A fish-specific new gene evolves new functions in cichlid fishes

Langyu Gu[1,2,3,]*, Chenzheng Li[2], Xiaobing Mao[2], Zongfang Wei[2], Youkui Huang[2], Ximin He[1], Wenjun Zhou[1], Li Li[2],* Deshou Wang[2]*

[1] State Key Laboratory of Biocontrol, School of Life Sciences and School of Ecology, Sun Yat-sen University, Guangzhou, Guangdong, China.

[2] Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education), Key Laboratory of Aquatic Science of Chongqing, School of Life Sciences, Southwest University, Chongqing, China

[3] Zoological Institute, Department of Environmental Sciences, University of Basel, Basel, Switzerland

*Correspondence:

Langyu Gu: langyugu@gmail.com

Li Li: alisir@swu.edu.cn

Deshou Wang: wdeshou@swu.edu.cn

Abstract

Genes whose dysfunction do not affect normal survival is a common phenomenon. Are they only meaningless residues during evolution? Here, we identified a fish-specific new gene, which we named *lg*. Gene knockout resulted in no obvious phenotype in zebrafish, but *lg* evolves an amino acid mutation that is under positive selection in the modern haplochromine (MH) cichlid fish lineage, the well-known adaptive radiative lineage. Moreover, the cichlid fish-specific upstream region of *lg* can drive new eGFP expression profiles in tissues related to adaptation. Besides, this homologous region from different cichlid fishes drive different eGFP expressions which is simply due to three MH-segregated SNPs mutations that are predicted to bind a hormone related transcription factor. We thus revealed an initially redundant gene evolving new functions in an adaptive radiative lineage. This further illuminates the mechanism about the emergence of new gene function with respect to evo-devo in a broad way.

Key words

lg, new gene, lineage-specific, evo-devo, fish

From the evolutionary point of view, the most adaptive organism is the one that spends the least resources but produce the most outputs. However, it seems that this rule does not apply to the genome organization, which is occupied by a large proportion of genes whose dysfunction does not affect normal survival. Does it make sense?

Nothing in biology makes sense except in the light of evolution [1]. As raw materials for natural selection, the large portion of redundant genes provides good opportunities for functional innovation. This is especially true for lineage-specific genes, which can even occupy up to one-third of the genome in many organisms [2]. Studies gradually revealed that many of them can evolve new functions related to novelties and environmental adaptation in the particular lineage. For example, in early brain development in human [3,4], in the creation of novel tissues in the sea anemone [5], and in the development of a novel fan trait on the leg in a water spider genus to explore new environment [6]. There is even a whole lineage-specific gene family in water flea related to environmental stimuli [7]. Lineage-specific gene thus provides a new perspective of new gene function emergence.

It is well accepted that different organisms share deep homologous developmental molecular bases (molecular toolkit), such as highly conserved signalling pathways and transcription factors [8–10]. Then, how do new genes gain new functions, and in what way are they integrated in the existing networks? Although usually restricted due to pleiotropic effects [11], protein coding region evolution can take place [12–14]. Compare to this, regulatory evolution is a well-known strategy, since changes at the expression level are more tolerable and can offer immediate phenotypic consequences [11,15]. Even a single mutation in the upstream region can induce new expression profiles [16,17]. The relaxed selection

86 pressure induced by the changing expression profiles can even further prompt
87 the accumulation of mutations at the protein level.
88
89 In this study, we identified a fish-specific new gene, which we named *lg*.
90 Comparative -omics analyses and functional experiments revealed that multiple
91 losses of *lg* occurred in different fish lineages, and no obvious phenotype was
92 detected in gene knockout zebrafish, which make us assume that the ancestral
93 function of *lg* is not indispensable. *lg* mainly expressed in the brain and eyes in
94 most fishes, but evolved new expression patterns in cichlid fish lineage,
95 especially in the most-species richness lineage, the modern haplochromine
96 (MH) lineage, such as in pharyngeal cartilages and epidermal serotonin-positive
97 neuroepithelia cells. Further comparative genomic analyses across East African
98 cichlid fish phylogeny revealed that *lg* evolved a cichlid fish-specific upstream
99 region possessing three MH lineage-segregated SNPs and one MH lineage-
100 segregated amino acid mutation which is under positive selection. Transgenic
101 experiments confirm the new regulatory functions of this upstream region to
102 drive eGFP expression in zebrafish. Noticeably, the homologous sequence from
103 different cichlid fishes drive different eGFP expression patterns, which can be
104 simply due to the three MH lineage-segregated SNPs mutations that are
105 predicted to bind to a hormone related transcription factor. Together, our work
106 revealed an initially redundant fish-specific new gene evolves new regulatory
107 functions in an adaptive radiative cichlid fish lineage for the first time.
108
109 **Results**
110
111 ***lg* is a fish-specific gene**
112 To see *lg* sequence evolution, we conducted tblastx search in all the
113 species with available genomes in Ensembl 100. Results showed that *lg* only
114 appears in fishes (Fig. 1A, Table S1). To further characterize its evolution

115 within teleost fishes, we conducted tblastx search and gene predictions in draft
116 genomes assembly across the whole phylogeny of teleost fishes [18]. Results
117 showed that *lg* has been lost multiple times in different lineages (Fig. 1B),
118 which inspired the assumption that the ancestral function of *lg* is not
119 indispensable for fish survival.
120

121 **_lg_ evolves new expression profiles in cichlid fishes, especially in the MH**
122 **lineage**

