The dye mutation was recovered from a cross between two F2 fish heterozygous for a single proviral insertion designated 404 [Fig. 1]. Defects are first apparent in embryos at day 2 of development when mutant embryos display necrosis of the tectum and the eyes. By day 5, most of the anterior head structures are reduced, including the eyes and the forebrain. Most, if not all, of the pharyngeal skeleton is absent, and the tectum and cerebellum are barely discernible. The embryos display edema and fail to develop a swim bladder, characteristics present in most embryonic lethal mutations in fish.

Because our zebrafish are not inbred and background mutations are present, we first sought evidence of whether the dye mutation was caused by proviral insertion. PCR analysis showed that in crosses of fish heterozygous for the 404 provirus insertion, all mutant embryos were transgenic, a result consistent with linkage of the insertion to the mutation. To determine whether the

Figure 1. Wild-type (top) and dead eye (bottom) zebrafish embryos at 72 hr after fertilization. dye mutants are recognized by the small head structures and protruding eyes. Bar, 100 μm.
proivirus and mutant phenotype are tightly linked, we tested whether mutant embryos are invariably homozygous for the insertion-bearing chromosome. Using inverse PCR we cloned a fragment of fish DNA adjacent to the provirus and identified a single-copy probe from this fragment. This probe yielded bands on a Southern blot that were diagnostic for either the transgenic chromosome or for its nontransgenic homolog. We used this probe on a Southern blot of DNA samples isolated from 57 mutant embryos and 110 phenotypically wild-type siblings (data not shown). All the mutant embryos were homozygous for the chromosome carrying the 404 insertion, whereas wild-type embryos were either heterozygous or were homozygous for the chromosome lacking the insertion. This result indicates that the insertion is tightly linked to the mutation and suggests that the insertion caused the mutation.

To obtain a preliminary characterization of the dye mutant phenotype, we prepared sections of mutant and wild-type embryos at day 5 after fertilization (Fig. 2). By this stage, dye/dye embryos show severe defects in the structure of the brain and cranial skeleton (Fig. 2B,D). Although the major brain subdivisions (fore-, mid-, and hindbrain) can be recognized in the mutant (Fig. 2B,D), they are substantially reduced in size compared with the same structures in the wild type (Fig. 2A). There is also a near complete absence of the pharyngeal skeleton, and only the posterior neurocranium is evident.

Low-power microscopic analysis revealed necrosis in the brain of dye/dye mutant embryos relative to wild-type controls. Necrosis was particularly evident in the brain and cranial skeleton of dye mutants (Fig. 2B). At this stage, the forebrain (f) and tectum of the midbrain (t) are grossly reduced, the hindbrain (hb), albeit smaller, appears less affected. Cartilage can be seen underlying the caudal brain in the dye mutant but anteriorly it is short and fragmented when compared to wild type (A vs. B. white arrowhead). The trabeculae, normally connecting the ethmoid plate and the posterior neurocranium, jut into the eyes (e, in D), which are recessed mediocaudally in dye mutants [D]. In pes mutant animals, the brain, particularly the tectum (t), is smaller than in wild type, the neurocranium is shorter and thicker, and the posterior pharyngeal arches (arrows) lack differentiated cartilage (A vs. C). At the position of the five branchial arches, however (arrows in C), mesenchymal tissue is organized into segmental bundles. The liver (li) is very reduced and the yolk (y) has not been consumed. (e) Eye; (f) forebrain; (h) heart; (hb) hindbrain; (li) liver; (t) tectum; (hb) hindbrain; (y) yolk; (asterisk) Meckel’s cartilage; (black arrowhead) hyoid; (black arrows) branchial arches; (white arrowhead) anterior neurocranium. Bar, 100 μm.

**Figure 2.** Sagittal sections of 5-day-old wild-type (A), dead eye (B,D), and pescadillo (C) mutant embryos. Anterior is to the left and dorsal is up. The plane of section is medial in A,B, and C, and mediolateral (to include the eye) in D. dye mutant embryos lack most structures in the ventral head such as elements of the visceral skeleton (indicated by an asterisk, black arrowhead and arrows in A, cf. B). At this stage, the forebrain (f) and tectum of the midbrain (t) are grossly reduced, the hindbrain (hb), albeit smaller, appears less affected. Cartilage can be seen underlying the caudal brain in the dye mutant but anteriorly it is short and fragmented when compared to wild type (A vs. B, white arrowhead). The trabeculae, normally connecting the ethmoid plate and the posterior neurocranium, jut into the eyes (e, in D), which are recessed mediocaudally in dye mutants [D]. In pes mutant animals, the brain, particularly the tectum (t), is smaller than in wild type, the neurocranium is shorter and thicker, and the posterior pharyngeal arches (arrows) lack differentiated cartilage (A vs. C). At the position of the five branchial arches, however (arrows in C), mesenchymal tissue is organized into segmental bundles. The liver (li) is very reduced and the yolk (y) has not been consumed. (e) Eye; (f) forebrain; (h) heart; (hb) hindbrain; (li) liver; (t) tectum; (hb) hindbrain; (y) yolk; (asterisk) Meckel’s cartilage; (black arrowhead) hyoid; (black arrows) branchial arches; (white arrowhead) anterior neurocranium. Bar, 100 μm.
Isolation of the dye gene and demonstration that its expression is disrupted in mutants

