Current Biology, Volume *31*

Supplemental Information

Modulation of bioelectric cues in the evolution of flying fishes

Jacob M. Daane, Nicola Blum, Jennifer Lanni, Helena Boldt, M. Kathryn Iovine, Charles W. Higdon, Stephen L. Johnson, Nathan R. Lovejoy, and Matthew P. Harris

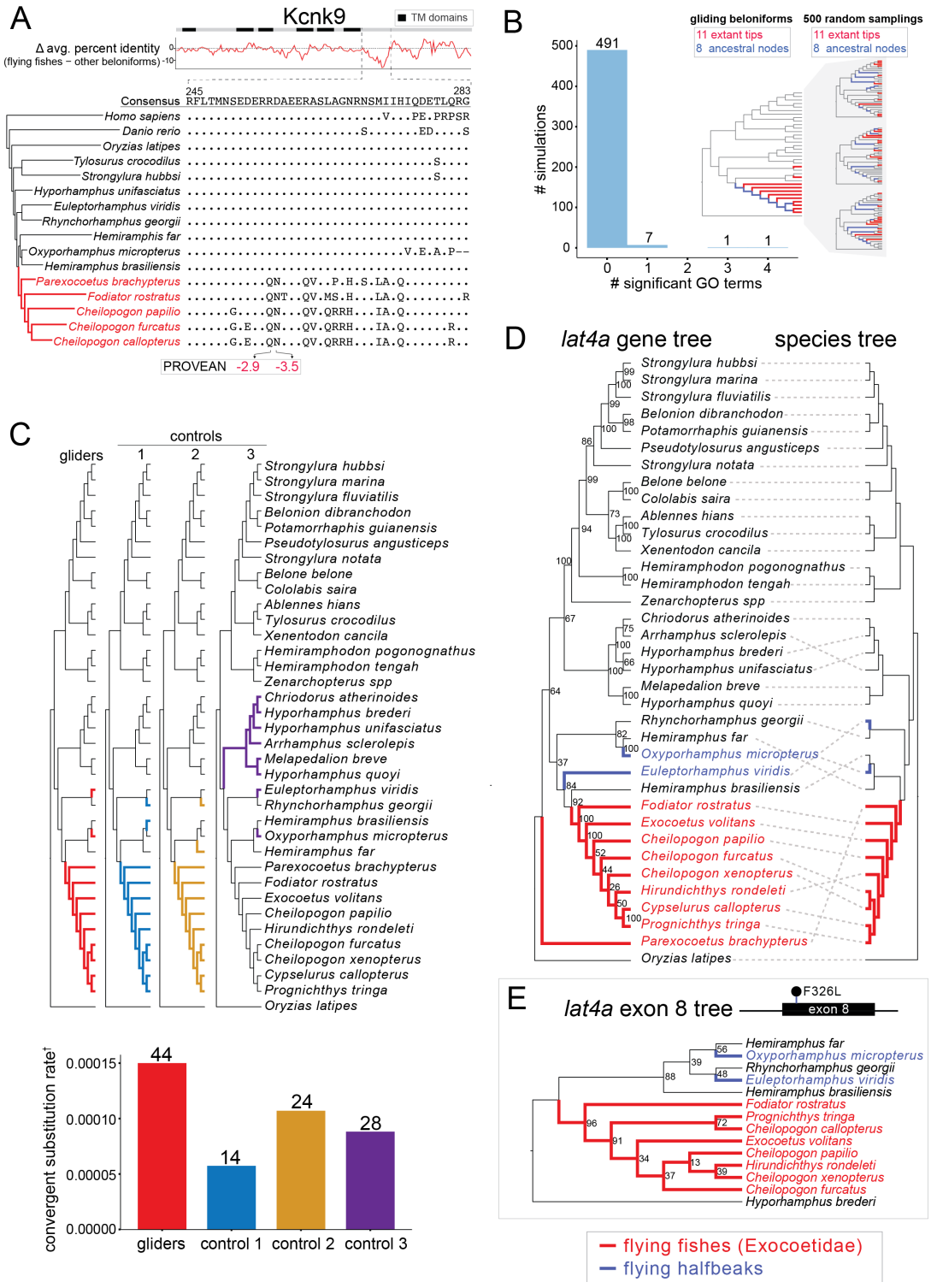

Figure S1. Comparative genomic trends across gliding beloniforms and control species samplings, related to Figure 2.

