

Supporting Information

Mo et al. 10.1073/pnas.1710254114

SI Materials and Methods

Generation of UAS:Gal4 Fish. A sequence encoding for the Gal4-activator (KalTA4) was inserted within the 5' UTR of *ltk*, *aug-α1*, *aug-α2*, and *aug-β* genes. To do so, embryos carrying *sox10:mCherry* and *UAS:eGFP* transgenes were injected at the single-cell stage with a mixture containing (i) sgRNAs targeting specific 5' UTR sequences near the endogenous gene's start codon; and (ii) Cas9 mRNA; (iii) a donor plasmid with the equivalent 5' UTR sequence followed by a viral 2A peptide and KalTA4 (32). The donor plasmid, which is based upon Auer et al. (33), was specifically inserted into the 5' UTR of the *ltk* and *augmentor* genes using two sgRNAs specific to that region as previously described. This donor plasmid includes a coding sequence of the transcriptional transactivator Gal4 (KalTA4) preceded by an E2A peptide linker for multicistronic expression. Upon injection of the above mixture into a *UAS:eGFP* transgenic fish, Cas9 cleaves the genomic locus as well as the donor plasmid sequence as specified by the sgRNAs. DNA repair triggers integration of the donor plasmid into the genomic locus through nonspecific ligation of cleaved DNA. Since the indel sequence preceding the Gal4-cDNA cannot be predicted upon integration, the presence of the E2A will allow the in-frame insertions of the Gal4-cDNA (33).

Determination of On-Targets and Off-Targets. Genomic DNA from zebrafish embryos at 24 hpf and from the tails of adult fish was extracted by boiling at 95 °C for 20 min in 50 μL of 100 mM sodium hydroxide and then neutralized with 10 μL of 1 M Tris-HCl (pH 7.5). Then 1 μL of genomic DNA was used as template to amplify an ~200-bp sequence surrounding the on-target and off-target regions identified by CRISPR RGEN Tools (www.rgenome.net) (Fig. S24). Mutations were detected by T7EI assay (31). Mutation frequencies were determined by quantifications of T7EI-undigested and -digested PCR products. For each quantification, the intensity of the PCR bands of the undigested product and of a T7EI-digested fragment of the same size were analyzed by ImageJ (<https://imagej.nih.gov/ij/>). The

mutation frequencies were calculated by the ratio of T7EI-digested to T7EI-undigested PCR products.

Sox10⁺ Cell Isolation. Sox10⁺ cells were isolated from Sox10:mCherry by using FACS. Sox10:mCherry transgenic zebrafish embryos are embryos expressing mCherry as driven by the *sox10* promoter. mCherry-positive and -negative cells were separated as described previously (41). Briefly, 16-hpf and 24-hpf *sox10:mCherry* embryos and age-matched wild-type embryos were collected and dechorionated, and the yolk was mechanically removed and disassociated with pronase followed by further digestion to single cells with Librase (Roche) at 28 °C for 10 min then trypsin (Gibco) for 5 min. The reaction was terminated by the addition of FBS. Cells were pelleted, washed with suspension buffer containing Leibovitz medium, 0.8 mM calcium chloride, penicillin streptomycin, and 0.5% FBS (all from Gibco), and filtered with a 35-μm cell strainer into 5-mL FACS tubes. FACS was performed by a five-laser BD FACSAria II cell sorter equipped with BD FACSDiva software v8.0.1 (BD Biosciences). DAPI was added to the sample before FACS to collect only viable cells. Viable mCherry-positive and -negative cells were sorted using a 70-μm nozzle at 60 psi. Sorted cells were collected, and total RNA was extracted using TRIzol reagent (Life Technologies).