If insertion 404 is responsible for the dye mutation, a prediction is that it lies in or near, and disrupts the expression of, a gene expressed during embryonic development. We cloned sequences adjacent to the 404 insertion and sequenced 1.4 kb on the 3’ and 2.1 kb on the 5’ side of the provirus. This 3.5-kb sequence was used to search the GenBank data base using the BLAST algorithm (Altschul et al. 1990). This search identified a region of 180 nucleotides whose predicted translation product was highly homologous to the first 60 amino acids of proteins encoded by genes from human, mouse, and Xenopus. This sequence begins 450 bp to the 5’ side of the provirus, such that the transcription units of the provirus and the putative gene are divergent (Fig. 4A). The human gene was cloned from the myeloid cell line KG1 and designated KIAA0095 (Nagase et al. 1995; GenBank accession no. D42085). The Xenopus gene, which is 84% identical to the human gene, was identified by differential display PCR and designated An4a (Hudson et al. 1996; GenBank accession no. U63919). In addition, translations of expressed sequence tags (ESTs) from both human and mouse were highly homologous to these 60 amino acids. Both the human and Xenopus genes share large regions of homology with the Saccharomyces cerevisiae NIC96 gene, whose product is part of the nuclear pore complex (Grandi et al. 1993, 1995); however, the BLAST search with the fish genomic sequence did not identify the NIC96 gene as there is no homology between the yeast and vertebrate genes in the first 60 amino acids.

Two nested oligonucleotides corresponding to sequences within the putative zebrafish protein coding region were used to perform 3’ rapid amplification of cDNA ends (RACE) from 48-hr embryonic RNA (see Materials and Methods; Fig. 4A), resulting in the isolation of a 2.4-kb cDNA fragment, approximately the size expected based on the human and Xenopus genes. This product was subcloned and both strands were sequenced. In addition, RT–PCR was performed between a downstream primer matching sequences within the 3’ RACE product and an upstream primer corresponding to the genomic sequence, which produced a fragment of the expected size (450 bp; Fig. 5). This fragment was also sequenced, confirming that the RNA that produced the 3’ RACE product includes the genomic sequences identified adjacent to the proviral insertion. Because both the 3’ RACE and RT–PCR products were generated from 48-hr RNA, the dye gene must be transcribed by this time, when the phenotype is first apparent in the mutant.

The RT–PCR and 3’ RACE sequences were compiled and translated. The predicted amino acid sequence is shown in Figure 4B, along with the predicted amino acid sequence of the human and Xenopus homologs. The zebrafish gene is 83% identical to the human gene (with one gap) and 79% identical to the Xenopus gene over its entire length. In addition, when the full-length predicted protein sequence of dye was used in a data-base search, significant but limited homology was found to the yeast NIC96 gene product, as was found previously with the other vertebrate dye homologs [Nagase et al. 1995; data not shown].

To determine whether the 404 proviral insertion affects expression of the dye gene, we performed RT–PCR on RNA extracted from phenotypically wild-type or mutant embryos obtained from a cross of two fish heterozygous for this insertion. As shown in Figure 5, a dye-specific RT–PCR product is detected when RNA from phenotypically wild-type embryos is used as a template, whereas no band can be detected when RNA isolated from phenotypically mutant embryos was used. This indicates that expression of the dye gene is abolished or is decreased beyond the level of detection by proviral insertion 404.

The pes gene encodes a highly conserved novel protein

Previously we reported that the proviral insertion designated 67D is linked genetically to, and presumably caused, the zebrafish pescadillo mutation (Gaiano et al. 1996b). This insertion lies within the first coding exon of a gene that was discovered because of high homology between genomic sequences flanking the 67D insertion and a human EST [GenBank accession no. R13806]. The regions of homology were presumed to be exon sequences and were used to design primers for RT–PCR. Analysis of the amplified RT–PCR products confirmed the predicted intron–exon structure in the region and showed that the provirus lies 80 bp upstream of the putative methionine initiation codon [Gaiano et al. 1996b; data not shown].

To learn more about the pes gene and its encoded protein, we screened a zebrafish cDNA library prepared from day 3 embryonic mRNA [kind gift of Dr. Kai Zinn, California Institute of Technology, Pasadena] using a 300-bp cDNA fragment amplified by RT–PCR as a probe. Two clones with cDNA inserts of apparently equal size were obtained and one was sequenced. This clone includes a 2214-bp insert. A putative protein coding region was identified in this sequence based on homology with the human EST sequence. This coding region corresponds to the longest open reading frame (ORF) found and encodes a protein of 582 amino acids [Fig. 6]. The region surrounding the presumed methionine ATG initiation codon conforms to the consensus translation start site [Kozak 1984] and is preceded by an in-frame
Zebrafish insertional mutants pcs and dye