(A) Accelerated sequence evolution in the potassium channel Kcnk9 in flying fishes. Top panel, 30 bp genomic sliding window of average percent identity to the Japanese medaka genome (*O. latipes*) across

kcnk9 in flying fishes compared to the other beloniforms. Bottom panel, protein multiple sequence alignment of accelerated region in Kcnk9. Predictions of deleterious amino acid substitutions by PROVEAN (≤ -2.5) are indicated[S1].
(B) Simulation of branch topology to assess background levels of divergence in evolutionary rate across gene ontology groupings. For each species sampling, the difference between the average relative evolutionary rate across each GO-term were compared between the selected branches and the background branches. Significance was determined by Wilcoxon signed-rank test and p-values corrected by Benjamini-Hochberg procedure.
(C) Highlighted branches for the gliding beloniforms (red), and three topologically similar control groupings of flying fishes and halfbeaks (blue, yellow, purple). Convergent substitution rate (†) refers to the number of identical amino acid substitutions found in all three highlighted clades normalized by the number of total unique substitutions found within at least one of the three clade selections and not within background branches. The number above each bar indicates total number of convergent amino acids.
(D) Comparison of *lat4a* gene tree and species tree. The gliding beloniforms are specifically highlighted, with the flying fish branches (Exocoetidae) in red, and flying halfbeaks in blue. Node labels indicate ultrafast bootstrap support values for maximum likelihood tree (IQTree).
(E) *lat4a* phylogenetic tree centered on the exon and the flanking non-coding sequence surrounding the F326L mutation in gliding beloniforms.

