

# A Genetic, Deletion, Physical, and Human Homology Map of the *long fin* Region on Zebrafish Linkage Group 2

M. Kathryn Iovine\* and Stephen L. Johnson

\*To whom correspondence and reprint requests should be addressed. Fax: (314) 362-7855. E-mail: [kiovine@genetics.wustl.edu](mailto:kiovine@genetics.wustl.edu).

**Mutation of the gene *long fin* causes overgrowth of zebrafish fins. Thus, molecular identification of the gene *long fin* may reveal the mechanisms involved in normal growth control. We have therefore developed genetic and physical maps of the corresponding region on linkage group 2 (LG2). A single deletion allele (*lof<sup>g61</sup>*) of the *long fin* gene was also generated. Examination of this deletion for the presence or absence of ESTs independently mapped to LG2 revealed a contiguous deletion of SSLP, STS, and gene-based markers spanning a physical distance of approximately 500 kb, including at least 10 zebrafish genes. Human orthologs of the zebrafish genes in the *long fin* region were identified and revealed two synteny segments from human chromosome 1 (Hsa1) and Hsa19. Homology searching for additional genes linked to the human orthologs revealed one additional gene in the *long fin* deletion region. Thus, our development of the genetic, physical, deletion, and human homology maps of the *long fin* region provides one of the first high-resolution comparisons of a zebrafish region with a homologous human region, and facilitates the molecular identification of the *long fin* gene.**

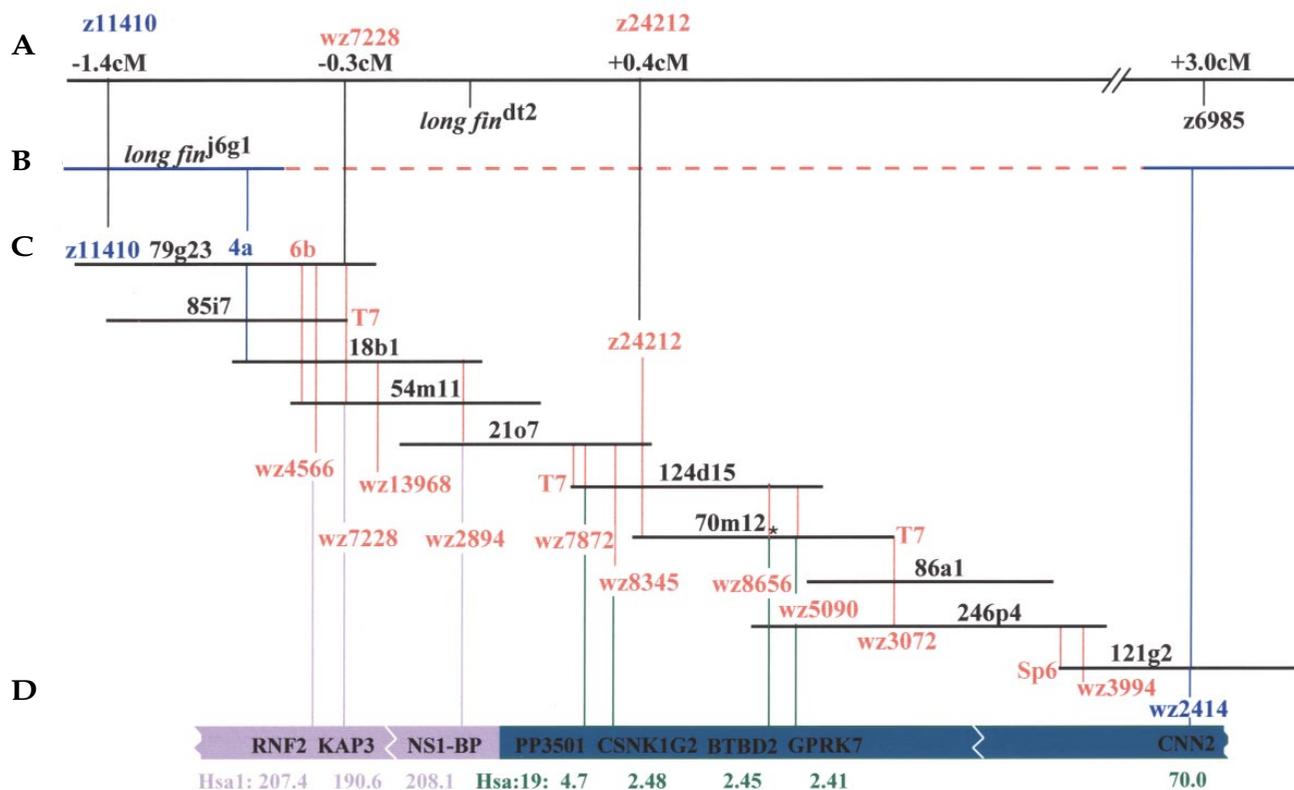
Zebrafish fins grow isometrically with body growth throughout the life of the animal. As the rate of increase in body length slows in older fish, the rate of increase in fin length also slows. The commensurate decrease in the rate of fin growth may be the result of growth control mechanisms that act to prevent addition of new fin ray segments to the end of fins at inappropriate times. The dominant *long fin* mutation causes too many segments to be added to the fin ray, and thus may reveal the mechanisms responsible for controlling fin ray segment initiation [1]. To aid in the molecular identification of the *long fin* gene, we have developed genetic and physical maps of the corresponding region.