123 To see the tissues where *lg* expressed, we investigated its expression
124 profiles based on the combined results of available RNAseq data retrieved from
125 NCBI and *in situ* hybridization in different fishes (Fig. 2, Table S2). *lg* mainly
126 expressed in the brain and eyes in most fishes. Interestingly, it evolved new
127 expression profiles in cichlid fishes in pharyngeal cartilages, as well as in the
128 skin in the most species-richness lineage of cichlid fishes, the MH lineage [19].
129 These results suggest that the ancestral function of *lg* can be related to neurons,
130 but new functions can evolve in other tissues in cichlid fishes, especially in the
131 MH lineage.
132

133 **No obvious phenotype was detected in _lg_ gene knockout (KO) zebrafish**

134 To further investigate the function of *lg*, KO was conducted using
135 CRISPR/Cas9 in zebrafish. We successfully obtained F2 homozygous mutant
136 (MUT) zebrafish and its wild-type siblings (WT). Sanger sequencing and
137 Western-blot analyses confirmed the successful gene dysfunction with
138 frameshift (5 bp deletion) and protein interruption (stop codon appeared at the
139 18th amino acid) (Fig. 3A). To see the affected gene expressions after KO,
140 comparative transcriptomic analyses were conducted between MUT and WT at
141 2 dpf (days post fertilization) and 4 dpf zebrafish, respectively. The results
142 showed that most differential expressed (DE) genes are related to immune,
143 neurons and energy (Table S3). To see whether there are different immune

144 responses between MUT and WT, we examined the immune responses of larvae
145 zebrafish to lipopolysaccharide (LPS) by counting the numbers of *lyz*$^{+}$
146 neutrophils in the caudal hematopoietic tissue at one day after injection [20]. No
147 significant difference was detected (Fig. 3B). To see whether KO affect neurons
148 development in the brain and eyes where *lg* expressed in zebrafish, we
149 compared the phenotypic changes between MUT and WT using the brain
150 neuron marker, HuC [21]. Still no difference was detected (Fig. 3C). Then we paid
151 our attention to one top DE gene, *slc16a6*, which has been reported to be related
152 to energy metabolism under fasting condition [22]. With the same strategy,
153 however, we still could not find obvious metabolism phenotypic changes
154 between MUT and WT under fast condition using ORO staining (Fig. 3C). This
155 can be explained by the fact that *slc16a6* is only down-regulated here instead of
156 entirely disrupted as in the [22], so that it can still execute its funciton. Above all,
157 mutation of *lg* resulted in no obvious phenotypes in zebrafish. These results
158 further support our assumption that the ancestral function of *lg* is not
159 indispensable.
160

161 **_lg_ evolves a cichlid fish-specific upstream region possessing MH-segregated**
162 **SNPs and one MH-segregated amino acid mutation that is under positive**
163 **selection**

164 Noticeably, as we mentioned above, *lg* evolved new gene expression
165 profiles in cichlid fishes (Fig. 2), the classical model species for studying
166 adaptive radiation, which explosively radiated into hundreds to thousands of
167 species within a short time with very closely related genomes [23]. To characterize
168 its sequence evolution within cichlid fishes, we amplified the whole *lg* gene
169 including its upstream region in a total of 20 cichlid fishes representing main
170 lineages across East African cichlid fish phylogeny [24] (Fig. 4A). Interestingly,
171 both upstream region and the coding region exhibited segregation patterns with
172 the MH lineage, the most species-richness lineage of cichlid fishes [25] (Fig. 4A).

173 For example, there is one nonsynonymous amino acid mutation (A to S) located
174 at the 38th residue specific to the MH lineage which is under positive selection
175 (Figure 4A, Table S4). Furthermore, *lg* evolved an upstream region specific to
176 cichlid fishes (Fig. S1). Within this region, there are three MH lineage-
177 segregated SNPs mutations (Fig. 4A). Transcription factors (TFs) binding
178 prediction revealed that these three SNPs can set up the binding opportunities to
179 a hormone related TF, the glucocorticoid modulatory element binding protein 2
180 (GMEB2) (Fig. 4A). These results stimulate us to pay special attention to the *lg*
181 of cichlid fishes in the next steps.
182

183 **The cichlid fish-specific upstream region indeed has regulatory functions**
184 **and performs different roles in different cichlid fish lineages**

185 To see whether the cichlid fish-specific upstream region of *lg* indeed has
186 regulatory functions, and whether this region of MH lineage has different
187 regulatory roles compared to other cichlid fish lineages, we constructed two
188 transgenic lines with the homologous upstream region from different species as
189 the promoters to drive eGFP expression in zebrafish (Fig. 4A, Fig. S2). One line
190 is with the upstream region of a MH cichlid fish*, Pundamilia nyererei* (pT2AL-
191 Pn-eGFP transgenic line, Pn line). The other line is with the corresponding
192 homologous region of a cichlid fish belonging to a sister lineage of East African
193 Great lakes cichlid fishes, Nile tilapia (*Oreochromis niloticus*) [23] (pT2AL-On-
194 eGFP transgenic line, On line). We successfully obtained F2 generations for
195 both transgenic lines. Both lines can drive eGFP expressions in zebrafish,
196 confirming their regulatory functions, but with different patterns (Fig. 4B and
197 Fig. 4C). Pn line showed strong expressions in neurons in most regions of the
198 brain (olfactory epithelium, olfactory bulb, forebrain, optic tectum, cerebellum
199 and hindbrain), eyes, spinal cord, and heart valves. Besides, it also showed new
200 strong eGFP expressions in the hatching gland in embryos, pharyngeal arches,
201 chondrocytes cells in the fin and scattered epidermal serotonin-positive cells.
202 The positive eGFP cells were further confirmed by available antibodies (Fig.
203 4B). Interestingly, compared to the Pn line, eGFP fluroscence is pretty weak in

