Honors College Thesis

Genetic mapping of

the *golden (gol^b1)* mutation in zebrafish

Summer Lee Street

Approved [redacted] Date: 26 June, 1998

John H. Postlethwait (Thesis Advisor)

Abstract

The pigment pattern pathway can be studied in zebrafish to further the understanding of early vertebrate development. To fully genetically characterize the pigment mutation *golden*^b1 in zebrafish, the gene must be mapped and the disrupted molecule identified. In 1986, George Streisinger hypothesized that the golden^b1 mutation would map far from its centromere. This hypothesis has been confirmed by mapping *golden*^b1 to zebrafish linkage group 18, far from the centromere. Tyrosinase-Related Protein-2 (TRP2), a molecule involved in the production of the black pigment melanin, is hypothesised to be the molecule disrupted by the *golden*^b1 mutation. Mapping TRP2 in the zebrafish genome will provide further understanding of the pigment pattern pathway in zebrafish and possibly humans.

Acknowledgments

Many thanks to John H. Postlethwait for all of his help and patience over the past three years. I will be ever-grateful for his guidance and support. Nathan Tublitz, Eric Selker and Henry Alley waded through many copies of this thesis and provided me with helpful suggestions. Thanks! Thanks also to all of the other researchers in the Postlethwait Lab. Lastly, thanks to all of those trees who supplied me with endless reams of paper.

Table of Contents

Figure Legends

Introduction

Every three-and-a-half minutes, a baby is born in the United States with a birth defect (http://www.cdc.gov/nceh/programs/infants/brthdfct/prevent/dd_pub.htm). Birth defects affect more than 150,000 babies each year, and are considered the leading cause of infant death--as well as a major cause of disability in young people (http://www.modimes.org/bdefects.htm). These birth defects range from treatable conditions such as spina bifida[1] and cleft lip[2] to devastating problems like cyclopopia[3]. While some of these conditions are due to environmental factors such as the mothers' use of illicit drugs, tobacco or alcohol, or a lack of folic acid--others are genetically based. The causes of many congenital abnormalities are yet unknown.

The investigation of human birth defects is difficult, partly because scientists cannot normally conduct experiments on humans or their fetuses. Such experiments, however, are easy to perform on certain model systems such as mouse and fish. Importantly, genetic control mechanisms are shared between fish and all other animals with backbones (known as

[1] A neural tube defect that causes a failure of fusion of the arches of the vertebrae, typically in the lumbar region (Thompson *et al.*, 1991).
[2] A common congenital malformation that originates due to a failure of fusion of the frontal process with the axillary process; generally occuring around the 35th day of gestation (Thompson *et al.*, 1991).
[3] A condition where the forebrain does not fully develop and only one eye is formed in the middle of the head (Roessler *et al.*, 1996).

chordates), including humans. Yet even in the mouse, one has difficulty studying early development, the period when most birth defects occur, because the embryo cannot be easily seen, since it is hidden in the uterus.

In contrast to mammals, fish embryos develop externally and are transparent. Zebrafish, in particular, have a high reproduction rate, large clutch size and small body size. All of these characteristics help facilitate the study of early vertebrate development, making them much easier to study than most vertebrates. In addition to the similar developmental abnormalities, vertebrate embryos have many other similar traits. The 28-day-old human embryo and the 20-hour-old zebrafish embryo have similar somites[4], gill arches, notochord, and tail bud, as well as the placement of heart, ears, eyes--to name only a few of the similarities (Fig 1). Furthermore, more than five thousand zebrafish mutants that block embryo development have been recently described and are available for extensive investigation. Because genes controlling development are similar among animals (Fouquet *et al.*, 1997), many of these zebrafish mutants are very similar to human developmental abnormalities. One striking example of this is the mutation *cyclops*, which occurs in both zebrafish (Roessler *et al.*, 1996) and humans. Because of this high degree of similarity, the zebrafish is a very good model system. Thus studying

[4] Somites form the repetitive segmented structures of the vertebrate body (Alberts et *al.*, 1989).

zebrafish will give us a better understanding of general vertebrate development.

Chapter One:

The *golden* mutation: Defining the problem

One aspect of development which is readily visible is pigmentation. Pigmentation is caused by the production of melanin. Melanin is a dark brown/black pigment found in the skin, retina and hair. Melanin is produced in a melanocyte, an epidermal cell which comes from the neural crest (Gilbert et al., 1997). The cells of the neural crest break loose from the epithelium along the line where the neural tube pinches off from the future epidermis. These cells migrate independently out through the mesoderm, and will form most of the peripheral nervous system (Alberts et al., 1989).

A defective production of melanin results in albinism, a complete lack of pigment in hair, retina and skin cells Albinism is caused by a mutation in tyrosinase function, an enzyme involved in melanin synthesis (Boissy *et al.*, 1996). Tyrosinase is the key enzyme in melanin synthesis Tyrosine is a precursor of melanin. The oxidation of tyrosine by tyrosinase leads to the formation of melanins. Tyrosine metabolizm in the melanocytes causes the appearance of spots and stripes in many animals as well as the varied shading of yellow, red, brown and black mammalian hair (Halaban et al., 1997).

The *golden (gol^b1)* mutation in zebrafish was the first zebrafish mutation discovered by Charlene Walker and George

Streisinger in the late 1970's (Streisinger et al., 1981). This mutation occurred spontaneously and was isolated from a descendent of zebrafish from a local pet store. Two day old embryos expressing the *golden* mutation contain no melanophores[5] on the body and no black pigment in the eyes (Streisinger et al., 1986). This makes the *golden* embryo easily distinguishable from the wildtype[6] embryo which has large black melanophores on the body and black pigmented retina (Fig 2). In the adult *golden* mutant fish, there are rows of unusually small, black pigment cells, and the xanthophores (yellow pigment cells) are much more prominant than in the wildtype fish (Streisinger et al., 1986)(Fig 3). These changes make the adult fish appear much more golden in color than a wildtype adult, and hence the name "golden" for this mutation. There are two possible explanations for the expression of the mutation. The gene for tyrosinase could be directly effected; or, a gene for a product which modifies the activity of tyrosinase could be affected. Because there is some pigmentation present in *golden* mutant embryos, we suspect that it is not a mutation that completely eliminates the activity of the tyrosinase enzyme.