We first mapped *long fin* to linkage group 2 (LG2) by screening for linked random amplified polymorphic DNAs (RAPD) markers [2,3]. This revealed close linkage of *long fin* with RAPD marker A8.1300 (data not shown). Development of the MGH simple sequence length polymorphism (SSLP) map [4] and integration of that map with the RAPD map (data not shown) then allowed us to map *long fin* on a panel of 984 meioses to a 1.8 cM region between SSLP markers z11410 and z24212 (Fig. 1A).

To aid in the development of the map, we generated a deletion of the *long fin* region. Because the *long fin* mutation is dominant, we reasoned that deletions of the mutant locus might revert the phenotype. Accordingly, we exposed sperm from *long fin* homozygotes to  $\gamma$ -irradiation (250 rads) to cause deletions [5]. We then used  $\gamma$ -irradiated sperm to fertilize eggs from the C32 inbred genetic background [6], and reared 1177 individuals to maturity. Most fish developed with long fins (typical of *long fin* heterozygotes), but one fish developed wild-type fins, suggesting that it now carried a reversion (*lof<sup>g81</sup>*) of the locus. Analysis with flanking markers z11410 and z6985 showed that this individual indeed carried the mutagenized *long fin* chromosome, but had lost the *long fin* haplotype for the closely linked marker z24212 (data not shown). This was confirmed by showing that haploid embryos that have the mutagenized *long fin* chromosome fail to amplify the z24212 marker (data not shown).

To enhance the *long fin* map and identify candidate genes for the *long fin* mutation, we screened ESTs and genes mapped on the zebrafish transcript maps [7] (S.L.J. *et al.*, manuscript submitted) in an  $\sim 7$  cM region (see [http://www.genetics.wustl.edu/fish\\_lab/cgi-bin/published/human\\_int\\_map.cgi](http://www.genetics.wustl.edu/fish_lab/cgi-bin/published/human_int_map.cgi)) between flanking markers z11066 and z6985 for the ability to amplify from *lof<sup>g81</sup>* DNA. Of 80 unique genes tested, we found 9 that amplified from embryos with unmutagenized chromosomes, but failed to amplify from embryos homozygous or hemizygous for the deletion chromosome (Table 1). We next used the STS markers for genes absent from deletion chromosomes to identify zebrafish PAC genomic clones [8] corresponding to this region. PCR screening of PAC pools (S.L.J., unpublished data) revealed 11 PACs. STS content mapping of these PACs for genes and SSLPs in the region allowed us to build a contiguous physical map of the region (Fig. 1C). The average depth of the contig map was 2.7 clones per STS marker in the deletion. Although we have not restriction-digest mapped [9] this contig, examination of the contig map suggests the deletion is on the order of 4–5 PAC clones in length. Assuming an average clone insert size of  $\sim 115$  kb [8], this suggests that the deletion is approximately 460–575 kb.

When we rescreened the larger set of 80 genes in the *long fin* region for presence in the clone contig map, we identified a single additional gene (*wz2414*) that amplified from the rightward-most clone of the contig. The STS for this gene was amplified from deletion animals, indicating that the contig spans the rightward breakpoint of the deletion. SSLP marker