Figure 3. Cell death in dye mutants. [A,B] Two-day-old wild-type (wt, A) and dye mutant (dy, B) embryos are shown. Necrosis is evident in mutants but is apparent restricted to the ectoderm (indicated by arrowheads) and eyes (arrows). (C,D) Antimycin chloride, which causes the onset of apoptosis within the wild-type (white arrowhead) and mutant (white arrowhead) embryos, is given at the time of wild-type (black arrow) and mutant (white arrow) injection. (E,F) Serial sections through the head of day 2 wild-type (E) and dye mutant (F) embryos. Note the necrotic zone in the ectoderm (black arrow) and eye (white arrow) of mutant embryos in F. Antimycin A is to the left and dorsal to up, (C,D) Bar, 100 μm; (E,F) bar, 2 μm.
Figure 4. Proviral insertion 404 lies adjacent to the dye gene. (A) A sequence of 3.5 kb flanking proviral insertion 404 was cloned by inverse PCR (see Materials and Methods). The region of homology to the human KIAA0095 and Xenopus An4a genes is shaded. The provirus has inserted 450 bp upstream of the putative ATG and its transcription runs in the opposite direction to that of the open reading frame. The position of primers 404-2 and 404-3, used in 3' RACE, is indicated. (B) The predicted amino acid sequence of the dye gene was compared with that of the human and frog genes using the Lasergene alignment tool. Amino acids identical between any two or all three of the genes are shaded. There is a one amino acid gap at position 215 in the human gene. Amino acid identity between the fish and human genes is 83%, between the fish and Xenopus 79%, and between the human and Xenopus 84%. Amino acids 1-60 of the dye gene were obtained from the genomic sequence originally found to lie adjacent to the proviral insertion, amino acids 49-820 were deduced from the 3' RACE product. Four independent RT-PCR reactions were carried out between primers 404-2 and 404-8 or 404-9 (see Materials and Methods) and the product from each was subcloned and sequenced to confirm the continuity of the expressed sequence. All four RT-PCR isolates contained two base pair changes relative to the 3' RACE product. Amino acids 111 and 114 reflect the sequence contained in these RT-PCR products. The 5' end of the dye gene was not determined, thus the assignment of the translation start site is based on the homology to the human and Xenopus genes. The zebrafish dye cDNA sequence has been deposited in the GenBank database (accession no. U77595).

stop codon located 21 nucleotides upstream. The 5' and 3' untranslated regions are 99 and 368 bp, respectively. The entire cDNA sequence upstream of the proviral insertion point is contiguous with the genomic DNA, suggesting that the insertion is within the first exon. It cannot be ruled out, however, that the cDNAs obtained are incomplete and that there are additional exons further upstream. No poly[A] sequence was found at the 3' terminus of this cDNA.

Three human cDNA clones containing the EST identified by computer search were obtained from ATCC and the longest was sequenced in its entirety. Comparison of the putative proteins encoded by the human and zebrafish transcripts show 74% identity (Fig. 6).

Further data base searches using the predicted pes amino acid sequence were carried out against the translated GenBank database (TBLASTN) [Altschul et al. 1990]. Significantly homologous sequences were identified in cosmids clones of genomic DNA isolated in the C. elegans and yeast genome projects [GenBank accession no. D75131 and yeast ORF name YGR103W, respectively], and in mouse EST sequences [GenBank accession no. AA003101]. Amino acid identity between the zebrafish pes protein and the S. cerevisiae homolog is ~54% in the amino-terminal third of the protein, 39% overall (Fig. 6).

Analysis of the predicted pes protein sequence did not reveal any known motifs that would suggest a possible

Figure 5. Proviral insertion 404 disrupts transcription of the dye gene. Embryos from a cross between fish heterozygous for insertion 404 were sorted by phenotype on day 3, and RNA was prepared from pools of wild-type or mutant embryos. RT-PCR was performed either with primers to detect the dye transcript (lanes 2-5), or the pes transcript (lanes 6-9) as a positive control. Lanes 2, 3, 6, and 7 used wild-type RNA as template; lanes 4, 5, 8, and 9 used dye/dye RNA as template. Wild-type embryos contain both dye and pes transcripts (lanes 2 and 6), whereas dye/dye embryos express pes but not dye (lanes 8 and 4). Lanes 3, 5, 7, and 9 are the products of reactions lacking reverse transcriptase. Both products span introns, and in both cases the bands were isolated, subcloned, and sequenced to confirm that they truly represent their respective genes. (Lane 1) 123-bp ladder.
Zebrafish insertional mutants pes and dye

A highly acidic region at the carboxyl terminus of the protein is conserved among the zebrafish, human, and yeast genes (underlined in Fig. 6). Most of the pes protein is predicted to have an α-helix structure as determined by analysis with the Robson/Garnier secondary structure algorithm (M. Robinson, pers. comm.).

Expression of the pes gene is developmentally regulated

If mutation of the pes gene is responsible for the mutant phenotype in pes/pes embryos, gene expression would be expected to occur at or before the mutant phenotype becomes visible at day 3 of embryogenesis. Thus, we analyzed the timing and tissue distribution of expression of the pes gene during embryogenesis by Northern blot and in situ hybridization.

A 300-bp pes cDNA fragment was labeled radioactively and hybridized to RNA prepared from several embryonic stages and adult fish in a Northern blot [Fig. 7]. Two transcripts, 2.2 and 1.9 kb, were detected, with the larger being more prevalent during embryogenesis. Unfertilized eggs and gastrulating embryos (6 hr after fertilization) have low levels of the transcripts, whereas strong zygotic expression is seen by 12 hr. Transcript levels decrease after ~24 hr. In adult fish, only females have detectable pes RNA and dissection of the ovaries shows that it is restricted to this organ [Fig. 7].