Another mutation in zebrafish with a similar phenotype to *golden* is that of *brass*, which mapped to linkage group 13 (Postlethwait *et al.*, 1994). Two day old embryos homozygous for the *brass* mutation contain large melanophores with light-

[5] Cells containing melanin.
[6] The normal or usual form.

brown pigment, rather than black, along with brown pigmented retina in the eye (Striesinger *et al.*, 1986). In the adult, the *brass* mutant contains no black pigment, prominent rows of xanthophores are expressed, and the eye is deep ruby red (Striesinger *et al.*, 1986). The phenotype of both of these zebrafish mutants is similar to the human condition Brown Oculocutaneous Albinism (OCA3). This mutation is located on Hsa 9p23. Recently, pigment mutations in humans have been associated with a protein called Tyrosinase Related Protein-1 (TRP1). OCA3 is a mutation in the TRP1 gene that causes a lack of expression of the TRP1 protein in their melanocytes (Boissy *et al.*, 1996). A TRP1 mutation has been discovered in a pair of African-American human fraternal twins. The affected twin is much lighter in pigmentation than his unaffected, normal type brother. The affected twin suffers from Brown Oculocutaneous Albinism (Brown OCA). The mutants *golden* and *brass* in zebrafish both have defective melanocyte expression. This suggests that *golden* or *brass* may be a mutant in a TRP gene. The goal of my thesis is to test this hypothesis.

Testing the hypothesis

One test of the hypothesis that *golden* or *brass* is a mutant in a tyrosinase related protein, is to compare the location of these mutations with the location of the TRP gene on the zebrafish genetic map. Mapping consists of placing

genetic markers[7], genes and mutations on a representation of the chromosomes of a particular species. Our lab has generated an extensive map of the zebrafish genome (Postlethwait *et al.*, 1998). Because *brass* is already mapped to linkage group 13, the work in this thesis focuses on the mapping of *golden* and the mRNA that encodes the TRP2 gene. Mapping these mutations and their genes may aid our understanding of a particular developmental process. Three types of markers are useful for making gentic maps: genes, mutations and anonymous DNA polymorphisms[8]. Genes, like the TRP gene encode RNAs. Mutations alter phenotype, like the *golden* or *brass* mutations. Anonymous DNA polymorphisms change the sequence of DNA in a detectable way. I used two types of anonymous DNA polymorphisms, CA repeats and RAPDs. The description of these techniques will be provided in the Materials and Methods section. It becomes possible to identify the molecule that is disrupted by the mutation, once that mutation is mapped. Therefore, mapping provides us with a tool to understand how a mutation on a specific gene interfers in development. An additional benefit to mapping is that the genetic map provides insights into the overall function and grouping of genes on the chromosome. From an evolutionary standpoint, much insight can be gained by determining the specific arrangement of genes on a chromosome and comparing genomic maps among related organisms.

---

[7] Sequences used to mark specific places on a chromosome.
[8] The occurrence of differences in sequence between homologous chromosomes.

Although about 150 zebrafish genes and 20 mutations have been mapped, previous work had failed to map *golden*. The hypothesis that *golden* is located far from the centromere was based on half-tetrad analysis, similar to tetrad analysis done in fungal genetics (Chakrabarti *et al.*, 1983). The work presented in this thesis confirms this hypothesis. To test whether *golden* is a mutation in a TRP gene, it was first necessary to map the *golden* mutation.

Chaper Two:

Methods and Materials

Animals

To map *golden*, a purebreeding *golden* male was mated to a female of the SJD strain with normal pigment. The SJD strain is genetically highly divergent from the *golden* strain. This mating produced F1 hybrid individuals that had one normal copy of the *golden* gene, the *golden* plus allele, and one mutant copy of the *golden* gene, the *golden* minus allele. To map TPR2, I used a deletion mutant of *golden*, the *golden^b6* allele, which removes much of the chromosome bearing the *golden* gene. The *golden^b6* siblings were mated to each other. This mating produced F1 hybrid individuals that were homozygous for the *golden^b6* mutation.

Wildtype and mutant stocks are maintained as described by Westerfield (1995). The *golden^b6* allele is a recessive mutation in zebrafish. Wildtype and mutant *golden* embryos are easily separable on the basis of pigmentation. The haploid embryos were used in haploid-based bulk segregant analysis (BSA) as described elsewhere (Postlethwait *et al.* 1994).

Mapping *golden*

Two types of anonymous DNA polymorphisms where used: the RAPDs (Postlethwait *et al.*, 1998) as well as CA repeats (Knapik *et al.*, 1998). Primers were used to specify which portion of the DNA was amplified. The RAPD primers used in this study were 10 base pairs long and had arbitrary sequences. A total of 700 RAPD primers were tested, yet only one was found to be linked to the *golden* gene. PCR machines were used to carryout the Polymerase Chain Reactions (PCR) needed to amplify specific portions of the DNA sequence. I used program CM92.2, with the following types of reaction, temperatures and times: initial denaturation 92 degrees, 1 minute; denature 94 degrees, 5 seconds followed by 92 degrees, 55 seconds; anneal 36 degrees, 1 minute; extend 72 degrees, 1 minutes and 45 seconds; repeat steps 2-4 thirty-seven times; final extension 72 degrees, 7 minutes. The amplified DNAs were then separated on a 1% agarose gel using gel electrophoresis; the KNA was stained with ethidium bromide to visualize the bands. The distance between the *golden* gene and its centromere was determined through linkage analysis as described (Streisinger *et al.* 1981).

Mapping TRP2

To map TRP2, primers were designed to identify polymorphisms. These primers bind specifically to a particular region of the TRP2 sequence. The first primer pair used was TRP2.RNK16R and 17F, designed by Robert Kelsh

(personal communication). Primer pair TRP2.F1 and R1 were then designed, followed by primer pairs F2/R2 and F3/R3. All of the possible combinations of the eight primers were tested. I used PCR program ZFPCR, with the following types, temperatures and times: initial denature 94 degrees, 5 minutes; denature 94 degrees, 30 seconds; anneal 50 degrees, 30 seconds; extend 72 degrees, 1 minute; repeat steps 2-4 fourty-five times; final extension 72 degrees, 7 minutes. The amplified DNA was then run on a 2% agarose gel using gel electrophoresis; ethidium bromide stained DNA to visualize the bands. The distance between golden and its centromere was determined through linkage analysis as described (Streisinger *et al.* 1981).

Chapter Three:

Explanation of Methods

To place *golden* on the zebrafish genomic map, I used haploid embryos in haploid-based bulk segregant analysis (BSA). I will use an example (Fig 4) to help explain the concept of haploid-based BSA. Figure 4 shows a hybrid individual with three different genetic markers on the chromosome that contains the *golden* gene. The first marker (g) is the *golden* gene. Two alleles[9] are shown with a plus (+) representing the wildtype (wt) and a minus (-) representing the mutant (mt). The other two markers are anonymous DNA polymorphisims. Marker g and 1 are closely linked (they are near each other on the same chromosome), while g and marker 2 are more distantly linked on the same chromosome (although marker 2 could also be on a different chromosome). The figure illustrates the gene arrangement with the 1- and 2- alleles on the *golden* minus chromosome while the 1+ and 2+ alleles are on the *golden* plus chromosome. These alleles are in coupling, so that all of the minus alleles are on one chromosome and all of the plus alleles are on the other homologous[10] chromosome.

[9] One of any of a group of possible mutational forms of a gene.
[10] Having the same linear sequence of genes as another chromosome.