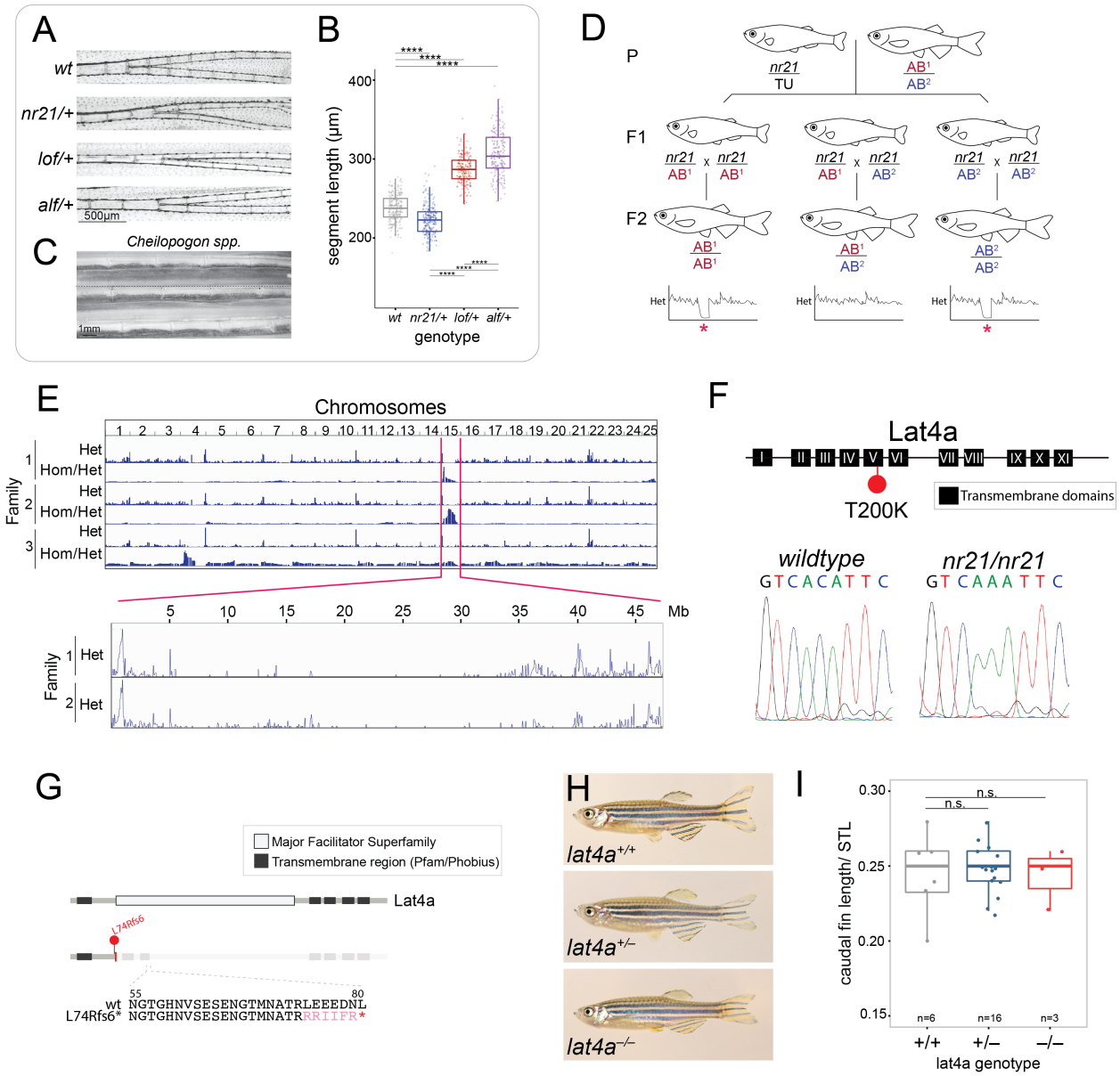

Figure S2. Mapping and detailed phenotyping of the *nr21* shortfin zebrafish mutant, related to Figure 3.

(A-C) Analysis of fin lepidotrichia segment patterning in zebrafish mutants.

(A) Images of pectoral fin lepidotrichia segments in wildtype (*wt*), short finned (*nr21*) and long finned (*lof*, *alf*) zebrafish mutants.

(B) Segment length across pectoral fins in zebrafish mutants. Roughly 30 segments were measured per fin across n=8 individuals for *wt* and *nr21* and from n=6 individuals for *lof* and *alf*. Note variable segment length in *alf* fins. **** indicates Tukey HSD adjusted p-value <0.0001.

(C) Pectoral fin lepidotrichia from a flying fish, *Cheilopogon spp*.

(D-F) Mapping of *nr21* mutation.

(D) Map cross strategy for identifying *nr21* mutation. As *nr21* is dominant, mapping was performed on F2 sibling pools for the wildtype (recessive) phenotype. However, given SNP diversity in wildtype

zebrafish strains, three separate options were possible depending on the F1 parental cross. Some of the F1 crosses would be homogeneous for particular wildtype haplotype (AB1/AB1 or AB2/AB2), but others may be heterozygous at the locus (AB1/AB2), obfuscating the mapping signal.
(E) Genome-wide patterns of heterozygosity and the ratio of heterozygous to homozygous SNPs along a 15 centimorgan (cM) sliding window. Of the three wildtype families, two families showed a strong ratio homozygosity/heterozygosity on chromosome 15 between 17-33Mb. Within this interval, three non-synonymous SNPs (*med13b*: Q1531H, *lat4a*: T200K, *si:dkey-285b23.3*: E354G) were identified in *nr21* carriers. There were 0 recombinants among 145 chromosomes for the *lat4a* T200K SNP.
(F) Sanger sequencing traces for the causative *nr21* non-synonymous point mutation (T200K).
(G-I) *Lat4a* is not required for normal fin growth.
(G) Truncating frameshift allele (L74Rfs6) in *lat4a* generated through CRISPR/Cas9.
(H) Images of wildtype, heterozygous and homozygous *lat4a* knockouts with wildtype fin patterning and no obvious phenotype.
(I) Measurements of caudal fin length normalized to fish standard length (STL) showing no significant difference in fin size in the absence of *lat4a* (Tukey's HSD).