Whole-Mount in Situ Hybridization. Zebrafish embryos at different stages were fixed, and whole-mount in situ hybridization was performed as previously described (42). To make riboprobes, fragments for riboprobe synthesis were amplified from wild-type full-length zebrafish cDNA and cloned into pBluescript II KS (-) (Agilent Technologies). To generate digoxigenin (DIG)-labeled antisense riboprobes, linearized constructs were reverse transcribed by T3 RNA polymerase with a DIG Labeling Kit (Roche). Meanwhile, sense-sequence riboprobes were generated using T7 RNA polymerase with the DIG Labeling Kit. The DIG signal was revealed by the NBT/BCIP system (Roche), and images were acquired with Leica Application Suite (Leica) using the Leica MacroFluo system and Leica LED5000RL light (Leica).

A

Full-Length Augmentors

zebrafish_Aug_a1	1	-----	12
zebrafish_Aug_a2	1	-----	59
zebrafish_Aug_b	1	-----	12
zebrafish_Aug_a1	13	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	65
zebrafish_Aug_a2	60	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	114
zebrafish_Aug_b	13	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	72
zebrafish_Aug_a1	66	-HHLKTAGSTLLEFPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	123
zebrafish_Aug_a2	117	-HHLKTAGSTLLEFPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	174
zebrafish_Aug_b	73	-HHLKTAGSTLLEFPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	132
zebrafish_Aug_a1	124	YKRCARLLRLAGSQRTE--	143
zebrafish_Aug_a2	176	YKRCARLLRLAGSQRTE--	195
zebrafish_Aug_b	133	YKRCARLLRLAGSQRTE--	152

Full-Length Aug-a

human_AUG_a	1	-----	11
zebrafish_AUG_a1	1	-----	12
zebrafish_AUG_a2	1	-----	59
human_AUG_a	12	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	69
zebrafish_AUG_a1	13	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	69
zebrafish_AUG_a2	60	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	114
human_AUG_a	70	RKAAAGLSPSPORVETVPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	129
zebrafish_AUG_a1	64	DV-HHLKTAGSTLLEFPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	122
zebrafish_AUG_a2	115	DV-HHLKTAGSTLLEFPRDLRREKRETTLLDGLVYGPCKKDDVYLYNTRDCTIPAH	174
human_AUG_a	130	YKRCARLLRLAGSQRTE--	152
zebrafish_AUG_a1	123	YKRCARLLRLAGSQRTE--	143
zebrafish_AUG_a2	175	YKRCARLLRLAGSQRTE--	195

Percent Identity Matrix - created by Clustal2.1

1: human_AUG_a	100.00	44.29	48.97
2: zebrafish_AUG_a1	44.29	100.00	50.70
3: zebrafish_AUG_a2	48.97	50.70	100.00

Aug-a Aug Domain

hAug_a_AugDomain	1	-----	60
zAug_a1_AugDomain	1	-----	60
zAug_a2_AugDomain	1	-----	60
hAug_a_AugDomain	61	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	72
zAug_a1_AugDomain	61	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	72
zAug_a2_AugDomain	61	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	70

Percent Identity Matrix - created by Clustal2.1

1: hAug_a_AugDomain	100.00	68.57	72.86
2: hAug_a_AugDomain	68.57	100.00	78.57
3: hAug_a2_AugDomain	72.86	78.57	100.00

Full-Length Aug-β

human_Aug_b	1	-----	35
zebrafish_Aug_b	1	-----	35
human_Aug_b	36	KVLDKPKPFLFLPAAGAKTSCSSAETPRSSNLRKRFKHEHGEVYSPSSSHH	94
zebrafish_Aug_b	59	YSVQSDAKSSKENTAPLPEKSRPTEIPRDTSKKRFLEHFAAGVYSPSSSHH	117
human_Aug_b	95	HRLYNTRDCTIPAHYKRCARLLRLAGSQRTE--	129
zebrafish_Aug_b	118	HRLYNTRDCTIPAHYKRCARLLRLAGSQRTE--	152

percent identical: 41.29%

Aug-β Aug Domain

hAug_b_AugDomain	1	-----	59
zAug_b_AugDomain	1	-----	60
hAug_b_AugDomain	60	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	72
zAug_b_AugDomain	61	LVLLCTAAQSDASAKVETLREHREHRENSADDESA-----QRTSAPFPHPT	73