After meiosis[11], haploid eggs with four different genotypes are produced. If no crossovers[12] have occurred between the *golden* gene and marker 2, then the alleles are in the original parental combinations. However, if a crossover event has occurred between *golden* and marker 2, the mutant embryo will possess the plus allele of marker 2 and the wildtype embryos will possess the minus allele. Recombination between g and 1 is rare, because markers g and 1 are so close together.

As I stated earlier, the wildtype and the mutant embryos are easily distinguished. After collecting DNA from each embryo, I pooled together DNA from twenty wildtype embryos and DNA from twenty mutant embyros. While I only showed two mutants in Fig 4, all (or nearly all) twenty embryos in the mutant pool should have the 1 minus allele. This is so because the *golden* gene and marker 1 are closely linked. On the other hand, *golden* and marker 2 are more distantly linked, so that it is equally likely that an embryo will receive the plus or minus allele. Likewise, in the wildtype pool, all 20 embryos usually have the 1 plus allele while only half have the 2 plus allele because of recombination. Therefore, both will have about equal amounts of the plus and minus alleles of marker 2. In contrast, for marker 1 there will be a difference (polymorphism) in the presence of the 1 plus allele from one pool to the other. The plus allele of

[11] The cell division in sexually reproducing organisms that reduces the number of chromosomes in reproducing cell, leading to the production of eggs or sperm.
[12] The exchange of genetic material between homologous chromosomes.

marker 1 is present in wildtype and absent in mutant pools. Thus, when I found a marker with this pattern, I knew that it was closely linked to *golden*.

My next step was to look for DNA polymorphisms that mimicked the pattern of the locus[13] in my sketch (Fig 5). Two methods can be used to look for the polymorphisms. I chose to use the RAPD technique (Random Amplified Polymorphic DNA). In order to identify DNA polymorphisms, PCR machines are used to carryout the Polymerase Chain Reactions needed to amplify specific portions of a DNA sequence. This then allows us to actually see the DNA fragment. Primers are used to specify which portion of the DNA will be amplified. The RAPD primers are 10 base pairs long and have an arbitrary sequence. Hundreds of these primers are commercially available. The primer (P1) has the sequence AGCGTATAGG. P1 will bind to a complementary base pair sequence of DNA (Fig 6A) and begin elongating in the 3' direction (Fig 6B). Because DNA is double stranded, by chance another primer will also bind to a complementary base pair sequence in the opposite direction on the opposite strand (Fig 6C). This strand will also elongate (Fig 6D). The DNA between the primers will be amplified, making lots of 550 base pair fragments (Fig 6E). Fragments of other sizes will also be produced, because the short sequence to which the RAPD primer binds occurs by chance many times in a genome the size of zebrafish.

[13] The position that a gene occupies on a chromosome.

The chromosome from the *golden* stock is homologous to that of the wildtype stock, but has a slightly different sequence at some point along its length. For example, near to the golden gene in wildtype may be the sequence:

AGCGTATAGG

Yet near to golden on the mutant bearing chromosome the sequence may have an A instead of a G.

AACGTATAGG

If the sequence is not completely complementary, as would happen if a mutation from G to A occurred, then the primer will not bind well. Due to this, DNA amplification will not occur, and a 550 base pair fragment will not be formed. This will show as a difference in amplified fragment contents in wildtype and mutant DNA pools In order to physically observe this polymorphism, it is necessary to run the amplified DNA on an agarose gel using gel electrophoresis.

Chapter Four:

Mapping *golden* Results

In search of a polymorphism, I tested exactly 700 RAPD primer pairs comparing pooled wildtype DNA to that of pooled *golden*. Of the 700 RAPD primers tested, seventy-nine putative polymorphisms were found. These putatives were re-tested with the pooled DNA to confirm the polymorphisms found. After the re-test, only four of the seventy-nine primers were found to show putative polymorphisms (Figure 7). These three primers were then tested on fourteen embryos who are siblings to our Mop mapping cross. Only one of these primers, V14.550, expressed a true polymorphism in this cross.

Primer V14.550 was checked against the approximately 500 markers on our map. I found that a CA repeat marker, Z5231, mapped close to RAPD primer V14.550. This placed V14.550 on linkage group 18. The distance between Z5231 and V14.550 was found to be about 14 cM. This distance was calculated from the 12 recombinant events that occurred out of 82 tested embryos. I then tested the primer V14.550 on a golden cross to find the distance between the gene and the primer. Out of seventy-two embryos tested, only two were recombinant (Fig 8). This indicates that the V14.550 locus is less than 3cM away from *golden*, which places the golden gene on linkage group 18.

However, the exact location still needed to be deduced. I tested Z5231 on a golden cross to find the distance between the gene and the marker (Fig 9). The distance between Z5231 and *golden* was found to be 11 cM. This places the *golden* gene between markers V14.550 and Z5231; V14.550 is furthest from the centromere. The distance between *golden* and its centromere was then determined through linkage analysis. As predicted, *golden* mapped far from the centromere of linkage group 18 (Fig 10).

Mapping *golden* Discussion

Mapping *golden* accomplished many goals. First, this information can contribute to the overall understanding of pigment patterns. This could help us understand the mechanism of pigment pattern control in zebrafish, and possibly humans, giving us a better understanding of how animal embryos develop. Secondly, mapping *golden* to linkage group 18, far from the centromere, confirmed Streisinger's hypothesis. Thirdly, mapping this mutation identifies another gene location on the zebrafish genetic map. Also, mapping this mutation provides us with a visible marker in further genetic studies. Finally, mapping *golden* puts us one step closer to finding out what type of molecule the *golden* gene encodes.

Mapping TRP2 Results

Now that the *golden* mutation has been mapped, I am interested in finding out what molecule is disrupted by the *golden* mutation. We know that the molecule is involved in the production of melanin and there is reason to believe that this molecule may be in a tyrosinase-related protein similar to TRP1 in humans. Currently, the only tyrosinase-related protein sequenced in zebrafish is TRP2, which was sequenced by Robert Kelsh (personal communication, Figure 11). The hypothesis that the TRP2 mutation is the *golden* mutation predicts that they will map to the same genetic location. To test this prediction, I am mapping TRP2. In order to map TRP2, I needed to design the appropriate primers to amplify the gene and then check the amplified fragment for a DNA polymorphism.

The primers I used for this analysis are not RAPD primers. I designed these primers to bind specifically to particular regions of the TRP2 sequence. However, the general amplification mechanism is the same as in that of the RAPD primers.

With my first primer pair TRP2.RNK16R and 17F (designed by Robert Kelsh) I expected to see a fragment of 178 base pairs in size. However, all I saw after running the amplified DNA out on the electrophoresis gel was my primer piled up at the bottom of the gel (Fig 12). The failure to

find a band indicates that DNA was not amplified from the TRP2 gene, because the primers were not working as expected. So I designed another primer pair (TRP2.F1 and R1). Unfortunately, these primers produced the same unpromising results. I then designed primer pairs F2/R2 and F3/R3, but both of these pairs also gave unpromising results. I then tested all of the possible combinations of the eight primers--with no workable results.