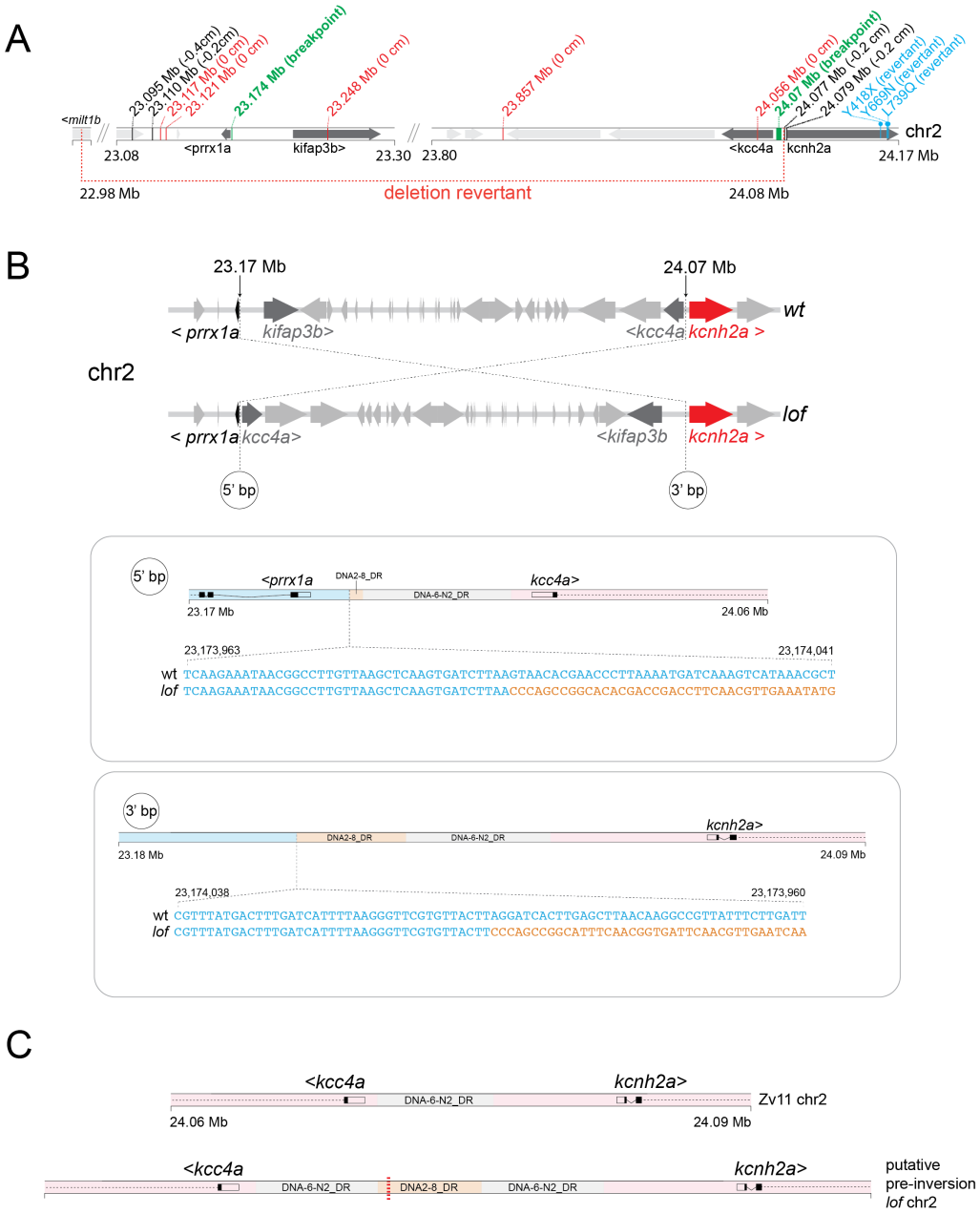

Figure S3. Mapping of the *longfin* zebrafish (*lof*^{dt2}), related to Figure 3.

(A) Combined recombinant and revertant data for *lof*^{dt2}. Linkage mapping identified a large region (~1 Mb) on chromosome 2, without recovery of additional recombinants. Four total reversion alleles were isolated, three point mutations in *kcnh2a* (Y418X, Y669N, L739Q) and a large deletion between 22.98 Mb and 24.08 Mb, just upstream of *kcnh2a* (*lof*^{fi6e1}).

(B) Local assembly of PacBio sequencing reads identifies inversion breakpoints (bp) upstream of *prrx1a* and upstream of *kcnh2a*.

(C) Reconstructed ancestral *lof*^{dt2} sequence prior to inversion, showing predicted transposon expansion relative to the zebrafish reference genome from the Tübingen strain (Zv11). Coordinates represent Zv11 positions.