204 the brain and pharyngeal arches in the On line ($p<0.01$), and is hard to detect in
205 epidermal serotonin-positive cells and fin chondrocytes (Fig. 4C). Considering
206 that we obtained different F1 individuals of each line possess the same stable
207 eGFP expression patterns, and the eGFP expression patterns of different lines
208 are in consistent with the *in situ* hybridization and RNAseq results above, the
209 transgenic lines we constructed are stable and credible. These results indicate
210 that the cichlid fish-specific upstream region of *lg* indeed evolved new
211 regulatory functions, i.e. not only execute the ancestral expression profiles in
212 neurons, but also evolve new expressions in the hatching gland and pharyngeal
213 arches in cichlid fishes. Strong and new regulatory function in the scattered
214 epidermal serotonin-positive cells further evolved in the MH cichlid fish
215 lineage.
216
217 **Different regulatory roles of the cichlid fish-specific upstream regions are**
218 **simply due to three MH lineage-segregated SNPs**
219 To see whether the strong and new expression profiles in the Pn
220 transgenic line are caused by the three MH lineage segregated SNPs in the
221 upstream region, we constructed another transgenic line with the same upstream
222 region as in the Pn line but deleted these three SNPs (pT2AL-Pn-noSNPs-eGFP
223 line, Pn-noSNPs line) (Fig. 4C, Fig. S2). Interestingly, only with these three
224 SNPs deletion, Pn-noSNPs line showed clearly significant different expression
225 patterns with the Pn line, but similar patterns with the On line (Fig. 4C, Fig. 4D
226 and Table S5). These results confirmed that these SNPs are the key reason for
227 the different expression patterns between Pn line and On line.
228
229 **Discussion**
230
231 Above all, we unravelled the evolution of a fish-specific new gene, *lg*, for
232 the first time. No obvious phenotypic changes was detected in the *lg* KO
233 zebrafish. This can be due to the compensation by other genes in the
234 correponding gene network. This robust nature is not unexpected for immune

235 and brain network system [26,27], as the DE genes of the *lg* KO line are mainly
236 neurons and immunity related genes. Besides, lineage-specific new genes
237 initially usually located at the edge of the gene network [28], its disruption thus do
238 not affect the core gene network much. No matter what it is, mutation of *lg* does
239 not affect zebrafish survival. Together with its multiple losses in different fish
240 lineages during evolution, we propose that the ancestral function of *lg* is not
241 indispensable for fish survival.
242
243 Interestingly, *lg* evolved new regulatory functions in cichlid fishes,
244 especially in its most species-richness lineage, the MH lineage. For example, in
245 the hatching gland which is a key innovation to regulate the hatching period of
246 embryos to become free-living larvaes [29]; in pharyngeal cartilages, the well-
247 known tissues related to feeding and ecological niche exploration, are key to
248 cichlid fish adaptive radiation [30–32]; in scattered epidermal serotonin-positive
249 cells which are neuroepithelia cells related to environmental stimuli, such as
250 water flow, pheromone and hormone [33,34]; in the olfactory epithelium which is
251 well-known for detection of odorants, such as amino acid, hormone and
252 pheromone [35]. Considering that some of these tissues are derived from a key
253 vertebrate innovation, the neural crest cells [36], *lg* can be also causally linked to
254 neural crest cell development and differentiation. Together, *lg* gained new
255 regulatory functions in the most species-richness lineage, the MH lineage,
256 which can be causally linked to its adaptive radiation. Noticeably, although we
257 focused mainly on the regulatory function evolution of *lg* in cichlid fishes here,
258 the newly important function can be also evolved in their coding region.
259 Molecular evolutionary analyses have already shown that there is one amino
260 acid mutation segregated with the MH cichlid fish lineage and is under positive
261 selection, which might be causally linked to the adaptive radiation. Further KO
262 of *lg* in MH cichlid fishes will help test this hypothesis and unravel the
263 corresponding involved pathways. Besides, the evolution of *lg* gene function
264 can be also still on the way in cichlid fishes, considering that they are
265 evolutionarily young species [37].

266

267 Actually, the originally thought junk DNA in the genome, such as
268 transposable elements, non-coding regions and non-functional genes keep
269 giving us surprises. How can a new gene function be evovled and integrated
270 into the existing gene network? Here, we propose one hypothesis of the
271 functional evolution of *lg* that can be applied in a broad way (Fig. 5): the
272 relaxed selection pressure on the lineage-specific gene in the ancesral lineage
273 free it to accumulate mutations in both the regulatory region and the coding
274 region. The evolution of the regulatory region in *lg* offers immediate effects, i.e.
275 only with three SNPs mutations in the upstream region can provide new TFs
276 binding opportunities to build up new gene network connections to affect gene
277 expression profiles in the most species-richness MH lineage of cichlid fishes. It
278 will be interesting to test our predictions of the binding of a glucocorticoid
279 hormone related TF with these SNPs in the future. If this is true, it can connect
280 the organism with external environments efficiently, since glucocorticoid is a
281 well-known stress related hormone to regulate animal behavior to response to
282 the environment [38,39]. The *lg* gene we investigated here thus provides a good
283 example of new gene function evolution in an initially redundant lineage-
284 specific gene with respect to evo-devo in a broad way.