I next investigated the structure of the mouse TRP2 gene structure and compared it to that of the zebrafish. I found a very high degree of sequence similarity between the two genes (Fig 13). The mouse TRP-2 sequence had seven introns[14]. Because of the high degree of similarity between the mouse and zebrafish sequence, it seemed reasonable to expect that zebrafish would have introns in identical locations. By comparing the two sequences, I was able to predict the placement of seven introns in the zebrafish TRP2 sequence.

Introns can be thousands of base pairs long, making them nearly impossible to amplify readily. Only a single intron, less than 1000 base pairs long, would amplify in our standard system. The primer pairs that failed to amplify fragments each span a couple of introns, thereby providing a possible explanation of the disappointing results. Putative primers F2 and R3, however, flank the position in which the human sequence predicted intron 6 to be located in the zebrafish

[14] Segments of a gene that are initially transcribed but are then spliced out of the primary RNA transcript, leaving the sequence (exons) on either side of it.

TRP2 gene. If no intron is present in this region, I would expect to see a fragment of 235 base pairs. Yet the reaction amplified a fragment of 400 base pairs. This indicates that a small intron of 165 (400-235) base pairs is likely to be present in the region predicted.

Because of the presence of the introns, my method of mapping TRP2 onto the zebrafish genome changed slightly. Now I am interested in finding a polymorphism in one of the intron sequences in order to map TRP2. I designed new primer pairs to amplify four of the suspected intron regions. Of these four pairs, TRP2.I2.F and R gave a strong band at 700 base pairs. I then used enzyme digests to cut the fragment in the hopes of detecting a polymorphism when the fragment is not so large, but found none (Fig 15). The primer pair TRP2.I2 was also run out on an acrylamide gel[15] by Yi-Lin Yan. A putative polymorphism was detected, but upon further study it was found to be inconclusive.

To supplement my information, I am also crossing heterozygous *gol^b6*/AB siblings and harvesting their embryos. If *golden* embryos are produced in this cross, they will have a deletion of the golden gene. Using TRP2.I2.F/R and other primers that have previously given strong bands, I will look for the presence or absence of a band (instead of a size difference). Finding a polymorphism by this method will allow me to map the TRP2 gene. If I am able to map the TRP2

[15] These dense gels can detect a 1 base pair difference, or even a change from A to C.

gene using my original method, then the deletion mapping will confirm the placement of the TRP2 gene.

Mapping TRP2 Discussion

Although there are not yet results from these experiments, three results are possible. First, TRP2 may map to *golden*. Second, TRP2 may map to *brass*. In either of these cases, I will isolate the DNA from the mutants and sequence it in order to see if there is a change in sequence. If *golden* or *brass* is the mutation that affects TRP2, then I would expect to see a change in the sequence of the gene when comparing wildttype to mutant. However, the TRP2 protein may map to somewhere else in the zebrafish genome entirely. In any of these cases, mapping TRP2 in the zebrafish genome will provide further understanding of the pigment pattern pathway in zebrafish and possibly humans.

REFERENCES

Alberts, B. (1989) Molecular Biology of the Cell, second edition. Garland Publishing, Inc., New York.

Boissy, R. E., Zhao, H., Oetting, W. S., Austin, L. M., Wildenberg, S. C., Boissy, Y. L., Zhao, Y., Sturm, R. A., Hearing, V. J., King, R. A., and Nordlund, J. J. (1996). Mutation in an dlack of expression of tyrosinase-related protein-1 (TRP-1) in melanocytes from an individual with Brown Oculocutaneous Albinism: a new subtype of albinism classified as "OCA3." *Am. J. Hum. Genet.* **58**: 1145-1156.

Chakrabarti, S., Streisinger, G., Singer, F., and Walker, C. (1983) Frequency of gamma-ray-induced specific site and recessive lethal mutations in germ cells of the zebrafish *Brachydanio rerio.* *Genetics* **103**: 109-123.

Fouquet, B., Weinstein, B., Serluca, F. C., and Fishman, M. C. (1997) Vessel Patterning in the Embryo of the Zebrafish: Guidance by Notochord. Developmental Biology **183**:37-48.

Gilbert, S. F. (1997) Developmental Biology, fifth edition. Sinamer Associates Inc, Publishers, Massachussetts.

Halaban, R., Cheng, E., Zhang, Y., Moellmann, G., Hanlon, D., Michalak, M., Setaluri, V., and Hebert, D. (1997) Aberrant retention of tyrosinase in the endoplasmic reticulum mediates accelerated degradation of the enzyme and contributes to the dedifferentiated phenotype of amelanotic melanoma cells. Cell Biology **94**: 6210-6215.

Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M., and Jacob, H. (1998) A microsatellite genetic linkage map for zebrafish (*Danio rerio*). Nature Genentics **18**: 338-343.

Postlethwait, J. H., Johnson, S. L., Midson, C. N., Talbot, W. S., and Gates, M. (1994). A genetic linkage map for the zebrafish. *Science* **264**: 699-703.

Postlethwait, J. H., Yan, Y., Gates, M., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T., Yelick, P., Beier, D., Joly, J., Larhammar, D., Rosa, F., Westerfield, M., Zon, L., Johnson, S., and Talbot, W. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics* **18**: 345-349.

Roessler, E., Bellon, E., Gaudenz, K., Jay, P., Berta, P., Scherer, S. W., Tsui, L., and Muenke, M. (1996) Mutations in the human *Sonic Hedgehog* gene cause holoprosencephaly. Nature Genetics **14**: 357-360.

Streisinger, G., Walker, C., Dower, N., Knauber, D., and Singer, F. (1981) Production of clones of homozygous diploid zebrafish (Brachydanio rerio). *Nature* **291**: 293-296.

Streisinger, G., Walker, C., Dower, N., Knauber, D., and Singer, F. (1986) Segregation analyses and gene-centromere distanced in zebrafish. *Genetics* **112**: 311-319.