A 2-kb cDNA fragment of the pes gene was used to synthesize digoxigenin-labeled RNA probes for whole mount in situ hybridization. A sense-strand-specific probe did not produce signal at the stages examined. Using an antisense-strand probe, pes transcripts could not be detected in 3- or 6-hr embryos, but were detected beginning at 12 hr in the eye and brain primordia (not shown). Between 18 hr and 3 days of development, the expression of pes message is highly dynamic. At 18–24 hr, strong expression is detected in the eye, forebrain, tectum, and somites, whereas lower levels of transcript are seen in the hindbrain and in cells flanking the hindbrain [Fig. 8A, data not shown]. During the second day of development [28–36 hr] expression levels begin to decrease and by 48 hr the distribution of transcripts is restricted to the ganglion cell layer of the eye, the ventral forebrain, cells in the posterior tectum at the midbrain–hindbrain boundary, rows of cells in the pharyngeal arches, the pectoral fin buds, the liver and pancreatic primordia, and the presumptive gut [Fig. 8B, data not shown]. At 72 hr, the tectal row of cells continues to express pes transcripts. RNA levels in the liver are lower than at 48 hr, whereas the developing gut continues to express high levels of pes mRNA. Striking expression is seen in a series of stripes coincident with the expanding branchial arches [Fig. 8C]. The distribution of pes transcripts in embryos older than 72 hr was not analyzed in

Figure 6. Amino acid sequence alignment of the predicted zebrafish pes protein with the human and yeast homologs as compiled by the Lasergene alignment tool. Shaded regions correspond to identities among the sequences. Underlined amino acids correspond to a highly acidic region conserved in pes homologs. The zebrafish and human cDNA sequences have been deposited in the GenBank data base (accession nos. U77627 and U78310, respectively).

Figure 7. Northern analysis. A pes cDNA fragment was hybridized to total RNA isolated from the indicated embryonic stages and adult zebrafish by Northern blot analysis. The numbers shown correspond to hours after fertilization; (ue) unfertilized egg RNA; (ov) ovary. The last four lanes on the right correspond to RNA from adult female, adult male, adult female without ovary, and ovary. A 2.2-kb product is prevalent during embryogenesis, whereas a 1.9-kb band is expressed at low levels in the embryo but is the more abundant product in adult ovary. RNA from adult male and from females whose ovaries have been removed have no detectable pes product.
Figure 8. Restricted expression of pes between days 1 and 3 after fertilization. (A) Twenty-eight-hour-old embryo whose yolk has been removed, observed dorsally; anterior is to the left. Note heavy expression in the eye primordia, in the forebrain, the tectum (arrowhead), and in the asymmetrically localized liver primordium (arrow). (B) Thirty-hour-old embryo was dissected as in A and shows heavy hybridization in the developing liver (arrow) and in the fin buds (arrowheads). Anterior is to the left. (C) Seventy-two-hour-old embryo. Expression of pes in the tectum is reduced to a row of cells (arrowhead), compare with stain at 28 hr (A). Expression in the pharyngeal arches can be seen at this stage (arrow). The pancreas and gut express heavily (dark stain to the right of arch staining). Anterior is to the left and dorsal is up. (A, C) Bar, 100 μm; (B) bar, 50 μm.
Figure 9. Pharyngeal cartilage formation is defective in pes mutants. (A) Ventral views of 3 (top) and 5 (center) day-old wild-type (wt) and mutant (pes) embryos stained with alcian blue. Note the differing head sizes between wild type and mutant at day 3 and the lack of growth of cartilage in the mutant by day 5. The principal cartilaginous elements of the wild-type and mutant pharyngeal skeleton are shown schematically below the stained embryos: the first arch (P1), including Meckel’s cartilage (m) and palatoquadrate (pq); the second arch (P2), including the hyomandibular (hm), the paired ceratohyals (ch) and the median basisi hyal (bh), and branchial arches, P3–P7. In the mutant, both elements of Meckel’s cartilage have fused at the midline, the hyomandibular is reduced, the ceratohyals are pointing ventrocaudally, and branchial arch cartilages are absent. (B) Lateral views of day 5 wild-type (wt) and mutant (pes) embryos stained with alcian blue. The short, ventrally protruding Meckel’s cartilage (m) and ceratohyal (ch) are the only stained structures seen in the pharyngeal region of the mutant. Bar, 100 μm.
Allende et al.

head region of pes/pes embryos (see Fig. 2B). Occasionally, a few cells within the first (P3), and less often the second (P4), branchial arch stain lightly by day 5, and in some mutant embryos, muscle fibers can be seen juxtaposed to the undifferentiated branchial arch primordia (not shown). In addition to the arch defects, the ethmoid plate, the anterior portion of the neurocranium, appears shortened in mutant embryos and cartilage in the pectoral fins is reduced, resulting in short fins (Fig. 9A, top; data not shown).

Transverse sections through the trunk region reveal that the internal organs of mutant and wild-type embryos are indistinguishable in size at day 3 of embryogenesis [Fig. 10 cf. A with B]. However, between days 3 and 5 of development striking differences appear. In wild-type embryos the liver grows and extends over the yolk surface and, concomitantly, the yolk is rapidly consumed (Fig. 10C). In mutants the liver does not grow substantially after day 3 and the yolk is not consumed (Fig. 10D). Moreover, the gut is markedly reduced in mutants and the anterior expansion of the intestine that forms the stomach in cyprinid fishes (Takashima and Hibiya 1995) does not develop. In addition to the striking differences in the expansion of the liver and gut, the pancreas, which appears darkly labeled by the counterstain in sections of 5-day-old wild-type embryos, was not