percent identical: 69.44%

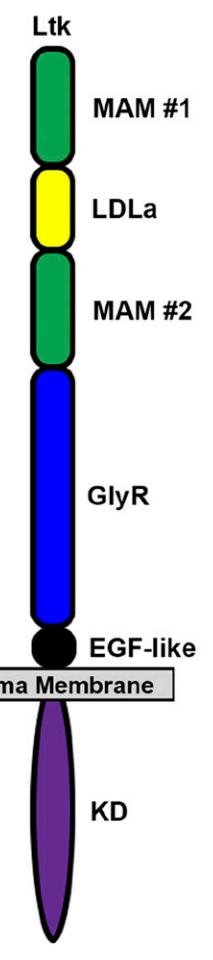
B

Zebrafish Ltk Primary Sequence:

```

MDYITRQTFVVKLALFIFTVVRSSCALLEKAAESVPHNPLQSSPAEDSDVSPCDFESPSCS
WTLSSHSTGGGEWFIITSAQHRNRRDTPQIRDYSTGKSEGHFLLKPSSSHLSAARCSFH
MTSPVVLSSGPFCHLQARFQPEPHAGNISAFVKHTDSIDIKPIDLTIKEQESDSSQWEV
LEAVIQGLNEPFQVTVQYSAAGSHVGFALFDSLELRKNCVMGDDYVDLGSDEKYSLSLQC
HSGGCIKQRVCFHDTDCPEGEDEGLICSTLPLGSCYCFELGSCGWSAADTQSSWRLVSG
QQLIEDTHLLGTTLKNTQGHFLFLKVRGHDREALVQSPALPSTISNQDCQLQFSLYRY
GDFNGTVLLSVVESGASAPALIWERSGHKDAWQEIITLPIITELNGFHLKVQAFWTSGSK
ADIALDDISLSAACFDTELNELLHEGLPHDLDFSLPEPSASEASPIITWFTSCGASGPF
GPTQAQCDAYSANTVSVVVGKGLRQVQMKVPAINTYKISAYGAAGGKAKNHNKRS
HGVIISATFPLEKGDILYLIGHQGEDACPRGNPQTHKICLGESSVIEDGFDSDGSALKW
AGGGGGGGGATYIYRMENQPLPLLAAGGGKAYLEDPSSQDQSFREQYENDTIVSGV
SGRSGAAGGGGWSVSSLSWAGKSLVGGGGSSCPALSVLQWATFGGFGGGGACSA
GGGGGGRGCDAPLLDDISADGQDGLSFVHPMGKIFLQSLAAMESHGAEIVVYLNCSHC
KTQSCKRDEDTKLLCLCDSDEVLADPNVTCAGTKHSLCQFINKHLQHNSSPLVCPPLVV
PMGSLADGPPSLVFMVAVSVTVVTVGVVLTFCASLTLIYRKNHHLHAVRIRLQSPYKLS
KIRSSITIMTDYNPYGFYFGKAASLSELKEVPRKNITLLRALGHGAFGEVYEGQVIGMNGE
NTAMQVAIKTLPEICSEQDEMDFLMEALIMSKFSHQNIIVRCIGVSLQILPRFILLELMTG
GDMKSFRLNRPRTHSSLSMLLELHMARDIALGCRYLEENHFHRDIAARNCLLTCPG
PDRVAKIGDFGMARDIYRASRYRKGGRAMLPVKWMPPEAFLEGIPTCKTDTWVSGVLLWE
IFSLGYMPYPCKTNQEVLEFVTGGGRMDPKSCPGPVYRIMTQCWQHCPEHRPNFTTILE
RINYCTQDPDVTINTPLPVECGPPVEEGGTVIRPDGSGSMTPLLVARSLSQDASPRASIT
SVTPQALKPRLQRPVHLTQEVGTRETLPCWAEVPPASGVCVCPWLVQVPEHRPCSR
SSSSGSQKLNKTKNLWNPYGSWVLESFGRKSAALCHTQSMPLSCNPTSVSAPSSTSEH
TDPVVEVNVANVSASPPSAAPSQTTLPTAAPSRSKSPGAGVSLATVMDLAKLQSFPCG
NVNYAYDEQSYETESLPVVVSKSLPESTSAATSSVALSQASSFTHKPLVKRHSYGH
DVRRYTQPEKPTRDRDGSFSLSEDSLVTPV