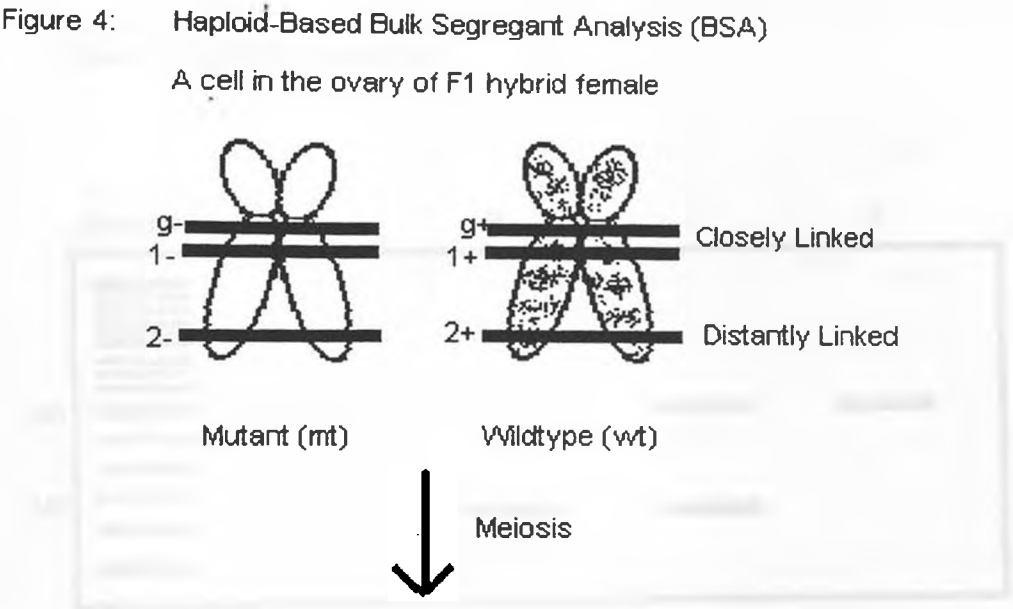
Thompson, M. W., McInnes, R. R., Willard, H. F. (1991) Genetics in Medicine, fifth edition. W.B. Saunders Company, Philadelphia.

Westerfield, M. (1995) The Zebrafish Book. University of Oregon Press.

28 day human embryo

20 hour zebrafish embryo

Figure 4: Haploid-Based Bulk Segregant Analysis (BSA)



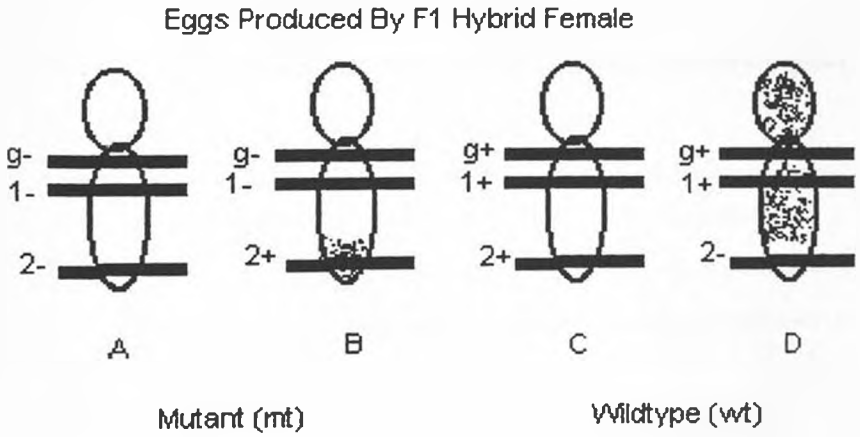


Figure 5: Gel Electrophoresis

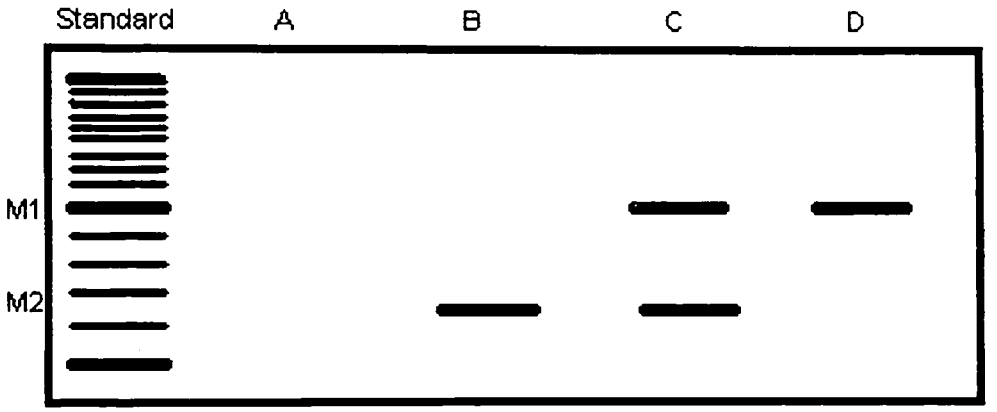
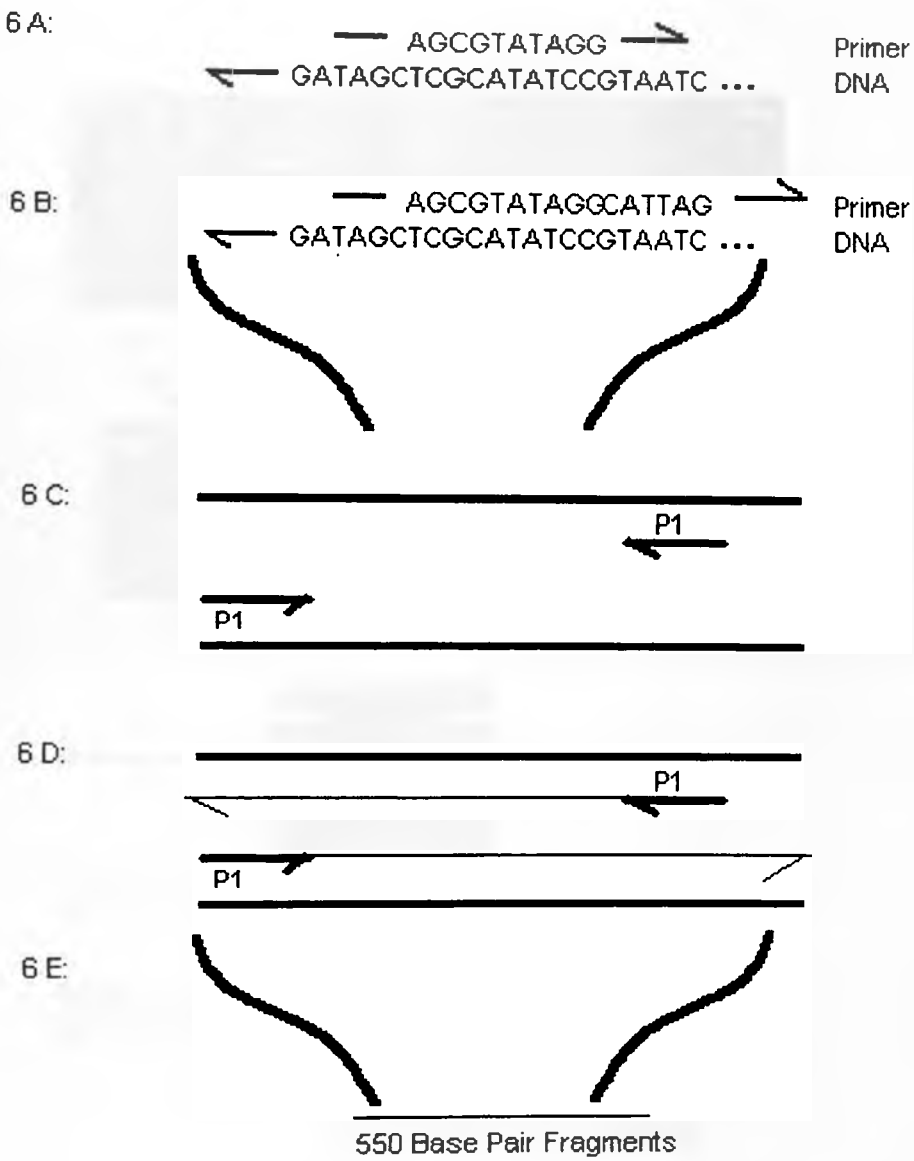
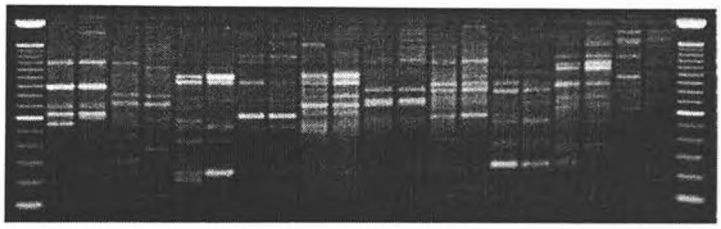