285

286 **Methods**

287

288 **sampling**

289 Samples of cichlid fishes for comparative genomic analyses were
290 provided by Prof. Walter Salzburger in University of Basel, Switzerland; cichlid
291 fish *Pundamila nyererei* for transgene plasmids construction was obtained from
292 IGV (Interessengemeinschaft Viktoriasee-Cichliden); Strains of AB (for
293 transgenic lines construction) and ABTü (for KO) zebrafish (*Danio rerio*) were
294 raised under standard condition (28.5°C, 14/10 h light/dark). Tilapia
295 (*Oreochromis niloticus*) were raised at 26°C, 14/10 h light/dark. Animal

296 experiments have been approved by the cantonal veterinary office in Basel
297 under permit number 2317 in University of Basel, Switzerland; the Committee
298 of Laboratory Animal Experimentation at Sun Yat-sen University and
299 Southwest University, China.
300
301 ***-omics analyses to characterize the evolution of lg across fish phylogeny***
302 To characterize the sequence evolution of *lg*, tblastx was conducted with
303 -*e* 0.001 in fishes with good genome assembly in Ensembl 100 (Table S1) and
304 local blast v2.9.0 (tblastx) was conducted with draft genomes assembly across
305 the phylogeny according to the reference [18]. After getting hits, the scaffold
306 possessing the corresponding sequence was extracted to predict the coding
307 regions with Augustus 3.3.3 [40,41]. The predicted *lg* coding sequences were
308 further confirmed by aligning the translated amino acids with the available
309 sequences from Ensembl 100 using Geneious 8.0.5 [42].
310 .
311 ***-omics analyses to characterize the evolution of lg within cichlid fishes***
312 To further characterize the sequence evolution of *lg* within East African
313 cichlid fishes, we amplified the whole gene sequence and its upstream region
314 (ca.10 kb in total) in different cichlid fishes. Combined with available cichlid
315 fish genomes, we thus studied the sequence evolution in a total of 20 cichlid
316 fishes representing main lineages across the phylogeny [24]. Genomic DNA was
317 extracted using DNeasy Blood&Tissue kit (Qiagen). PCR primers were
318 designed using Primer premier 5 (all the primers in this study were designed
319 with this software) [43]. High-Fidelity PCR Master Mix (BioLabs) and a touch-
320 down annealing process were used for the amplification. PGM Ion Torrent was
321 used to sequence long amplicons (400 bp) with CHIP 316 (Life Technologies).
322 Galaxy v20.09 [44] was used to trim low quality reads (window size 2, step size 1,
323 minimum quality score 20) and filter short reads with minimum size 40 bp. The
324 sequence of a cichlid fish, tilapia, retrieved from Ensembl database was used as

325 the reference to recheck the assembled sequences. The sequencing reads of *lg*
326 gene across cichlid fish phylogeny are available at NCBI (BioProject ID
327 PRJNA611690). The ion torrent sequencing of most species were successful
328 except *Limnochromis abeelei*, *Ophthalmotilapia ventralis*, and *Bathybates*
329 *graueri*, whose sequences were further obtained with Sanger sequencing
330 (Primers can be found in Table S6).
331
332 Genomic alignment of *lg* of cichlid fishes and other teleost fishes was
333 conducted with mVISTA [45,46] using the LAGAN alignment tool [47] with *P.*
334 *nyererei* as the reference. We applied the repeat masking option with fugu as the
335 reference. Transcription factors (TFs) binding prediction analyses were
336 conducted with MatInspector v8.4.1 with genomatix software suite
337 (https://www.genomatix.de) using default parameters with zebrafish as the reference
338 [48,49].
339
340 **Positive selection detection of *lg* in cichlid fishes**
341 To detect whether positive selection occurred in the evolution of *lg* in a
342 specific branch, a branch-site model within codeml in PAML v3.14 was
343 performed [50,51]. Fixed branch lengths (fix_blength = 2) retrieved under the M0
344 model were used. Alignment gaps and ambiguity characters were removed
345 (Cleandata = 1). The tree for PAML test is based on the phylogeny from the
346 study [24]. Rates of non-synonymous to synonymous substitutions (ω or dN/dS)
347 significantly larger than 1 represent that the sites in a specific target branch
348 (foreground branch) are under positive selection. A *chi*-square approximation
349 was calculated to test statistical significance. Detail statistics of branch-site
350 model comparisons can be found in Table S4.
351
352 **CRISPR/Cas9 gene KO in zebrafish**

353 The *lg* mutant line was generated by CRISPR/Cas9 system. The guide
354 RNA (gRNA) target site was selected using the CRISPOR v4.98 [52] with 20 bp
355 NGG PAM motif. The synthesized gRNA and Cas9 mRNA were co-injected
356 into the one-cell stage embryos of zebrafish at the concentration of 50 ng/ul.
357 The randomly picked 10 embryos at 1 dpf (days post fertilization) were lysated
358 and used to access functional gene knockout using PCR amplification with
359 primers designed at the target KO site (Table S6). Heterozygous F1 fish were
360 obtained by crossing F0 and wild-type (WT) zebrafish and were genotyped with
361 the same primers above using Sanger sequencing at the target site (Table S6).
362 Siblings of F1 generation harboring the same mutations were mated to generate
363 homozygous F2 mutants (MUT) and WT sisters (as the control).
364
365 **Western blot**
366 Since there is no available antibody for *lg*, we thus added a HA tag
367 (ATGTACCCATACGACGTCCCAGACTACGCT) in front of the coding
368 region of *lg* by PCR amplification (Table S6). The antibody of HA tag can thus
369 detect the fused *lg* protein if there is no disruption. To ligate the amplicons of
370 HA tagged coding regions of WT and MUT into pCS2 vector, the restriction
371 enzyme sites (*Bam*H I and *Xba* I) were added at the two ends of the amplicons
372 by PCR amplification (Table S6), followed by double digestion-system (NEB)
373 and ligation (Promega) (Table S6). The plasmids of pCS2, pCS2-lg-HAtag-WT
374 and pCS2-lg-HAtag-MUT (2 μg) (Fig. S2) were then transfected into HEK293T
375 cell line (ab255449, abcam), seperately. Protein was collected after 36 h
376 transfection to perform Western blot according to the standard procedure [53].
377 Antibodies of mouse anti-HA (901501 Biolegend, 1:2000) and mouse anti-
378 GAPDH (60004-1-lg Proteintech, 1:2000) were used for the detection.
379 Detections were performed with goat anti-mouse secondary antibody (31430
380 Thermo 1:5000) and enhanced chemiluminescence (ECL plus) reagent.
381