Figure 10. Comparison of day 3 and day 5 cross-sections through the anterior trunk region. (A,B) At day 3, relative sizes of liver (thick arrow) and gut (thin arrow) in the mutant (B) are virtually indistinguishable from wild type (A), as are trunk muscles (s) and hindbrain (h). Both animals were sectioned at the level of the pectoral fins (f). Circulating blood is visible (b). (C,D) At day 5, the liver (thick arrows) has greatly expanded in wild type (C), but not in the mutant (D). Also in the mutant the yolk (y) is unconsumed, the gut (thin arrows) has failed to expand, the body wall muscles (bm), which surround the abdomen, are absent, the axial musculature has started to degenerate, and all body cavities have expanded and filled with fluid. The renal tubules (arrowheads), however, appear to be of similar size and position in mutant and wild type. (A,B) Bar, 100 µm; (C,D) bar, 50 µm.
discernible in mutants. Furthermore, in mutants the axial musculature shows signs of degeneration and the body wall muscle, which surrounds the abdominal cavity in wild-type embryos, is absent (Fig. 10 cf. C with D).

At the level of resolution presented here, there appears to be a striking correlation between defects in pes mutant embryos and regions where pes is normally expressed at high levels earlier in development: the branchial arches, liver and gut, as well as brain, eyes, fin buds, and ethmoid plate. In most cases the primordia develop, but between days 3 and 5 they fail to expand. Importantly, other structures appear normal in mutants at day 5, although many other regions of the animal are severely affected. Tissues that appear normal include the notochord, the pronephros, and renal tubules (Fig. 10 cf. C with D; data not shown). Expression of pes message was not detected in these tissues at any stage in wild-type animals. To establish firmly that the pes phenotype is restricted entirely to cells that normally express pes, product will require further experiments. Likewise, we cannot yet rule out that pes mRNA expression occurs at some point during embryogenesis in tissues that appear normal in mutant animals.

Discussion

We have presented evidence that dye and pes are insertional mutants and we have described the genes whose disruption is likely to be responsible for the mutant phenotypes. In the case of dye, we showed that [1] a single proviral insertion is linked genetically to the mutant phenotype, [2] the provirus lies just 5' of the putative ATG codon of a gene whose transcript is expressed in embryos; and [3] the insertion abolishes detectable expression of this gene. In the case of pes, previously we had shown genetic linkage of a single provirus to the mutant, shown that the provirus lies in the 5' exon of a gene that is expressed in embryos, and shown that the provirus abolishes detectable gene expression. Here we have provided additional strong support that pes is an insertional mutant in the gene we identified by showing that at least many sites of pes expression in wild-type embryos correspond to regions of the embryo that fail to reach normal size in pes/pes mutants. The evidence presented suggests that it is highly likely that mutations in the pes and dye genes are responsible for the mutant phenotypes observed. Definitive proof will require either rescue of the mutant phenotype by introducing the gene or its product into mutant animals, or possibly targeted mutation of these genes. Neither technology has yet been reported in zebrafish. Targeted disruption could be performed in mice (Mansour et al. 1988; Joyner et al. 1989; Schwartzberg et al. 1989; Zijlstra et al. 1989) as the pes and dye genes are so highly conserved among vertebrates. Only a positive result would be informative in a knockout experiment, however, as mice and fish may differ in their genetic redundancy and as homologous genes could serve different functions even among vertebrate species.

Ease of molecular cloning of genes disrupted by proviral insertion

The virtue of insertional mutagenesis is the ease of cloning the mutated genes. In mice, this virtue has not always been realized. DNA insertions in the mouse genome frequently cause complex rearrangements at the site of insertion, which can make it difficult to identify the mutated gene responsible for the observed phenotype (Jaenisch 1988). Retrovirus proviral insertions in mice can disrupt genes by integrating into coding or noncoding exons, introns, or regulatory regions. The location of the provirus relative to recognizable coding regions or transcripts determines whether it is easy to locate the gene whose disruption causes the mutant phenotype.

In our case, we identified rapidly genes disrupted in three of the four insertional mutants (Table 1). The most recent case, dye, required only 3 weeks of work by a

<table>
<thead>
<tr>
<th>Insertion no.</th>
<th>Mutant name</th>
<th>Allele</th>
<th>Disrupted gene</th>
<th>Onset of phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>38M</td>
<td>no arches</td>
<td>nar^hi1</td>
<td>nar, homologous to Drosophila clipper, a zinc finger ribonuclease^a</td>
<td>Early day 3</td>
<td>Gaiano et al. [1996b]</td>
</tr>
<tr>
<td>67D</td>
<td>pescadillo</td>
<td>pes^hi2</td>
<td>pes, homology to unknown genes in human, mouse and yeast unidentified</td>
<td>Late day 3</td>
<td>This work and Gaiano et al. [1996b]</td>
</tr>
<tr>
<td>80A</td>
<td></td>
<td>80A^hi3</td>
<td>unidentified</td>
<td>Late day 2</td>
<td>Gaiano et al. [1996b]</td>
</tr>
<tr>
<td>404</td>
<td>dead eye</td>
<td>dye^hi4</td>
<td>dye, homology to genes in human^b and frog, similarity to yeast NIC96 protein^c</td>
<td>Early day 2</td>
<td>This work</td>
</tr>
</tbody>
</table>

^aBai and Tolias [1996].
^bNagase et al. [1993].
^cHudson et al. [1993].
^dGrandi et al. [1993].
single individual. The reasons for this speed were (1) the proximity of the mutagenic proviral insertions to coding sequences, and (2) the fact that the coding sequence was homologous to sequences present in the database. We have not yet located a gene near the 80A insertion. Possibly we do not yet have enough sequence data (1 kb on one side, 4 kb on the other). Alternatively, this insertional mutant may involve a gene sequence that is not in the database or one that is not evolutionarily conserved, or the 80A insertion may lie in a large intron or in a regulatory sequence distant from the coding region of the putative disrupted gene.