Figure 6: Random Amplified Polymorphic DNA (RAPD) Primers





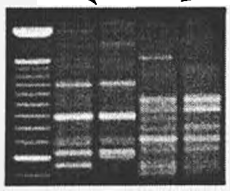


Figure 7a: Primer V14 shows a putative polymorphism between golden and wildtype pools at 550 base pairs. Primer X17 shows a putative polymorphism at 1550 base pairs.

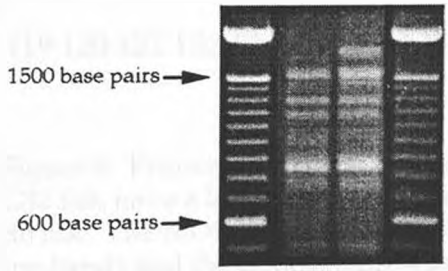
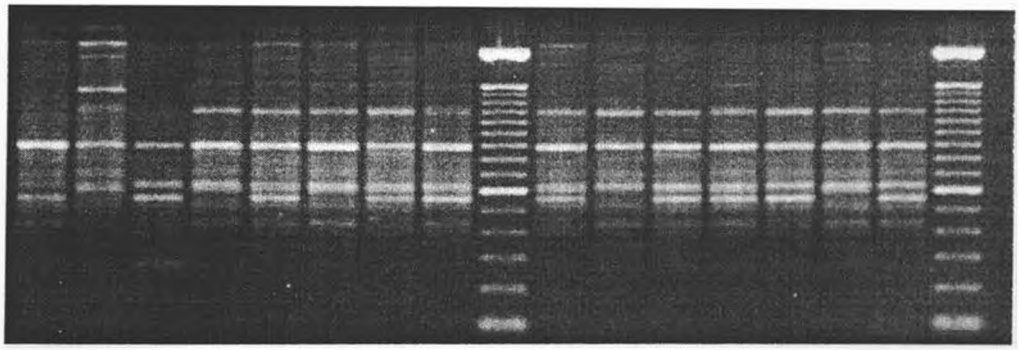


Figure 7b: Primer Z3 shows a putative polymorphism at 1550 base pairs.

Figure 7c: Primer K10 shows a putative polymorphism at 1575 base pairs.

V14.550

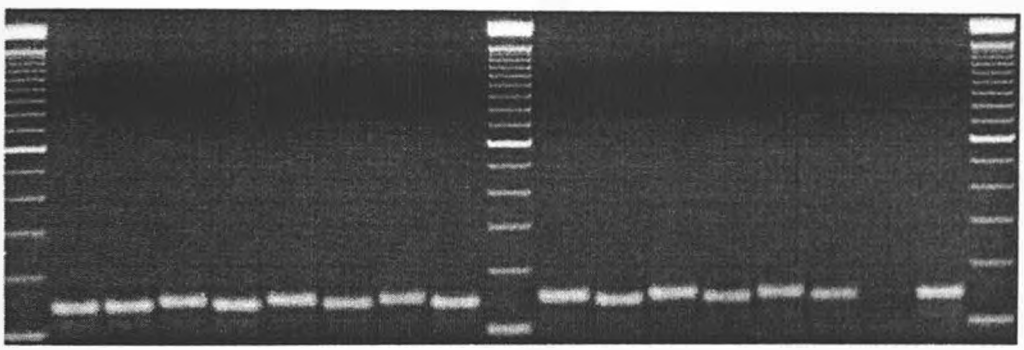


C32 SJD

MOP 119 120 121 122 123 124 125 126 127 128 129 130 131

Figure 8: Primer V14.550 on the MOP mapping cross. C32 fish have a band at 550 base pairs, while SJD fish do not. The MOP DNA was scored for + (band) or - (no band) and the information was entered into our Map Manager program. Primer V14.550 was found to be linked to Z5231, which placed the marker on linkage group 18.

Primer Z5231



gol wt gol wt gol wt gol wt gol wt gol wt gol wt gol wt

Figure 9: Primer Z5231 on golden and wildtype embryos. Usually, the golden embryos will express the "up" band, while the wildtype embryos express the "down" band. When recombinationevents occur, as happened in the first golden embryo, then the band will be the reverse.

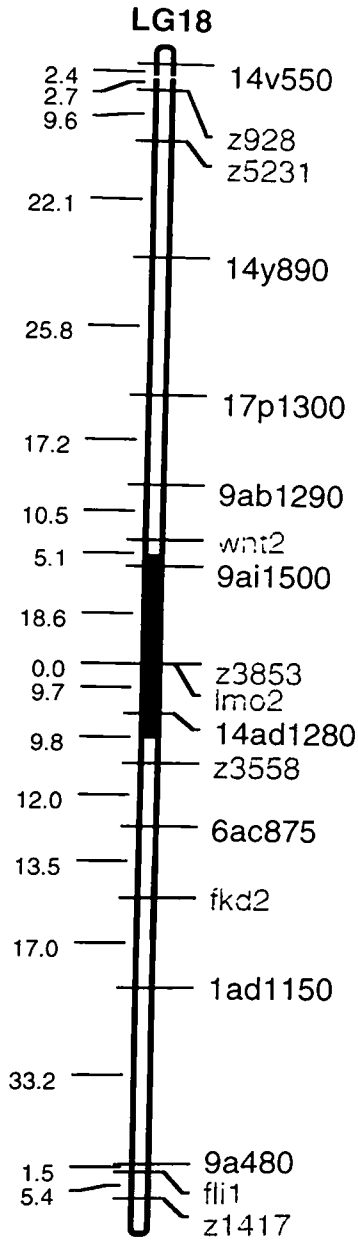


Figure 10 Map of linkage group 18. The *golden* gene is marked in orange, all other genes are marked in blue. Z markers are marked in green and RAPDs are marked in black. The numbers on the left of the chromosome indicate the centiMorgan distance between each marker. As is noted in the text, the *golden* gene lies between Z5231 and V14.550. The *golden* gene is located far from the centromere, as George Streisinger hypothesized.