```



Sequence Alignment

MAM #1			
hALK_MAM_1	1	LEDCDFEFLVYPPHDLRQCSNWRRIPESSQMDLDCGARRKEMPSSFLLL	60
zLtk_MAM_1	1	STLSEKSEVTSSEHPT---GRTSAGDHS---NRDQVTRVYQKSGREHLL	55
hALK_MAM_1	61	HTLSEKSEVTSSEHPTLAVVHRLQSGRYLQGLFNEAR---	110
zLtk_MAM_1	56	KPSSSHLSAARCSHMTSVLSSCPTEHQLARQPEHAGNISAFVKHTDSIDIKP	114
hALK_MAM_1	111	-HLLMPTPRGQTVLQKRGDPNFRVALDISSSNLSL---VYFPAKNGSE	164
zLtk_MAM_1	115	DITIKKQESRQVYVGLVGLDPEFVYVYSSSHSHPLELSEGLKQW	171
LDLa			
hALK_LDLa	1	---LQETLWHTVVLQDQEDFQCAQDHEGCKAK	37
zLtk_LDLa	1	DCEKYHSEQHSGCTIKQVDEHTDCPEDEKELIG-	39
MAM #2			
hALK_MAM_2	1	FKLHSDVQVTLSPPTKQVH-----TKDINQVQVALLCTVTPVA	51
zLtk_MAM_2	1	SYCFELSGDSSAADTQSSWRLVSGQQLIEDTHLLGTTLKNTQGHFLFLKVRGHD	57
hALK_MAM_2	52	SEKTTTATTAPKSSPTEHRLRVEHVEVIVVHKTQGRVHVVAEY	111
zLtk_MAM_2	58	SEKTTTATTAPKSSPTEHRLRVEHVEVIVVHKTQGRVHVVAEY	114
hALK_MAM_2	112	GLLQVHVEVIVVHKTQGRVHVVAEY	159
zLtk_MAM_2	115	WKADQETTELTELQVHVEVIVVHKTQGRVHVVAEY	162
GlyR			
human_ALK_GlyR	1	GGGGGGGATYIYRMENQPLPLLAAGGGKAYLEDPSSQDQSFREQYENDTIVSGV	56
human_Ltk_GlyR	1	GGGGGGGATYIYRMENQPLPLLAAGGGKAYLEDPSSQDQSFREQYENDTIVSGV	60
zebrafish_Ltk_G..	1	GGGGGGGATYIYRMENQPLPLLAAGGGKAYLEDPSSQDQSFREQYENDTIVSGV	60
human_ALK_GlyR	57	NSGAAGGGGNDNLSLQAGTATSHSPEKMKWNETKGGGGGSSGSSG	116
human_Ltk_GlyR	61	NSGAAGGGGNDNLSLQAGTATSHSPEKMKWNETKGGGGGSSGSSG	120
zebrafish_Ltk_G..	61	NSGAAGGGGNDNLSLQAGTATSHSPEKMKWNETKGGGGGSSGSSG	120
human_ALK_GlyR	117	GGGGGRRG	125
human_Ltk_GlyR	121	GGGGGRRG	129
zebrafish_Ltk_G..	121	GGGGGRRG	129

Percent Identical: 21.5%

Percent Identical: 37.5%

Percent Identical: 28.0%

Percent Identity Matrix - created by Clustal2.1

1: human_Ltk_GlyR	100.00	60.00	59.69
2: human_ALK_GlyR	60.00	100.00	64.00
3: zebrafish_Ltk_GlyR	59.69	64.00	100.00



Fig. S1. (Continued)

C

Alk

Sequence ID: Query_44940 Length: 426 Number of Matches: 2

Range 1: 99 to 168 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
130 bits(70)	2e-34	70/70(100%)	0/70(0%)	Plus/Plus