**FIG. 1.** Genetic, deletion, physical, and human homology map of the *long fin* region. (A) Genetic map of the *long fin* region places the dominant *long fin* mutation between markers *wz7228* and *z24212*. The oligonucleotide for *wz7228* that amplifies a single nucleotide polymorphism is 5'-GGCTACGGCTCTACC/CAGATC-TAAGACTACAGC-3'. At approximately nucleotide 190 in the SJD haplotype, we found cagcaaaa; in the *long fin* and C32 haplotypes, we found cacaaaa. (B) Deletion map of the  $\gamma$ -ray-induced revertant allele, *long fin*<sup>j6g1</sup>. The dotted line indicates the span of the deletion. (C) Contig map of genomic PACs in the *long fin* region. Markers represented in red are absent from the deletion chromosome; markers represented in blue are present on the deletion chromosome. The lines representing the individual PAC clones are not meant to indicate the size of the PACs. (D) Human orthology map of the genes found in the region. Human map positions were found from the UCSC assembly (April 2001 freeze) of the human genome [11]. Because the genes *CSNK1G2*, *BTBD2*, and *GPRK7* are missing from the Golden Path assembly, we took the NCBI map viewer positions for these genes and for *PP3501* in this homology segment. Asterisks indicate negative amplification from a PAC using that STS marker.

*z11410* and STS *4a* (developed from sample sequencing of PACs; data not shown) both amplify from the leftward-most PAC of the contig and from deletion animals, indicating that the contig also spans the leftward breakpoint of the deletion. Together, this indicates that the entire extent of the deletion is likely to be covered by our contig map.

We sought to use conserved synteny between zebrafish and human to identify additional zebrafish genes in the deletion region. We first identified human orthologs for genes in this region from the annotation of the WU zebrafish EST project ([http://www.genetics.wustl.edu/fish\\_lab/cgi-bin/search.cgi](http://www.genetics.wustl.edu/fish_lab/cgi-bin/search.cgi)). For 3 (*wz7228*, *wz2894*, and *wz7872*) of the 10 genes mapped in this region, human orthologs (*KIFAP3*, *NS1-BP*, and *PP3501*, respectively) had been predicted by stringent automated criteria (S.L.J. *et al.*, manuscript submitted). These genes belong to the LG2:Hsa1 or LG2:Hsa19 conserved synteny groups. We predict human orthologs (*RNF2*, *CSNK1G2*, *GPRK7*, and *CNN2*) for an additional four of the genes in the deletion region (*wz4566*, *wz8345*, *wz5090*, and

*wz2414*, respectively) based on their best human hit (with *e*-value < -20). In each of these cases, the best human hit belongs to one of the conserved synteny groups with members in the deleted region. This supports our conclusion that we have identified the correct orthologs for these zebrafish genes. A search of the corresponding human regions revealed a previously unmapped zebrafish gene (*wz8656*) orthologous to *BTBD2* on human chromosome 19 (Hsa19) between *CSNK1G2* and *GPRK7*, suggesting this gene might also belong to this zebrafish region. STS markers for *wz8656* failed to amplify from *loj*<sup>β6g1</sup> deletion animals, but amplified from the PACs *124d15* and *246p4*, confirming the utility of conserved synteny to identify candidate genes and enhance our maps.

The availability of the integrated zebrafish transcript map, with 6817 mapped genes and 1957 SSLP markers (S.L.J. *et al.*, manuscript submitted), provides a powerful starting point for developing physical maps for identification of mutated zebrafish genes. As this map provides an average marker density of 1 marker every 170 kb, it will be possible to