382 **RNAseq analyses**

383 To see gene expression profiles of *lg*, the available transcriptomic raw
384 data of different tissues in different fishes were extracted from SRA database
385 (Table S2). The raw reads were mapped to the corresponding transcriptome
386 retrieved from Ensembl using Hisat2-2.1.0 [54] with default parameters. The
387 reported SAM results were converted into BAM format, and finally were
388 recorded as count files using SAMtools v1.10 [55] followed by TPM values
389 calculation [56].

390

391 To see the affected genes after knockout, comparative transcriptomic
392 analyses were conducted between MUT and WT (Table S3). RNA extraction
393 was performed (each replicate pooled 30 individuals, six biological replicates
394 for each group) using TRIzol® reagent (Invitrogen). Six biological replicates of
395 WT and MUT of F2 generation were used for RNAseq analyses. Library
396 construction and sequencing were performed using Illumina TrueSeq RNAseq
397 technology at Novogene Company, China, with paired end reads length of 350
398 bp. The raw reads were mapped to the zebrafish transcriptome assembly using
399 the same methods mentioned above to get the reads count files. The differential
400 expressed (DE) transcripts between WT and MUT were further analyzed using
401 edgeR v3.20.7 (FDR<0.01). The raw data were deposited at NCBI BioProject
402 ID PRJNA611476. The top up- and down- 10 DE transcripts list can be found in
403 Table S3.

404

405 ***in situ* hybridization**

406 The gene fragments in the coding region of *lg* in zebrafish and tilapia
407 were amplified using PrimerSTAR (Takara) (Table S6). *lyz* probe synthesis was
408 conducted as described in the reference [57]. The amplified products were cloned
409 into pGEM®-T Easy vector (Promega). Plasmids were extracted and purified
410 using TIANprep Mini Plasmid Kit II (Tiangen). The insertion and direction of

411 probes were confirmed with Sanger sequencing using M13 PCR products
412 (Tsingke). RNA probes were synthesized with the DIG RNA labeling kit
413 (SP6/T7) (Roche). The experiment was conducted as described in the reference
414 [58] with biological replicates (n=30 for zebrafish, n=20 for tilapia, and n=10 for
415 *P. nyererei*). The embryos or larvae were mounted in 1% low melting point
416 Agarose (GBCBIO) and imaged with Zeiss Axio Imager Z2 microscopy.
417
418 **Transgenic plasmids and transgenic lines construction**
419 To see the regulatory functions of cichlid fish-specific upstream regions,
420 we constructed plasmids with different upstream regions as promoters to drive
421 eGFP expression in zebrafish (Fig. 4A and Fig. S2), i.e. 1) species in the MH
422 lineage (*P. nyererei*) (pT2AL-Pn-eGFP line); 2) the corresponding region of
423 species at the sister lineage of East African Great Lakes cichlid fishes (tilapia)
424 (pT2AL-On-eGFP line). 3) To see whether the three MH-segregated SNPs in
425 the cichlid fish-specific upstream region are important, we specifically deleted
426 these SNPs using site-directed mutagenesis by PCR primer extension (Table S6)
427 [59] to construct another transgenic line (pT2AL-Pn-noSNPs-eGFP line).
428 Schematic Plasmids construction can be found in Fig. S2.
429
430 DNA fragments for different plasmids construction were amplified with
431 PCR primers added with corresponding restriction endonuclease sites (Fig. S2,
432 Table S6). Fragments were then recovered (Tiangen), digested (NEB) and
433 ligated (Promega) into pT2AL vector following the instructions accordingly.
434 Successfully constructed recombinant plasmids were further confirmed by
435 Sanger sequencing (Tsingke). Injections were performed into one-cell stage of
436 zebrafish embryos. By outcrossing with WT zebrafish, we successfully obtained
437 F2 stable transgenic lines for all these lines (Pn line, On line, Pn-noSNPs line).
438 The fluroscent signals were imaged with Zeiss 880 confocal microscopy.
439

440 **Immunohistochemistry (IHC)**

441 The following primary antibodies were used: mouse anti-Pvalb7, 1:1000
442 provided by Prof. Masahiko Hibi, Nagoya University, Japan [60]; rabbit anti-
443 Serotonin, 1:400 (s5545, Sigma); rabbit anti-Calretinin, 1:400 (ab702, abcam);
444 goat anti-GFP, 1:400 (ab6673, abcam); mouse anti-HuC/HuD, 1:1000 (16A11,
445 Invitrogen); rabbit anti-SOX10, 1:400 (ab229331, abcam); rabbit anti-PAX6
446 1:400 (ab209836, abcam); mouse anti-Collagen type II (II-II6B3, DSHB),
447 rabbit anti-Cathepsin L/MEP (ab209708, abcam). Secondary antibodies are
448 donkey to goat IgG H&L (Alexa Fluor® 488) (ab150133, abcam), donkey to
449 mouse IgG H&L (Alexa Fluor® 647) (ab150107, abcam), donkey to rabbit IgG
450 H&L (Alexa Fluor® 647) (ab150063, abcam) and DAPI (ab228549, abcam).
451 Fish larvae were fixed in 4% Paraformaldehyde (Sigma) for 4 h at room
452 temperature, followed by detail protocol as described before [61]. The fluroscent
453 signals were imaged with Zeiss 880 confocal microscopy.