Primer Forward TRP2, Intron 1

5'GGCACGAGCGCACTTACACACACTGAGAGAGCGAGACTGTGTTTTCGCAGAAATGAAGCGCTTTGTGTTGATGGTGC

GTTTTGGATACATGTACCTGTCGGAGGCTCAGTTTCCCAGAGTCTGCTGTACGGTGGAGGGGATCGTGTCTAAAGAGTGC

TGCCCGGCGCTGGGATCCGATCCGGCAAACGTCTGCGGGCTGCTGGCGGGACgGGGAAACTGCAGTGATGTTCATGTGGA

CAGTAAACCCTGGGGAGGACCGTACAGTCTGAGGAACGTGGACGACAGGGAGAGATGGCCAACCAAGTTCTTCAACCGGA

Intron 1

CCTGCCGCTGCTTTG*GAAACTTTGCCGGCTATAACTGTGGCGAGTGTAAGTTCGG**CTGGACGGGGCTGAACTGTGACCA**

Primer Forward TRP2, Intron 2

CAAGCCGACCTGCCCCGACTTGAC

Primer Reverse TRP2, Intron 1

CGTAAAGCTCCGGTGCTGCGGAAAGACATTCACTCTCTGAGTCCGGAAGAGCTGCAGGAGTTTCTGAACGCACTGGATCT

CAGCAAAACCACCGTTCATCCTGATTATATGATCGCCACACAGCACTGGCTCGGCCTGCTGGGACCCAACGGCACAGAACCA

Intron 2

CAGTTCAGCAACATCTCCATATACGACTTCTTCGTCTGGCAGCATTACTACTCCGTCAGAGACACACTTCTGG*GTCCT

Primer Forward TRP2, Intron 3

GGACGTCCTTTTAAAGCCATCGACTTCTCGCAC**AAAGGCCcAGCGTTCATCACCT**GGCACAGATATCACCTGCTCAGCC

TGTTTTCCGGGTCGCAAGTAGTG

Primer Reverse TRP2, Intron 2

Intron 3

Forward Primer TRP2, Intron 4

GGAGAGAGAACTACAG*AAACTGACAGGTAATGAGAACT**TCGCcGTGCCGTACTGGAACTTT**GCGACAGGCCAGTCA

CTTGAAACGCTGTCCGGTCAGTT

Primer Reverse TRP2, Intron 3

GTGATGTGTGCACAGACTCTCTGCTGGGAGGACGAGACCCCAGTGACCCATCCCGAATCAGCCCGCGTTCACGgtTCGCC

Intron 4

Primer Forward

cGCTGGGGAGTCGtATGTGACAg*TCTGGATGAAtATAATcGTCTGGTGACTCTGTGtAACGGCACAAaTGA**GGGATCAcT**

GTGTTTACTCCCTAGTGA

Primer Reverse TRP2,

TRP2, F3

GCG**CAGAGGAACA**GCGATGAGCACTAATGTGACTTTACCAACCATGAATGATGTGAGAAGCTGCCTGAACCTCAGAGA

CGCGT

Intron 4

Intron 5

TCGACAGTCCACCGTATTTCACCAACTCCACCTTCAGCTTCAG*GAATGCACTGGAGGGATATGATAAGCCGGATGGAGAG

Primer Forward TRP2, F2

CTGGACACCTCGGTGTCCAATCTACACAACCTGGTTCATCAGTTCCTCAATGGGACCAGC**GCTCTGTCTCACTCCGCTG(**

TGGAGCCACAGGTTAGATG

Primer Reverse TRP2, Intron 5

Intron 6

CAATGACCCAATATTTcTG*GTGCTCCACGCATTCACCGACgCCATATTTGATGAATGGATGAGACGCTTCCGTGAAAACG

Primer Forward RNK17

AAAcTTTCCCGGATGAAATGGCTCCAATCGGAaCACAA**CAGGCATTATAATATGGTGCCC**TTCTTCCCACCTGTGAC(

GGGAAGAAGGGTGGACACTGGTT

Primer Reverse TRP2, R3

Intron 7

TGAGGAGATTTACATCGCATCTGAACAGCTCGGATATTCATACTCCATCGACCTCGAAG*AACTGGACAACAGACCGGA

Primer Forward TRP2, CA

TGTTTttGCTgGGCAGCACTCTGGgAGGAATAtTtCTgGGGCTGC**TCCTGCTGTTGaTCGTGTTG**GTGTTGTACATTCAG

CCCCGACGAGGACGACAACT

Primer Reverse RNK16

CAGAGGAGAAAaCGCGAGTTTGAGCCACTCCTGAACGCTGAGTTCACAAACMCAAAATATTCAGGGgAAGCATAACACAC

ACACACACACACTGTCCCAAATCCTATGGAGCTTAAAaTATGCTGAAaCAATCTAAATTTGAACTGTGGTTATTTAAATT

CTTAACATTGAAATGTAATAAAACTGATTATTATATGCCATTAAAAAAAAAAAAAAAAAAAAA3'

TTTTGACTAATAATATACGGT

Primer Reverse TRP2, R1

Figure 11: Tyrosinase-related protein 2 sequence in zebrafish. The sequence is 1744 base pairs long and includes seven introns. Introns are marked with an asterisk (*) and labeled in bold. Some of the forward and reverse primers are located in bold and labeled.