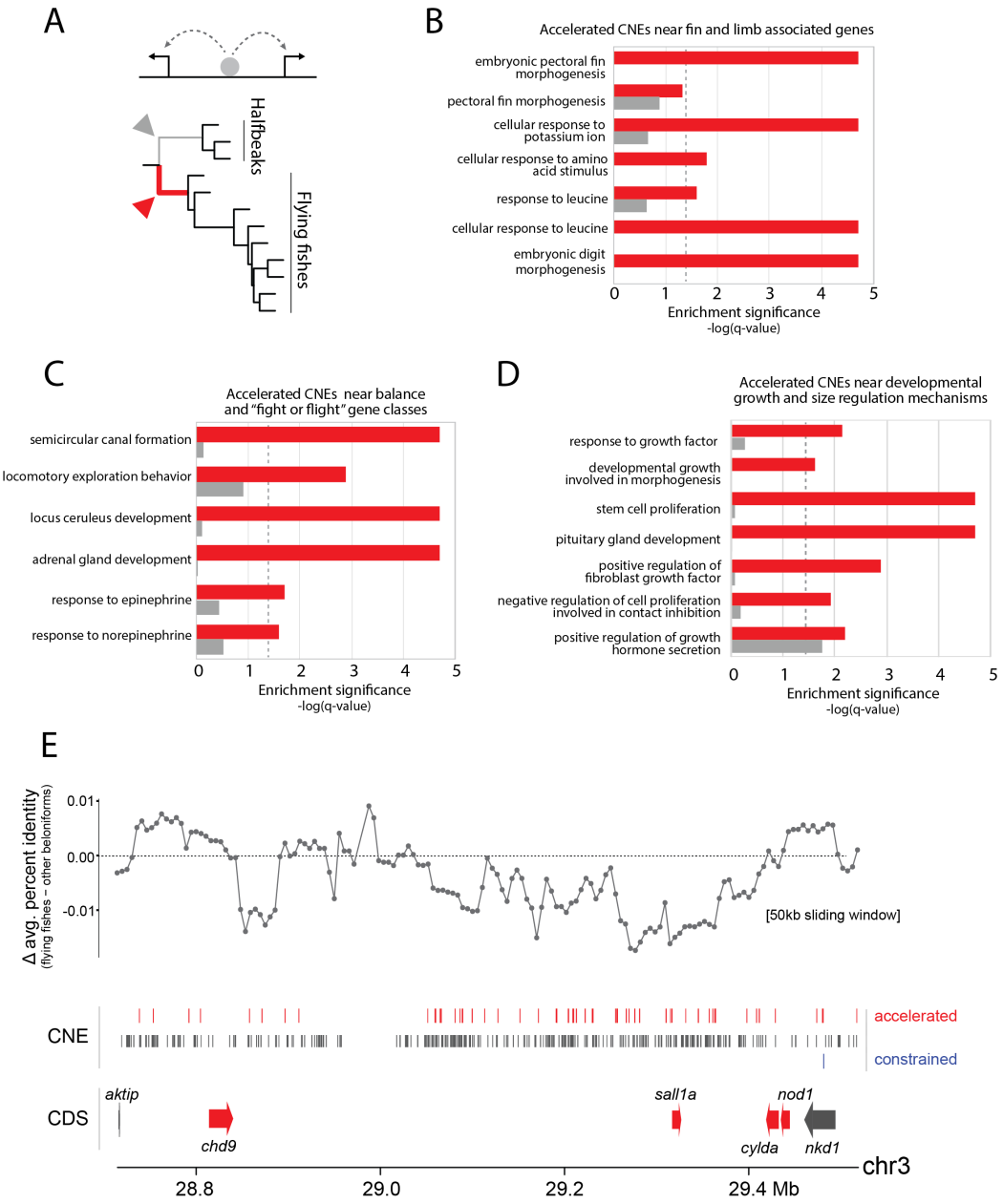

Figure S4. Accelerated sequence evolution in predicted gene regulatory regions that are associated with morphological and behavioral traits in flying fishes, related to Figure 2.

(A) Analysis of accelerated sequence evolution using the program phyloP, with a focus on the ancestral common ancestor of flying fishes (Exocoetidae; red) compared to the sister lineage of halfbeak fishes (gray). CNEs were assigned to neighboring genes through the GREAT approach[S2] for gene ontology enrichment analysis.