454

455 **Immune response detection**

456 To see whether there are different immune responses between MUT and
457 WT zebrafish, we injected lipopolysaccharide (LPS), the key inflammatory
458 components of Gram-negative bacteria into the circulation of 2 dpf zebrafish
459 and counted the numbers of *lyz*$^{+}$ neutrophils in the caudal hematopoietic tissue
460 at one day post injection. The *lyz*$^{+}$ whole-mount RNA *in situ* hybridization
461 signals across eight somites of the larval CHT (caudal hematopoietic tissue)
462 regions were manually counted. Statistics analyses can be found in Table S7.

463

464 **Data Availability**

465

466 The publicly available datasets that this study used are provided in the
467 supplementary information. The RNAseq data of WT and MUT are available at
468 NCBI (BioProject ID PRJNA611476). The ion torrent sequencing reads of *lg*
469 gene are available at NCBI (BioProject ID PRJNA611690). The scaffolds

470 possess *lg* of teleost fishes across the phylogeny, the assembly of *lg* including
471 upstream region of cichlid fishes across the phylogeny retrieved from ion
472 torrent sequencing and sanger sequencing can be found in Dryad
473 https://doi.org/10.5061/dryad.fxpnvx0qb. Source associated data of Fig. 3 and Fig. 4
474 are provided in the supplementary files. Other data are available upon
475 reasonable request.
476

477 **References**

478

479 1. Dobzhansky, T. Nothing in biology makes sense except in the light of
480 evolution. *Am. Biol. Teach.* **March,** 125–129 (1973).
481 2. Tautz, D. & Domazet-Lošo, T. The evolutionary origin of orphan genes.
482 *Nat. Rev. Genet.* **12,** 692–702 (2011).
483 3. Bae, B.-I., Jayaraman, D. & Walsh, C. A. Genetic changes shaping the
484 human brain. *Dev. Cell* **32,** 423–34 (2015).
485 4. Zhang, Y. E. & Long, M. New genes contribute to genetic and phenotypic
486 novelties in human evolution. *Curr. Opin. Genet. Dev.* **29,** 90–6 (2014).
487 5. Babonis, L. S., Martindale, M. Q. & Ryan, J. F. Do novel genes drive
488 morphological novelty? An investigation of the nematosomes in the sea
489 anemone Nematostella vectensis. *BMC Evol. Biol.* **16,** 114 (2016).
490 6. Santos, M. E., Le Bouquin, A., Crumière, A. J. J. & Khila, A. Taxon-
491 restricted genes at the origin of a novel trait allowing access to a new
492 environment. *Science (80-.).* **358,** 386–390 (2017).
493 7. Khalturin, K. *et al.* A novel gene family controls species-specific
494 morphological traits in Hydra. *PLoS Biol.* **6,** e278 (2008).
495 8. Garcia-Fernàndez, J. The genesis and evolution of homeobox gene
496 clusters. *Nat. Rev. Genet.* **6,** 881–92 (2005).
497 9. Wu, M., Chen, G. & Li, Y.-P. TGF-β and BMP signaling in osteoblast,
498 skeletal development, and bone formation, homeostasis and disease. *Bone*
499 *Res.* **4,** 16009 (2016).
500 10. Wagner, G. P. & Lynch, V. J. The gene regulatory logic of transcription
501 factor evolution. *Trends Ecol. Evol.* **23,** 377–85 (2008).
502 11. Wray, G. A. The evolutionary significance of cis-regulatory mutations.
503 *Nat. Rev. Genet.* **8,** 206–216 (2007).
504 12. Zhang, J. Parallel adaptive origins of digestive RNases in Asian and
505 African leaf monkeys. *Nat. Genet.* **38,** 819–23 (2006).
506 13. Hoekstra, H. E., Hirschmann, R. J., Bundey, R. A., Insel, P. A. &
507 Crossland, J. P. A single amino acid mutation contributes to adaptive
508 beach mouse color pattern. *Science* **313,** 101–4 (2006).