Query 1 GTGTTTCAGCTGTCAGTGGATATTTACACCTGTGGAGCGACAGGCAAGATGGCCCA 60
 Sbjct 99 GTGTTTCAGCTGTCAGTGGATATTTACACCTGTGGAGCGACAGGCAAGATGGCCCA 158

Query 61 ACACCAACCC 70
 Sbjct 159 ACACCAACCC 168

Range 2: 173 to 217 [Graphics](#) ▼ Next Match ▲ Previous Match ▲ First Match

Score	Expect	Identities	Gaps	Strand
84.2 bits(45)	1e-20	45/45(100%)	0/45(0%)	Plus/Plus

Query 136 GGGATCCAAATGTGGCAGGTTCCAGAGACAAGGAATACAGGTAC 180
 Sbjct 173 GGGATCCAAATGTGGCAGGTTCCAGAGACAAGGAATACAGGTAC 217

Amino Acid Alignment
 WT: 370 WIFHTCGATGQDGPPTQCSNSYRNTNVNVTGKGFPGKIOMWQVPETRKYRITAYGA
 ALK KO: 370 WIFHTCGATGQDGPPTRLGSKCGRFQRQNGTGRHMEQLVDAVFRWRTSHMGCT-

▲
stop

Ltk

Sequence ID: Query_68687 Length: 356 Number of Matches: 1

Range 1: 64 to 230 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
252 bits(136)	4e-71	166/178(93%)	11/178(6%)	Plus/Plus

Query 46 GCAGATGTAGTTTCACATGACCAAGTCCTGTTCTCAGCAGTGGCCATTTTGTCAATC 105
 Sbjct 64 GCAGATGTAGTTTCACATGACCAAGTCCTGTTCTCAGCAGTGGCCATTTTGTCAATC 123

Query 106 TCCAACCTGCTGCTTTCAACCAAGCCACAGCCTGGAACATATCAGCCTTCGTGAAGC 165
 Sbjct 124 TCCAACCTGCTGCTTTCAACCAAGCCACAGCCTGGAACATATC-----ACC 172

Query 166 ACACCGACTTACCGACATCAAAACCAATTGACCTTACAATCAAAGAACAAAGAAAGCGA 223
 Sbjct 173 ACACCGACTTATCGACATCAAAACCAATTGACCTTACAATCAAAGAACAAAGAAAGCGA 230

Amino Acid Alignment
 WT: 115 RCFHMTSPVVLSSGPFCHLQARFQPEPHAGNISAFVKHTDSTDIKPIDLTIKEQES
 LTK KO: 115 RCFHMTSPVVLSSGPFCHLQARFQPEPHAGNISAHRLYRHQTH-

▲
stop

aug-a1

Sequence ID: Query_7615 Length: 318 Number of Matches: 1

Range 1: 98 to 233 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
150 bits(81)	8e-41	131/152(86%)	16/152(10%)	Plus/Minus

Query 2 CACTCTGGTTTCGGGAGAGATGCGCGCGCTGCCAGCCCGGTGCTGTAATGGGCTCGT 61
 Sbjct 233 CACTCTGGTTTCGGGAGAGATGCGCGCGCTGCCAGCCCGG-----GGCTCGT 186

Query 62 ATGTGTAATCTGCACTGCCGACAAAGCGATGCCAGCCGAAACAAAGTGGAGAAAACGCT 121
 Sbjct 185 ATGTGTAATCTGCAAT-C-GTT-AAT-CGATGCCAGCCGAAACAAAGTGGAGAAAACGCT 130

Query 122 CAGACGGATCATGGAATCATGAGACAGGTGG 153
 Sbjct 129 CAGACGGATCATGGAATCATGAGACAGGTGG 98

Amino Acid Alignment
 WT: 1 MRALRAPVLMGLVLLICTAAQSDASANKVEKTLRRIMEIMRQVENSADESAQKTESAPEPKDT
 AugKO: 1 MRALRAPGLVLLICNR-

▲
stop

aug-a2

Sequence ID: Query_210167 Length: 411 Number of Matches: 1

Range 1: 98 to 318 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
335 bits(181)	1e-95	213/227(94%)	8/227(3%)	Plus/Minus