(B-D) Significant enrichment of accelerated sequence evolution in CNEs near genes associated with fin and limb development (B), balance and fear response (C), and general organ and tissue growth and size regulation (D). For full list of significantly enriched terms under acceleration and constraint, see **Data S1K-N**.

(E) Accelerated sequence evolution at the *sall1a* locus in flying fishes. Top panel, 50kb genomic sliding window of average percent identity to the Japanese medaka genome (*O. latipes*) across targeted

elements (CDS and CNE) in flying fishes compared to the other beloniforms. CNEs and CDS considered by the program phyloP to be under accelerated sequence evolution (red), neutral (gray) or under constraint (blue) along the ancestral branch of flying fishes.

Species	# Individuals Sequenced	CDS capture	CNE capture	Avg nucleotide diversity (π)
Ablennis hians	5	✓	✓	0.0025
Arrhamphus sclerolepis	2	✓	✓	0.0030
Belone belone	5	✓	✓	0.0044
Belonion dibranchodon	5	✓	✓	0.0020
Cheilopogon furcatus	7	✓	✓	0.0059
Cheilopogon papilio	7	✓	✓	0.0057
Cheilopogon xenopterus	8	✓	✓	0.0054
Chriodorus atherinoides	8	✓	✓	0.0027
Cololabis saira	1	✓	✓	0.0101
Cypselurus callopterus	8	✓	✓	0.0048
Euleptorhamphus viridis	5	✓	✓	0.0050
Exocoetus volitans	8	✓	✓	0.0030
Fodiator rostratus	5	✓	✓	0.0063
Hemiramphus brasiliensis	5	✓	✓	0.0042
Hemiramphus far	6	✓	✓	0.0053
Hemiramphus unifasciatus	8	✓	✓	0.0027
Hemirhamphodon pogonognathus	5	✓	✓	0.0049
Hemirhamphodon tengah	1	✓	✓	0.0043
Hirundichthys rondeleti	5	✓		0.0055
Hyporhamphus brederi	6	✓	✓	0.0036
Hyporhamphus quoyi	5	✓	✓	0.0031
Melapedalion breve	2	✓	✓	0.0024
Oxyporhamphus micropterus	8	✓	✓	0.0044
Parexocoetus brachypterus	6	✓	✓	0.0057
Potamorrhamphis guianensis	8	✓	✓	0.0032
Prognichthys tringa	7	✓	✓	0.0053
Pseudotylosurus angusticeps	6	✓	✓	0.0014
Rhynchorhamphus georgii	2	✓	✓	0.0030
Strongylura fluviatilis	5	✓	✓	0.0017
Strongylura hubbsi	3	✓		0.0075
Strongylura marina	5	✓	✓	0.0034
Strongylura notata	4	✓	✓	0.0024
Tylosurus crocodilus	5	✓	✓	0.0041
Xenentodon cancila	5	✓	✓	0.0059
Zenarchopterus spp	2	✓	✓	0.0099