14. Chen, L., DeVries, A. L. & Cheng, C.-H. C. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proc. Natl. Acad. Sci.* **94,** 3811–3816 (1997).
15. Carroll, S. B. Evolution at Two Levels: On Genes and Form. *PLoS Biol.* **3,** e245 (2005).
16. Hu, W., Chen, S., Zhang, R. & Lin, Y. Single nucleotide polymorphisms in the upstream regulatory region alter the expression of myostatin. *Vitr. Cell. Dev. Biol. - Anim.* **49,** 417–423 (2013).
17. Marceca, C., Pfob, M., Schelling, G., Steinlein, O. K. & Eggert, M. Single nucleotide polymorphism creating a variable upstream open reading frame regulates glucocorticoid receptor expression. *Gene* **563,** 24–28 (2015).
18. Malmstrøm, M. *et al.* Evolution of the immune system influences speciation rates in teleost fishes. *Nat. Genet.* **48,** 1204–1210 (2016).
19. Salzburger, W. The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol. Ecol.* **18,** 169–85 (2009).
20. SH, O. & LI, Z. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* **132,** 631–44 (2008).
21. Kim, C.-H. *et al.* Zebrafish elav/HuC homologue as a very early neuronal marker. *Neurosci. Lett.* **216,** 109–112 (1996).
22. Hugo, S. E. *et al.* A monocarboxylate transporter required for hepatocyte secretion of ketone bodies during fasting. *Genes Dev.* **26,** 282–93 (2012).
23. Salzburger, W. Understanding explosive diversification through cichlid fish genomics. *Nat. Rev. Genet.* (2018). doi:10.1038/s41576-018-0043-9
24. Meyer, B. S., Matschiner, M. & Salzburger, W. A tribal level phylogeny of Lake Tanganyika cichlid fishes based on a genomic multi-marker approach. *Mol. Phylogenet. Evol.* **83,** 56–71 (2015).
25. Salzburger, W., Mack, T., Verheyen, E. & Meyer, A. Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evol. Biol.* **5,** 17 (2005).
26. Hubel, P. *et al.* A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. *Nat. Immunol.* **20,** 493–502 (2019).
27. Denève, S., Alemi, A. & Bourdoukan, R. The Brain as an Efficient and Robust Adaptive Learner. *Neuron* **94,** 969–977 (2017).
28. Zhang, W., Landback, P., Gschwend, A. R., Shen, B. & Long, M. New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome Biol.* **16,** 202 (2015).
29. Nagasawa, T. *et al.* Evolutionary Changes in the Developmental Origin of Hatching Gland Cells in Basal Ray-Finned Fishes. *Zoolog. Sci.* **33,** 272 (2016).
30. Conith, A. J., Kidd, M. R., Kocher, T. D. & Albertson, R. C. Ecomorphological divergence and habitat lability in the context of robust patterns of modularity in the cichlid feeding apparatus. *BMC Evol. Biol.*

552 **20,** 95 (2020).
553 31. Roberts, R. B., Hu, Y., Albertson, R. C. & Kocher, T. D. Craniofacial
554 divergence and ongoing adaptation via the hedgehog pathway. *Proc. Natl.*
555 *Acad. Sci. U. S. A.* **108,** 13194–9 (2011).
556 32. Albertson, R. C., Streelman, J. T., Kocher, T. D. & Yelick, P. C.
557 Integration and evolution of the cichlid mandible: the molecular basis of
558 alternate feeding strategies. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 16287–92
559 (2005).
560 33. König, D. *et al.* Distribution and Restoration of Serotonin-
561 Immunoreactive Paraneuronal Cells During Caudal Fin Regeneration in
562 Zebrafish. *Front. Mol. Neurosci.* **12,** 227 (2019).
563 34. Coccimiglio, M. L. & Jonz, M. G. Serotonergic neuroepithelial cells of the
564 skin in developing zebrafish: morphology, innervation and oxygen-
565 sensitive properties. *J. Exp. Biol.* **215,** 3881–3894 (2012).
566 35. Kermen, F., Franco, L. M., Wyatt, C. & Yaksi, E. Neural circuits
567 mediating olfactory-driven behavior in fish. *Front. Neural Circuits* **7,** 62
568 (2013).
569 36. Green, S. A., Simoes-Costa, M. & Bronner, M. E. Evolution of vertebrates
570 as viewed from the crest. *Nature* **520,** 474–82 (2015).
571 37. Salzburger, W. Understanding explosive diversification through cichlid
572 fish genomics. *Nat. Rev. Genet.* **19,** 705–717 (2018).
573 38. Nyman, C., Fischer, S., Aubin-Horth, N. & Taborsky, B. Evolutionary
574 conserved neural signature of early life stress affects animal social
575 competence. *Proceedings. Biol. Sci.* **285,** (2018).
576 39. Culbert, B. M., Gilmour, K. M. & Balshine, S. Stress axis regulation
577 during social ascension in a group-living cichlid fish. *Horm. Behav.* **103,**
578 121–128 (2018).
579 40. Stanke, M. & Morgenstern, B. AUGUSTUS: a web server for gene
580 prediction in eukaryotes that allows user-defined constraints. *Nucleic*
581 *Acids Res.* **33,** W465–7 (2005).
582 41. Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B. AUGUSTUS: a
583 web server for gene finding in eukaryotes. *Nucleic Acids Res.* **32,** W309–
584 12 (2004).
585 42. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop
586 software platform for the organization and analysis of sequence data.
587 *Bioinformatics* **28,** 1647–9 (2012).
588 43. Lalitha, S. Primer Premier 5. *Biotech Software&Internet Rep.* **1,** (2000).
589 44. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and
590 collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46,**
591 W537–W544 (2018).
592 45. Mayor, C. *et al.* VISTA : visualizing global DNA sequence alignments of
593 arbitrary length. *Bioinformatics* **16,** 1046–1047 (2000).
594 46. Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. & Dubchak, I.