TABLE 1: Zebrafish genes surveyed in the *lof* region

WZ	Position	Amplify from <i>lof</i> <sup>681</sup>	WZ	Position	Amplify from <i>lof</i> <sup>681</sup>	WZ	Position	Amplify from <i>lof</i> <sup>681</sup>
<i>wz382</i>	2:41(1)	+	<i>wz8568</i>	2:41(28)	+	<i>wz4295</i>	2:44(14)	+
<i>wz7013</i>	2:41(2)	+	<i>wz2362</i>	2:41(29)	+	<i>wz4979</i>	2:44(15)	+
<i>w9884</i>	2:41(3)	+	<i>wz13355</i>	2:41(30)	+	<i>wz5573</i>	2:44(16)	+
<i>wz1581</i>	2:41(4)	+	<i>wz2863</i>	2:41(31)	+	<i>wz11277</i>	2:44(17)	+
<i>wz892</i>	2:41(5)	+	<i>wz8957</i>	2:41(32)	+	<i>wz7004</i>	2:44(18)	+
<i>wz7872</i>	2:41(6)	-	<i>wz4332</i>	2:43(1)	+	<i>wz12520</i>	2:44(19)	+
<i>wz13681</i>	2:41(7)	+	<i>wz7651</i>	2:43(2)	+	<i>wz6826</i>	2:44(20)	+
<i>wz2414</i>	2:41(8)	+	<i>wz883</i>	2:43(3)	+	<i>wz8524</i>	2:44(21)	+
<i>wz1190</i>	2:41(9)	+	<i>wz6809</i>	2:43(4)	+	<i>wz857</i>	2:44(22)	+
<i>wz4735</i>	2:41(10)	+	<i>wz252</i>	2:43(5)	+	<i>wz465</i>	2:44(23)	+
<i>wz5677</i>	2:41(11)	+	<i>wz4360</i>	2:43(6)	+	<i>wz7228</i>	2:44(24)	-
<i>wz3994</i>	2:41(12)	-	<i>wz 4895</i>	2:43(7)	+	<i>wz1567</i>	2:44(25)	+
<i>wz5090</i>	2:41(13)	-	<i>wz8163</i>	2:43(8)	+	<i>wz2881</i>	2:44(26)	+
<i>wz8345</i>	2:41(14)	-	<i>wz4773</i>	2:43(9)	+	<i>wz6207</i>	2:44(27)	+
<i>wz4566</i>	2:41(15)	-	<i>wz7929</i>	2:44(1)	+	<i>wz10450</i>	2:44(28)	+
<i>wz3295</i>	2:41(16)	+	<i>wz278</i>	2:44(2)	+	<i>wz7813</i>	2:44(29)	+
<i>wz3072</i>	2:41(17)	+	<i>wz3868</i>	2:44(3)	+	<i>wz8840</i>	2:46(1)	+
<i>wz4591</i>	2:41(18)	+	<i>wz2862</i>	2:44(4)	+	<i>wz2342</i>	2:46(2)	+
<i>wz9792</i>	2:41(19)	+	<i>wz9110</i>	2:44(5)	+	<i>wz8682</i>	2:46(3)	+
<i>wz7368</i>	2:41(20)	+	<i>wz4547</i>	2:44(6)	+	<i>wz5848</i>	2:46(4)	+
<i>wz13918</i>	2:41(21)	+	<i>wz7628</i>	2:44(7)	+	<i>wz6938</i>	2:46(5)	+
<i>wz2894</i>	2:41(22)	-	<i>wz2395</i>	2:44(8)	+	<i>wz13392</i>	2:46(6)	+
<i>wz13968</i>	2:41(23)	-	<i>wz3690</i>	2:44(9)	+	<i>wz437</i>	2:46(7)	+
<i>wz2786</i>	2:41(24)	+	<i>wz5834</i>	2:44(10)	+	<i>wz7318</i>	2:46(8)	+
<i>wz5313</i>	2:41(25)	+	<i>wz7451</i>	2:44(11)	+	<i>wz4897</i>	2:46(9)	+
<i>wz1377</i>	2:41(26)	+	<i>wz386</i>	2:44(12)	+	<i>wz10420</i>	2:46(10)	+
<i>wz8865</i>	2:41(27)	+	<i>wz4061</i>	2:44(13)	+	<i>wz512</i>	2:46(11)	+

Shown are relative marker order for genes on the integrated map (S.L.J. *et al.*, manuscript submitted). Position in cM is based on standard marker positions [10]. All oligo sequences may be found at [http://www.genetics.wustl.edu/fish\\_lab/cgi-bin/search](http://www.genetics.wustl.edu/fish_lab/cgi-bin/search).

generate physical contig maps for selected regions without laborious walking. Here, we have tested this idea by using markers shown to be absent from a deletion as the starting point for generating the clone contig map. From nine mapped genes and two SSLPs in the region used as probes, we were able to identify 11 PACs that completely spanned the region. We identified these 9 mapped genes from a relatively large population of 80 genes predicted to be close to our deletion. We chose to survey such a large number of markers because the radiation hybrid panels [7] (S.L.J. *et al.*, manuscript submitted) and low-resolution meiotic mapping panels [10] used in the integration may locally disorder individual markers. Further, the integration of the maps relied on markers shared between panels which are separated on average about every

1 million bp (which is at a lower resolution than the size of our deletion). We found that the nine genes shown to be absent from the *lof*<sup>681</sup> deletion are interspersed in the integrated transcript map with other genes that we have shown are present on the deletion chromosome (Table 1). As the contiguity of genes in the deletion was confirmed by the clone contig map, we conclude that although the radiation hybrid transcript maps [7] (S.L.J. *et al.*, manuscript submitted) are successful at identifying genes in local regions, the relative marker order may not be reliable.