Table S1. Pooled populations and targeted sequence capture, related to Figure 1

	Reference genome of target													
	Medaka		Platyfish		Amazon Molly		CDS*		CNE*†		miRNA*		UCNE*	
Species	Cov.	Depth	Cov.	Depth	Cov.	Depth	Cov.	Depth	Cov.	Depth	Cov.	Depth	Cov.	Depth
Xenentodon cancila	80.5%	40.2	63.8%	24.5	41.8%	49.2	81.3%	35.3	78.1%	48.9	91.1%	51.1	97.5%	78.6
Ablennis hians	80.4%	34.2	64.6%	22.2	44.0%	26.8	81.1%	32.3	78.5%	37.7	89.3%	36.5	97.3%	51.3
Tylosurus crocodilus	75.3%	43.8	57.9%	23.7	40.8%	44.8	73.2%	32.8	78.8%	65.2	89.0%	52.3	96.5%	78.3
Strongylura notata	81.9%	46.7	66.8%	32.9	45.9%	47.0	83.5%	48.1	78.1%	43.1	89.2%	53.1	97.7%	70.3
Belone belone	81.4%	43.7	66.3%	34.3	46.6%	53.1	83.2%	48.2	77.2%	34.4	89.6%	41.9	96.4%	49.9
Cololabis saira	80.2%	41.3	63.3%	29.1	44.5%	54.0	81.3%	43.4	77.4%	36.4	89.8%	39.9	96.5%	58.1
Pseudotylosurus angusticeps	81.8%	38.5	63.1%	24.9	32.9%	22.1	81.8%	38.5	-	-	-	-	-	-
Strongylura marina	76.4%	37.4	59.0%	24.8	39.2%	31.0	75.7%	35.8	77.1%	40.0	88.2%	37.4	96.6%	53.5
Strongylura hubbsi	80.4%	28.7	62.4%	18.3	35.2%	21.6	80.4%	28.7	-	-	-	-	-	-
Strongylura fluviatilis	81.1%	40.2	65.0%	24.4	44.2%	26.7	82.2%	33.8	78.2%	51.5	89.3%	57.4	97.7%	88.8
Potamorrhamphis guianensis	78.7%	33.0	61.2%	20.0	40.4%	36.6	78.8%	27.2	77.7%	43.8	90.0%	46.1	97.5%	71.1
Belonion dibranchodon	76.6%	23.6	59.4%	17.7	36.0%	17.1	78.6%	27.1	71.8%	16.1	85.0%	21.7	95.4%	29.2
Zenarchopterus spp	82.6%	59.2	67.9%	38.6	46.9%	46.8	84.5%	55.7	78.1%	64.6	90.5%	63.6	97.4%	108.0
Hemirhamphodon pogonognathus	80.0%	49.5	63.9%	35.9	39.7%	31.3	82.7%	56.1	73.7%	35.4	87.3%	41.4	95.9%	66.2
Hemirhamphodon tengah	79.4%	43.2	62.8%	29.6	37.6%	26.9	81.7%	45.4	73.9%	37.6	88.3%	41.6	95.6%	71.9
Melapedalion breve	81.6%	39.4	67.0%	32.2	46.5%	38.8	83.7%	46.2	76.7%	25.5	88.2%	30.2	96.2%	39.3
Hyporhamphus quoyi	82.0%	43.8	67.7%	36.1	47.9%	53.1	84.2%	51.5	77.0%	27.9	88.3%	36.5	96.6%	43.2
Arrhamphus sclerolepis	78.1%	29.1	63.0%	23.6	43.5%	33.1	80.3%	36.3	73.1%	14.4	83.2%	21.3	94.2%	22.9
Chriodorus atherinoides	81.6%	51.6	65.8%	32.6	46.1%	44.9	82.9%	45.3	78.2%	63.0	89.2%	67.2	97.0%	94.6
Hemiramphus unifasciatus	71.6%	25.5	52.7%	16.6	36.8%	19.0	69.1%	23.5	75.7%	28.7	85.8%	32.1	95.8%	46.3
Hyporhamphus brederi	72.4%	31.3	54.0%	16.7	39.3%	28.9	69.4%	20.7	77.8%	51.1	88.7%	53.6	96.8%	89.3
Euleptorhamphus viridis	75.1%	35.4	57.6%	21.3	40.7%	35.6	73.6%	31.2	77.5%	43.0	88.9%	39.6	97.0%	61.2
Rhynchorhamphus georgii	81.6%	47.3	66.2%	26.5	45.3%	31.6	82.6%	37.4	79.1%	66.7	89.5%	53.0	97.6%	71.4
Hemiramphus far	80.6%	35.9	64.1%	20.8	45.6%	34.2	80.4%	29.0	80.2%	49.0	89.3%	40.1	97.6%	63.5
Hemiramphus brasiliensis	72.6%	25.4	56.1%	19.5	39.3%	40.5	72.7%	30.5	71.5%	14.8	82.4%	18.3	95.3%	25.2
Oxyporhamphus micropterus	77.0%	32.6	59.1%	18.2	41.0%	32.5	76.1%	25.3	78.3%	47.0	89.1%	40.7	95.9%	56.1
Parexocoetus brachypterus	78.1%	36.4	61.3%	20.8	43.6%	32.4	78.4%	29.4	76.8%	50.0	89.4%	42.5	96.3%	59.5
Fodiator rostratus	72.1%	37.0	53.6%	20.7	37.0%	26.2	68.7%	29.8	78.1%	50.6	89.0%	49.2	96.7%	72.6
Exocoetus volitans	67.6%	25.3	48.8%	15.1	34.5%	33.8	63.5%	22.0	74.9%	31.3	86.3%	28.6	95.5%	40.9
Cheilopogon papilio	80.3%	38.5	63.6%	24.2	43.9%	29.1	80.7%	33.8	78.8%	46.9	89.9%	48.7	97.2%	72.1
Cypselurus callopterus	78.9%	44.3	61.4%	25.4	42.5%	32.8	78.5%	35.5	79.0%	60.6	89.2%	58.8	97.3%	92.1
Prognichthys tringa	72.5%	36.3	54.2%	20.9	37.9%	28.7	69.8%	30.2	77.2%	47.5	88.5%	44.0	97.0%	75.2
Cheilopogon furcatus	79.5%	35.0	62.7%	22.4	44.0%	31.6	79.8%	31.4	78.2%	41.3	89.0%	39.3	97.2%	60.2
Hirundichthys rondeleti	78.4%	45.2	59.9%	28.7	37.0%	62.9	78.4%	45.2	-	-	-	-	-	-
Cheilopogon xenopterus	80.6%	41.0	64.4%	26.0	43.8%	32.7	81.2%	36.6	78.8%	48.9	89.0%	46.3	97.3%	71.7