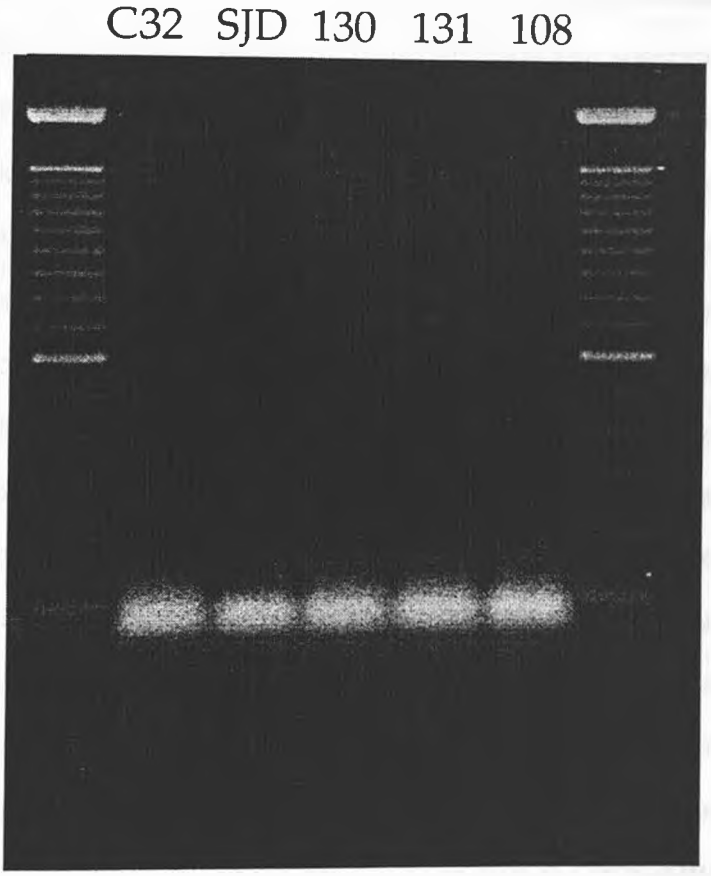


Figure 12: The RNK16, 17 primer does not amplify any DNA sequences.

```
emb|X63349|MMTYRP2 M.musculus tyrp2 mRNA for tyrosinase-related protein-2
          Length = 2182

 Plus Strand HSPs:

 Score = 2437 (673.6 bits), Expect = 3.2e-205, Sum P(2) = 3.2e-205
 Identities = 785/1157 (67%), Positives = 785/1157 (67%), Strand = Plus / Plus

Query:   106 GGCTCAGTTTCCCAGAGTCTGCTGTACGGTGGAGGGGATCGTGTCTAAAGAGTGCTGCCC 165
            |||||||||||||| ||||||||   ||  |||| ||  |  ||   || || ||||||||
Sbjct:   470 GGCTCAGTTTCCCCGAGTCTGCATGACCTTGGATGGCGTGCTGAACAAGGAATGCTGCCC 529

Query:   166 GGCGCTGGGATCCGATCCGGCAAACGTCTGCGGGCTGCTGGCGGGACGGGGAAACTGCAG 225
            | | |||||  ||||  |  | ||| |||| ||  | || | |||  ||||  | |||
Sbjct:   530 GCCTCTGGGTCCCGAGGCAACCAACATCTGTGGATTTCTAGAGGGCAGGGGGCAGTGCGC 589

Query:   226 TGATGTTCATGTGGACAGTAAACCCTGGGGAGGACCGTACAGTCTGAGGAACGTGGACGA 285
             || || ||    ||||  | ||||||| | || || || |  ||  | |||  ||| ||
Sbjct:   590 AGAGGTGCAAACAGACACCAGACCCTGGAGTGGCCCTTATATCCTTCGAAACCAGGATGA 649
                                                                *
Query:   286 CAGGGAGAGATGGCCAACCAAGTTCTTCAACCGGACCTGCCGCTGCTTTGGAAACTTTGC 345
            | | |||  |||||| |  || ||||||||||||| |||   |||   ||||||||||
Sbjct:   650 CCGTGAGCAATGGCCGAGAAAATTCTTCAACCGGACATGCAAATGCACAGGAAACTTTGC 709

Query:   346 CGGCTATAACTGTGGCGAGTGTAAGTTCGGCTGGACGGGGCTGAACTGTGACCAGCGTAA 405
             || ||||| ||||| |  || |||||||||||||| || |   ||||| | | |   ||
Sbjct:   710 TGGTTATAATTGTGGAGGCTGCAAGTTCGGCTGGACCGGCCCCGACTGTAATCGGAAGAA 769

Query:   406 AGCTCCGGTGCTGCGGAAAGACATTCACTCTCTGAGTCCGGAAGAGCTGCAGGAGTTTCT 465
              |  |  | ||  |      | || || || |||| | |  | |||  | || ||||  |
Sbjct:   770 GCCGGCCATCCTAAGACGGAATATCCATTCCCTGACTGCCCAGGAGAGGGAGCAGTTCTT 829

Query:   466 GAACGCACTGGATCTCAGCAAAACCACCGTTCATCCTGATTATATGATCGCCACACAGCA 525
            |  |||  | || ||   ||| |  |   | |||||| || ||  ||||| |||| || ||
Sbjct:   830 GGGCGCCTTAGACCTGGCCAAGAAGAGTATCCATCCAGACTACGTGATCACCACGCAACA 889

Query:   526 CTGGCTCGGCCTGCTGGGACCCAACGGCACAGAACCACAGTTCAGCAACATCTCCATATA 585
            |||||| || ||||| |||||||||||| ||  | || ||| ||  ||||  |  | | ||
Sbjct:   890 CTGGCTGGGGCTGCTCGGACCCAACGGGACCCAGCCCCAGATCGCCAACTGCAGCGTGTA 949
                                                              *
Query:   586 CGACTTCTTCGTCTGGCAGCATTACTACTCCGTCAGAGACACACTTCTGGGTCCTGGACG 645
             ||||| || || ||||  ||||| || || ||  |||||||| |  | |||||  |||||
Sbjct:   950 TGACTTTTTTGTGTGGCTCCATTATTATTCTGTTCGAGACACATTATTAGGTCCAGGACG 1009

Query:   646 TCCTTTTAAAGCCATCGACTTCTCGCACAAAGGCCCAGCGTTCATCACCTGGCACAGATA 705
             || | |||  ||||| || ||||| ||| |||| || || ||  |||| |||||||| ||
Sbjct:  1010 CCCCTATAAGGCCATTGATTTCTCTCACCAAGGGCCTGCCTTTGTCACGTGGCACAGGTA 1069
                                                  *
Query:   706 TCACCTGCTCAGCCTGGAGAGAGAACTACAGAAACTGACAGGTAATGAGAACTTCGCCGT 765
             || ||| |  | ||||| |||||||| |||| ||| || || ||||||  ||| ||  |
Sbjct:  1070 CCATCTGTTGTGGCTGGAAAGAGAACTCCAGAGACTCACTGGCAATGAGTCCTTTGCGTT 1129

Query:   766 GCCGTACTGGAACTTTGCGACAGGCCAGTCAGAGTGTGATGTGTGCACAGACTCTCTGCT 825
            ||| |||||||||||||| || ||  ||   |||||||| ||||||||||||     |||
Sbjct:  1130 GCCCTACTGGAACTTTGCAACCGGGAAGAACGAGTGTGACGTGTGCACAGACGACTGGCT 1189

Query:   826 GGGAGGACGAGACCCCAGTGACCCATCCCGAATCAGCCCGCGTTCACGGTTCGCCCGCTG 885
             ||||  |  |   |    |||||||  |  |  || || | |   ||  | ||| |   |||
```

Figure 13: Structure similarity between mouse (query) and zebrafish (subject) TRP2 genes. The intron/exon boundry is marked with an asterisk (*).Each matched base is designated by a verticle line.