Query 134 CACCTTTCAGCATGAGTGCAGTGCGCCACCTGTC-TTC-ATAGGCTCCTACTGCTGAT 191
 Sbjct 318 CACCTTTCAGCATGAGTGCAGTGCGCCACCTGTCCTACTTTAGGGCTCCTACTGCTGAT 259

Query 192 CCTCACCAACCGCTACTGCAACCGAGAGACAGAGACGAAACAGTCTGCTGGAGCTCTT 251
 Sbjct 258 CCTCACCAACCGC-----AACCGRAGRACGGGACGAAACAGTCTGCTGGAGCTCTT 205

Query 252 AATGGACAGATGAGACAGACACAGGAGCATCACAGTGGAGGAAACACACAGCATCCTCC 311
 Sbjct 204 AATGGAGAGATGAGACAGACACAGGAGCATCACAGTGGAGGAAACACACAGCATCCTCC 145

Query 312 TCAGATCATGAGCATTCGCTGGAAACAAAGACGTTAATAAGTCA 358
 Sbjct 144 TCAGATCATGAGCATTCGCTGGAAACAAAGACGTTAATAAGTCA 98

Amino Acid Alignment
 WT: 1 MCFALSLSIYKHTGTPPNKRLNSALKRNRATVRLDIGRGSGLCPPFSMSAVRPPVF IGLLLLL
 AugKO: 1 MCFALSLSIYKHTGTPPNKRLNSALKRNRATVRLDIGRGSGLCPPFSMSAVRPPVLL-

▲
stop

aub-b

Sequence ID: Query_197261 Length: 418 Number of Matches: 2

Range 1: 164 to 275 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
183 bits(99)	3e-50	112/117(96%)	5/117(4%)	Plus/Plus

Query 106 ACTCTGCTTAATCTAATCTGCAAGTGTGGGAAAAGCCCGCTCCAGACGCGTCACC 165
 Sbjct 164 ACTCTGCTTAATCTAATCTGCAAGTGTGGGAAAAGCCCGCTCCA-----GTCACC 218

Query 166 AGTGGCCTTACTCTGTATCTCAAGATGTAAGTTTCTCGCGTAAAACACTGCA 222
 Sbjct 219 AGTGGCCTTACTCTGTATCTCAAGATGTAAGTTTCTCGCGTAAAACACTGCA 275

Range 2: 95 to 155 [Graphics](#) ▼ Next Match ▲ Previous Match ▲ First Match

Score	Expect	Identities	Gaps	Strand
108 bits(58)	2e-27	60/61(98%)	0/61(0%)	Plus/Plus

Query 1 ATCGGGCTGAGAAGATGGCACATCTTGTGAGTATGATCTTTTGCTCATCACCTCC 60
 Sbjct 95 ATCGGGCTGAGAAGATGGCACATCTTGTGAGTATGATCTTTTGCTCATCACCTCC 154

Query 61 A 61
 Sbjct 155 A 155

Amino Acid Alignment
 WT: 2 RAEKRWHILLSMILLITSSQCMDSKEVKESEKRLNLLILQVIGEPASRRVTSGLYSVSQDAKFPSS
 AugKO: 2 RAEKRWHILLSMILLITSTSI TLLNLLILQVIGEPASSHOWPLLCISRC-

▲
stop

Fig. S1. (A) Comparison of the full-length and augmentor domain primary sequence of zebrafish and human augmentors. (B) Ltk primary sequence with domains highlighted. Sequence comparison of Ltk domains with their human ALK counterparts shows that the GlyR region of Ltk is more closely related to human ALK than to LTK. (C) Sequencing of knockout Alk, Ltk, and augmentor fish resulting in frame-shift mutations (red text denotes the mutant amino acid sequences) and subsequent successful introduction of a premature stop codon (red arrowheads).