It is interesting that the proviruses in three of the four mutants we isolated integrated just upstream of the ATG initiation codon of a gene and that all three point in the opposite transcriptional orientation from the gene they disrupt. These numbers are too small to allow conclusions, but the following points are relevant. Despite much effort, whether mouse and chicken C-type retroviruses integrate randomly into the host cell genome remains controversial [Vijaya et al. 1986; Rohdewohld et al. 1987; Shih et al. 1988; Scherdin et al. 1990, Withers-Ward et al. 1994]. Some studies suggest preferred integration into 5' ends of genes and into actively transcribed regions, whereas others suggest that integration is essentially random across the genome. In many studies biological selection is operating. For example, in our case only insertions that disrupt genes will be detected as mutants. The position of the integrations we have observed in mutants might reflect the most probable way for the virus to disrupt gene expression. Provirus insertions into introns, for example, might not affect transcription of fish genes as they sometimes do in mice. We have sequenced 1–2 kb of DNA adjacent to 42 randomly selected proviruses present in transgenic lines of fish that did not harbor embryonic mutations [K. Kawakami, N. Gaiano, D. Grosshans, M. Allende, A. Amsterdam, T. Becker, and N. Hopkins, unpubl.]. The data do not reveal preferred sequences and are consistent with random integration, although such data are limited.

It has been reported that in mice approximately one in 20 provirus integrations causes an embryonic phenotype [Jaenisch 1988]. The frequency we have obtained for fish insertional mutants [approximately one in 70] is considerably lower. However, our numbers are still very small, and the genetic requirements for embryonic development may not be comparable between mice and fish. In addition, it will be important to learn whether mutagenic insertions are limited to the 5' ends of genes in the fish. If so, reengineering the viral vector might increase its mutagenicity at other locations, increasing significantly the target size for gene disruption and thus the frequency of mutants.

Relationship of insertional mutants to previously identified zebrafish mutants

Even more important than the location of the provirus relative to the gene it disrupts is the related question of whether all genes can be mutated by proviral insertion. Most insertional systems show bias. In Drosophila, P elements readily target only about one-third to one-half of the genes that can be mutated to an embryonic phenotype by chemical mutagenesis [Spradling et al. 1995]. Even if this is the case with retroviruses, we might expect that the collection of insertional mutants will ultimately contain a distribution of phenotypic classes similar to that encountered in the chemical mutagenesis screens. So far, the four insertional mutants we identified show distinct phenotypes that become manifest relatively late in embryogenesis (although the dye phenotype begins to be visible by day 2). It is probably not surprising that these insertional mutants present posthatching defects, as most of the chemical mutants did.

Although most late-appearing mutants were discarded in the chemical screens [Driever et al. 1996; Haffter et al. 1996], some were kept that affected specifically the jaw, liver, and gut, all of which are relatively late-developing structures. Among the collection of chemically induced mutants classified as having specific defects in pharyngeal arch development, one designated babyface (bab) has deficiencies in the cranial skeleton that are strikingly similar to those seen in pes mutants [Schilling et al. 1996]. Complementation tests will be needed to determine whether bab or other chemically induced mutations are allelic to pes.

The chemical mutagenesis screens isolated a large number of mutants whose primary phenotype was necrosis restricted to the central nervous system [Furutani-Seiki et al. 1996]. However, because of the large number of mutants that fell into this class, complementation tests were not performed and therefore, the number of genes involved remains unknown. Based on the phenotype observed here, it appears that dye would belong to this class of mutants.

Understanding the dye and pes mutant phenotypes

The dye and pes mutant phenotypes described here are both distinctive. dye embryos exhibit extensive necrosis of the central nervous system, whereas pes embryos fail to expand normally at least many of the primordia in which the pes gene is expressed: the branchial arches, liver, gut, brain, eyes, fin buds, and ethmoid plate. To understand the molecular basis for the mutant phenotypes we observed, it will be necessary to understand the normal cellular function of the products of the dye and pes genes. pes is a novel gene and does not contain any motifs that might suggest a clearly identifiable function, making the task demanding. In the case of dye, a possible function is suggested by its relation to the yeast NIC96 gene, whose product is part of the nuclear pore complex. The NIC96 protein interacts physically with several nucleoporins, including nucleoskeletal-like protein 1 [NSP1], the yeast homolog of the vertebrate p62 nucleoporin that is required for the assembly of transport-competent nuclear pore complexes [Finlay et al. 1991; Grandi et al. 1993, 1995; Zabel et al. 1996]. NIC96 is an essential gene for cell growth in S. cerevisiae and appears to be required for the formation of nuclear pore com-
plexes (Grandi et al. 1993; Zabel et al. 1996). Although the degree of homology between dye and NIC96 is statistically significant, to determine whether they perform a similar biological function will require further experiments.