* - medaka (*Oryzias latipes*) reference genome targets

† - not inclusive of miRNAs and ultraconservative elements (UCNE)

Table S2. Coverage breakdown, related to Figure 1

	coverage†			
	1X	2X	4X	10X
Xenentodon cancila	80.5%	80.1%	78.0%	67.9%
Ablennis hians	80.4%	79.9%	77.4%	65.8%
Tylosurus crocodilus	75.3%	74.3%	70.5%	59.5%
Strongylura notata	81.9%	81.5%	79.9%	72.3%
Belone belone	81.4%	81.1%	79.4%	72.1%
Cololabis saira	80.2%	79.8%	78.0%	69.7%
Pseudotylosurus angusticeps	81.8%	81.4%	79.3%	69.4%
Strongylura marina	76.4%	75.6%	72.1%	61.4%
Strongylura hubbsi	80.4%	79.8%	76.7%	64.5%
Strongylura fluviatilis	81.1%	80.6%	78.4%	69.0%
Potamorrhamphis guianensis	78.7%	78.1%	75.2%	62.9%
Belonion dibranchodon	76.6%	75.9%	72.4%	58.5%
Zenarchopterus spp	82.6%	82.4%	81.4%	75.8%
Hemirhamphodon pogonognathus	80.0%	79.7%	78.5%	72.3%
Hemirhamphodon tengah	79.4%	79.0%	77.4%	69.6%
Melapedalion breve	81.6%	81.2%	79.5%	70.9%
Hyporhamphus quoyi	82.0%	81.7%	80.1%	72.5%
Arrhamphus sclerolepis	78.1%	77.5%	74.4%	60.8%
Chriodorus atherinoides	81.6%	81.2%	79.4%	71.8%
Hemiramphus unifasciatus	71.6%	70.4%	65.7%	53.0%
Hyporhamphus brederi	72.4%	71.3%	66.8%	54.5%
Euleptorhamphus viridis	75.1%	74.1%	70.1%	58.3%
Rhynchorhamphus georgii	81.6%	81.2%	79.3%	69.5%
Hemiramphus far	80.6%	80.0%	77.4%	65.9%
Hemiramphus brasiliensis	72.6%	71.4%	66.5%	51.9%
Oxyporhamphus micropterus	77.0%	76.4%	73.7%	60.5%
Parexocoetus brachypterus	78.1%	77.6%	75.0%	63.0%
Fodiator rostratus	72.1%	71.0%	66.7%	54.6%
Exocoetus volitans	67.6%	66.2%	60.8%	47.6%
Cheilopogon papilio	80.3%	79.8%	77.6%	67.1%
Cypselurus callopterus	78.9%	78.3%	75.7%	64.5%
Prognichthys tringa	72.5%	71.4%	67.0%	55.7%
Cheilopogon furcatus	79.5%	79.0%	76.4%	65.1%
Hirundichthys rondeleti	78.4%	77.6%	74.4%	63.9%
Cheilopogon xenopterus	80.6%	80.2%	78.1%	68.1%

Table S3. Sequencing read coverage by depth, related to Figure 1

Supplemental references

S1. Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLoS One *7*, e46688.

S2. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.