We also found that conserved synteny between zebrafish and human genomes, coupled with the nearly completed human sequence assembly [11], can be used to identify additional zebrafish genes in the region that have not yet been

mapped. Thus we identified *wz8656* as a candidate for the deletion region because it is an ortholog of *BTBD2*, a gene that falls between *CSNK1G2* and *GPRK7* on the human map. It is also worth noting that homology segments (that is, regions of common gene order between zebrafish and human chromosomes) shared between zebrafish and human may be relatively small. Whole genome comparison between zebrafish and human suggests that the average size of homology segments is on the order of 600 kb (S.L.J. *et al.*, manuscript submitted). In this study, we suggest that the LG2:Hsa1 conserved synteny group in the deletion region is composed of two adjacent homology segments. The order of markers *wz4566* and *wz7228* is likely to be correct as shown (Fig. 1C) because the gene *wz7228* was also found in T7-end sequence from PAC 85i7, thus placing it between *wz4566* and *wz13968*. However, on the human map, the gene *RNF2* (orthologous to *wz4566*) falls between genes *KIFAP3* (orthologous to *wz7228*) and *NS1-BP* (orthologous to *wz2894*), indicating a rearrangement in either the zebrafish or human lineage. Alternatively, these genes may have been mis-ordered on the human genome sequence assembly [11]. By similar logic, we suggest that the LG2:Hsa19 conserved synteny group is also composed of two adjacent homology segments (Fig. 1D). As homology segments between human and zebrafish genomes analyzed here show several local rearrangements, the search for additional zebrafish genes using conserved synteny should be carried out with some caution.

We have developed a genetic, deletion, physical, and human homology map for the zebrafish *long fin* region. The combination of the *lopf<sup>6g1</sup>* deletion with gene identification from the zebrafish EST sequencing project has provided valuable tools for this analysis of the *long fin* genomic region. Thus, we have described the first steps in the molecular

identification of the *long fin* gene. Furthermore, we found that it was possible to use human homology with zebrafish genes to identify an additional zebrafish gene in the *long fin* region.

#### ACKNOWLEDGMENTS

We thank Chad Clayton and Colleen Boggs for care of the zebrafish facility. We also thank the Washington University Genome Resources project (NIH grant RO1DK55379) and the Zon lab (NIH grant RO1DK55381) for access to RH mapping data before publication. This work was supported in part by the NIH GM56988, a HHMI-Washington University Pilot Project grant, and a Pew Scholars award to S.L.J. M.K.I. was supported by a post-doctoral fellowship from the National Institute of Child Health and Development (NICHD) HD08536-01.

RECEIVED FOR PUBLICATION OCTOBER 3, 2001;  
ACCEPTED APRIL 3, 2002.

#### REFERENCES

1. Iovine, M. K., and Johnson, S. L. (2000). Genetic analysis of isometric growth control mechanisms in the zebrafish caudal fin. *Genetics* **155**: 1321-1329.
2. Johnson S. L., *et al.* (1996). Centromere-linkage analysis and consolidation of the zebrafish genetic map. *Genetics* **142**: 1277-1288.
3. Johnson, S. L., Midson, C. N., Ballinger, E. W., and Postlethwait, J. H. (1994). Identification of RAPD primers that reveal extensive polymorphisms between laboratory strains of zebrafish. *Genomics* **19**: 152-156.
4. Knapik, E. W., *et al.* (1998). A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nat. Genet.* **18**: 338-343.
5. Fritz, A., Rozowski, M., Walker, C., and Westerfield, M. (1996). Identification of selected  $\gamma$ -ray induced deficiencies in zebrafish using multiplex polymerase chain reaction. *Genetics* **144**:1735-1745.
6. Nechiporuk, A., Finney, J. E., Keating, M. T., and Johnson, S. L. (1999). Assessment of polymorphism in zebrafish mapping strains. *Genome Res.* **9**: 1231-1238.
7. Hukriede, N., *et al.* (2001). The LN54 radiation hybrid map of zebrafish expressed sequences. *Genome Res.* **11**: 2127-2132.
8. Amemiya, C. T., Ahong, T. P., Silverman, G. A., Fishman, M. C., and Zon, L. I. (1999). Zebrafish YAC, BAC, and PAC genomic libraries. *Methods Cell Biol.* **60**: 235-258.
9. Marra, M. A., *et al.* (1997). High throughput fingerprint analysis of large-insert clones. *Genome Res.* **7**: 1072-1084.
10. Woods, I. G., *et al.* (2000). A comparative map of the zebrafish genome. *Genome Res.* **10**: 1903-1914.
11. Lander, E., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.