Transcripts for the pes gene are contributed maternally to the egg, but then decrease during the first hours after fertilization. Although this could reflect a “housekeeping” function for pes, some patterning genes in other organisms also display maternal and zygotic expression. During embryogenesis, pes expression increases rapidly during the first day of development, and later decreases gradually. No expression can be detected in adult animals suggesting a specific embryonic requirement for this gene. The results described here suggest that the pes gene product may be required for the growth of a subset of embryonic organs. Regions of the embryo that express the gene heavily subsequently fail to reach normal size in mutants. Particularly striking in this respect are the gill arches, liver, and gut. In contrast, regions of wild-type embryos in which expression was not detected by in situ hybridization appear normal in mutants, including the notochord, pronephros and renal tubules. The defects observed in pes/pes embryos suggest that cell types of diverse embryonic origins are affected and, moreover, that the expression pattern of the gene is not restricted to specific cell types.

The fact that the dye and pes mutations exhibit distinctive embryonic phenotypes, including defects in the growth or maintenance of specific tissues, suggests that they may be important in cell biological processes, although possibly not in pattern formation and morphogenesis. Such genes could also have medical relevance. Currently we are mapping their homologs on human chromosomes to determine whether these genes correspond to disease loci.

Feasibility of genetic screens

Forward genetic approaches are usually powerful only when large-scale screens are possible, because the number of genes that affect any one developmental process is small. Although this will also be true for insertional mutagenesis in zebrafish, the situation is somewhat different because the mutants that we isolate can be viewed within the broad picture provided by the chemical mutagenesis screens. Nonetheless, we have attempted to develop a technology that will make large scale mutagenesis screens possible, and we are continuing to try to increase the number of insertions that can be screened.

Although our current mutagenesis frequency is ~70-fold lower than that in the chemical screens, in chemical mutagenesis, as the mutations are unmarked, it was necessary to obtain about five successful pair matings for every mutant identified (Mullins et al. 1994; Solnicka-Krezel et al. 1994; Driever et al. 1996; Haffter et al. 1996]. Because we track insertions by PCR and Southern blot and only mate fish heterozygous for the same lesion, we require only one successful cross to identify a mutant. Thus, the amount of work, although clearly greater in our case, is not 70 times greater per mutant than in chemical mutagenesis screens.

Given our experience to date, and at the mutation frequency we have observed, we estimate that eight scientists will be able to screen ~25,000 insertions in 2–3 years, obtaining ~250–350 mutants. If the distribution of phenotypes among these mutants proves to be similar to that in the chemical screens, ~30% of these should have specific early developmental defects and at least several should affect the patterning and morphogenesis of almost every embryonic structure.

The zebrafish has long been admired for the ease with which early development can be visualized and it has been pursued as a model system because of the possibility that the organism might serve as a tool for rapidly identifying genes essential for vertebrate development. The results presented here suggest that insertional mutagenesis, in conjunction with chemical mutagenesis screens, may help the fish to realize this potential.

Materials and methods

Animals

Zebrafish (Danio rerio) were kept and raised essentially according to standard conditions (Westerfield 1995) and using practices established in our laboratory (Culp et al. 1991). The aquarium systems used were designed specifically for housing large numbers of animals in small containers (Mullins et al. 1994) and were purchased from K.-J. Schwarz Glas Aquarienbau (Göttingen, Germany). Fertilization was achieved by natural spawning and embryos were raised at 28°C and staged according to Kimelman et al. (1995). The insertional mutant pilot screen was carried out by inbreeding fish harboring identical proviral insertions (Gaiano et al. 1996b) and scoring their progeny for several morphological criteria under low magnification as described (Haffter et al. 1996). Identified mutants have been named according to the conventions established for zebrafish (Westerfield 1995; M. Mullins, pers. comm.); the superscript letters indicate the laboratory designation [b] and the insertional nature of the mutation [l].

Isolation of genomic sequence flanking proviral insertion 404

Inverse PCR was used to clone genomic DNA fragments on each side of proviral insertion 404. Genomic DNA from fish heterozygous for the 404 insertion was digested with either Ncol and BspHI (for the 3' flanking sequence) or BglII (for the 5' flanking sequence), extracted with phenol/chloroform, and ethanol precipitated. T4 DNA ligase was then added to the DNA diluted to 2 μg/ml to circularize the fragments. PCR was then carried out with high fidelity PCR system (Boehringer Mannheim) and pairs of primers from the provirus such that one primer was in the long terminal repeat (LTR) oriented 5' to 3' toward the genomic DNA and the other was in the middle of the virus oriented 5' to 3' toward either the Ncol or BglII site: N5 (GTAAGATCTCAGGTAGTTGACTACCCGT) and N1 (GTACTCTATAGCTTTACGSTCG) were used for the 3' flanking sequence and NU5 (GTAAGATCTCGAGGCCAACCTCTAGG) and NU1 (GTAAGATCTCGAGGCCAACCTCTAGG) were used for the 5' flanking sequence. The PCR program was 30 cycles of 94°C for 15 sec, 55°C for 30 sec, 68°C for 4 min [plus 20 sec per cycle after
cycle 10), preceded by a 2-min denaturation step at 94°C and followed by a 10-min extension step at 72°C.

3' RACE and RT–PCR

RNA was isolated from day 2 or 3 embryos by the guanidinium hydrochloride method [Westerfield 1995]. 3' RACE for the isolation of the dye gene and RT–PCR for the detection of both dye and pes transcripts was carried out with a commercially available kit (GIBCO-BRL Life Sciences) according to the manufacturer’s instructions, with the use of expand high fidelity PCR system (Boehringer Mannheim) for the PCR. For 3' RACE, first strand synthesis from 1 μg of day 2 RNA was performed using the supplied AP primer. One round of PCR was performed on 10% of this sample using the supplied AUAP downstream primer and an upstream primer [404-2, CATGGATACT-GAGGGTTTGGGGAGC] that overlaps the first 8 amino acids of the ORF found in the genomic sequence. This reaction was fractionated on a low melt agarose gel, DNA from a 2.6-kb band unique to the presence of both primers was purified and 2% of this sample was used as a template for a second round of PCR in which the upstream primer [404-3, CACCA-GAACCCTTCAGACACAGC] overlapped amino acids 49–56 of the same ORF. The 2.4-kb product was subcloned into PB-luecscript II (Stratagene) for sequencing. Both rounds of PCR used the same PCR program as the inverse PCR except that the annealing temperature was 60°C.

For RT–PCR, 0.5 μg of RNA from either wild-type or dye day 3 embryos was used for first strand synthesis with a primer specific to either the dye gene [404-9, TGCCTGGACCAG- CAGGACG] or the pes gene [P20, TACTCTCTGAAATTGGC- CAAGC]. PCR was performed on 10% of this sample using either 404-2 and 404-8 [TCCCAGCAGGGTGTCGAAC] or 404-2 and P20 and P15 [TGCAAAGCTTCTG- GAGCCACCGTATAG] that lies upstream of 404-9, or P20 and P15 that overlaps amino acids 49–56 of the pes ORF. The PCR program was 30 cycles of 94°C for 30 sec, 60°C for 1 min, 72°C for 2 min, preceded by a 2-min denaturation step at 94°C and followed by a 5-min extension step at 72°C.

Northern blot and cDNA isolation

For the Northern blot analysis, total RNA (15 μg) from each sample was fractionated on a 2 M formaldehyde agarose gel and transferred to a nylon filter (Hybond N+, Amersham). The blot was probed with a radioactively labeled 292-bp RT-PCR cDNA fragment [corresponding to nucleotides 61–353 of the pes cDNA]. Exposure was performed on Kodak BioMax MS film for 6 days. The same cDNA fragment was used as a probe to screen a 3-day embryonic cDNA library (gift of Dr. Kai Zinn). Plaques [5 × 103] were screened and two positive clones were identified and isolated. Both clones contained inserts of identical length by PCR and one of them was sequenced in its entirety on both strands. The human cDNA clones were purchased from American Type Culture Collection (Rockville, MD), and were sequenced on both strands. Sequence alignment was accomplished by using the Lasergene software (DNAStar, Inc.).

Histology and acridine orange labeling

Alcian blue staining was done essentially as described [Dingerkus and Uhler 1977; Gaiano et al. 1996b]. The nomenclature for skeletal elements is that described by Schilling et al. [1996]. For tissue sectioning, embryos were fixed in 4% paraformaldehyde/PBS, dehydrated and embedded in Polybed 812 epoxy resin (Polysciences). Specimens were cut into 1- to 2-μm sec-

tions that were counterstained subsequently with a solution of 0.05% crystal violet, 0.01% methylene blue, and 0.05% borax at 95°C.

For detection of apoptotic cells, anesthetized embryos were injected with a solution of 1 μg/ml of acridine orange [Sigma] into the yolk sac (Furutani-Seiki et al. 1996). Embryos were allowed to recover for 30–60 min and were visualized under a Nikon Microphot SA microscope with an EPI-FL3 fluorescence attachment using a 450- to 490-nm excitation filter and a 520-nm long pass emission filter.

In situ hybridization

UTP-11 digoxigenin-labeled RNA probes were prepared as suggested by the manufacturer (Boehringer Mannheim Biochemicals). The probe used was an in vitro transcription product of a 2-kb fragment of the pes cDNA (corresponding to nucleotides 214–2214 of the pes cDNA). Anesthetized embryos were fixed in 4% paraformaldehyde in PBS at 4°C for 12–16 hr and were dehydrated in methanol at −20°C for at least 1 hr. In situ hybridization was carried out essentially after Jowett and Lettice [1994]. Protease K treatment was for 10 min at 10 μg/ml for embryos up to 24 hr old and at 25 μg/ml for older embryos. Prehybridization and hybridization temperature was 65°C with a probe concentration of 1 μg/ml. Hybridized embryos were cleared in glycerol or in methyl salicylate and were photographed under a Leica Wild M3Z dissecting scope or a Nikon Microphot SA microscope.

Acknowledgments

We thank Dean Thompson, Elizabeth Abrams, and Brett Hayward for fish care, Kai Zinn for the cDNA library, Catherine Willet and Agustin Zapata for advice on histology and comments on the manuscript, Patricia Reilly for instruction on tissue sections, David Grosshans for sequencing help, and Murray Willen and Agustin Zapata for advice on histology and preparation of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References


Bai, C. and P.P. Tolias. 1996. Cleavage of RNA hairpins medi-

3154 GENES & DEVELOPMENT
Zebrasfish insertional mutants pcs and dye


Insertional mutagenesis in zebrafish identifies two novel genes, pescadillo and dead eye, essential for embryonic development.

M L Allende, A Amsterdam, T Becker, et al.

Genes Dev. 1996, 10:
Access the most recent version at doi:10.1101/gad.10.24.3141

References
This article cites 36 articles, 18 of which can be accessed free at:
http://genesdev.cshlp.org/content/10/24/3141.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.