595 VISTA: computational tools for comparative genomics. *Nucleic Acids*
596 *Res.* **32,** W273–9 (2004).
597 47. Brudno, M. *et al.* LAGAN and Multi-LAGAN: efficient tools for large-
598 scale multiple alignment of genomic DNA. *Genome Res.* **13,** 721–31
599 (2003).
600 48. Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. MatInd and
601 MatInspector: new fast and versatile tools for detection of consensus
602 matches in nucleotide sequence data. *Nucleic Acids Res.* **23,** 4878–84
603 (1995).
604 49. Cartharius, K. *et al.* MatInspector and beyond: promoter analysis based on
605 transcription factor binding sites. *Bioinformatics* **21,** 2933–2942 (2005).
606 50. Yang, Z. PAML: a program package for phylogenetic analysis by
607 maximum likelihood. *Comput. Appl. Biosci.* **13,** 555–6 (1997).
608 51. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol.*
609 *Biol. Evol.* **24,** 1586–91 (2007).
610 52. Concordet, J.-P. & Haeussler, M. CRISPOR: intuitive guide selection for
611 CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids*
612 *Res.* **46,** W242–W245 (2018).
613 53. Zhao, F. *et al.* Irf8 regulates the progression of myeloproliferative
614 neoplasm-like syndrome via Mertk signaling in zebrafish. *Leukemia* **32,**
615 149–158 (2018).
616 54. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner
617 with low memory requirements. *Nat. Methods* **12,** 357–360 (2015).
618 55. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools.
619 *Bioinformatics* **25,** 2078–9 (2009).
620 56. Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance
621 using RNA-seq data: RPKM measure is inconsistent among samples.
622 *Theory Biosci.* **131,** 281–285 (2012).
623 57. Hou, Y. *et al.* Systemic inoculation of Escherichia coli causes emergency
624 myelopoiesis in zebrafish larval caudal hematopoietic tissue. *Sci. Rep.* **6,**
625 36853 (2016).
626 58. Westerfield, M. *The zebrafish book. A guide for the laboratory use of*
627 *zebrafish (Danio rerio). 4th edition.* (University of Oregon Press, Eugene,
628 2000).
629 59. Reikofski, J. & Tao, B. Y. Polymerase chain reaction (PCR) techniques
630 for site-directed mutagenesis. *Biotechnol. Adv.* **10,** 535–47 (1992).
631 60. Nimura, T. *et al.* Role of Reelin in cell positioning in the cerebellum and
632 the cerebellum-like structure in zebrafish. *Dev. Biol.* **455,** 393–408 (2019).
633 61. Barresi, M. J. F., Stickney, H. L. & Devoto, S. H. The zebrafish slow-
634 muscle-omitted gene product is required for Hedgehog signal transduction
635 and the development of slow muscle identity. *Development* **127,** 2189–
636 2199 (2000).
637

638 **Acknowledgements**

639

640 Many thanks to Prof. Walter Salzburger, Attila Rüegg, Zhen Liu,
641 Zhiqiang Zhou and Hesheng Xiao for the sampling and suggestions. We thank
642 Prof. Masahiko Hibi and Prof. Kiyoshi Naruse for providing the antibody
643 Pvalb7. We also thank Prof. Lingfei Luo, Prof. Hua Ruan, Prof. Yang Liu and
644 Prof. Tienming Lee for the laboratory supports. Many thanks to Prof. Zuzana
645 Musilova, Dr. Xiaogui Yi, Dr. Yandong Zhan, Dr. Fangying Zhao, Yuepeng He,
646 Xinkai Cheng, Qing Li and Yue Gao for the supports. Thanks to Huansheng
647 Tao and Jialing Xu for the Confocol microscopy support, Zhao Long for
648 fluroscent microscopy support. This study was funded by grants from the
649 National Natural Science Foundation of China (31802283), Guangdong Basic
650 and Applied Basic Research Foundation (2020A1515010330); the Fundamental
651 Research Funds for the Central Universities (20lgpy109); China Postdoctoral
652 Science Foundation (2018 M633311) and Doctoral Scholarship of University of
653 Basel, Switzerland to LG; Grants from the National Natural Science Foundation
654 of China (31630082) to DW, and (31771623 and 31571500) to LL. Part of -
655 omics analyses was supported by National Supercomputer Center in Guangzhou,
656 China.

657

658 **Authors contributions**

659

660 LG, DW and LL conceived this work. LG carried out –omics sequencing
661 and evolutionary analyses. LG, CL, XM and ZW conducted the *in situ*
662 hybridization. LG carried out the IHC. XH and WZ participated the IHC
663 experiments of serotonin-positive cells and olfactory epithelium cells. LG, CL
664 and XM constructed plasmids of the Pn line and the On line. LG constructed
665 plasmids of the Pn-noSNPs line. LG leads the transgenic lines and mutant lines
666 construction with the help from CL and XM. LG took the confocal photos. YH

667 performed the western-blot experiment. ZW carried out the LPS immune
668 response experiments. LG conducted the fasting ORO staining experiments. LG
669 wrote the initial manuscript draft, and all authors commented on the manuscript
670 and approved the final version.
671
672 **Competing interests**: The authors declare no competing interests.
673
674 **Figure legends**
675
676 **Fig.1 Sequence evolution of *lg*.** a. Comparative genomic analyses showed that
677 *lg* is a fish-specific gene. b. Multiple losses occurred in different lineages within
678 teleost fishes. The reference tree is from [18].
679
680 **Fig.2 Expression profiles of *lg* in fishes.** a. Transcriptomic analyses of *lg* gene
681 expression patterns in adult fish tissues. *lg* mainly expressed in the brain and
682 eyes in most fishes, but evolved new expression patterns in the skin in the
683 modern haplochromine (MH) lineage of cichlid fishes. b. *in situ* hybridization in
684 zebrafish and cichlid fish larvae. *lg* mainly expressed in the brain, ganglion and
685 inner plexiform cell layers in the eyes of zebrafish. No expression was detected
686 in the hatching gland of 24 hpf (hours post fertilization) embryos of zebrafish.
687 *lg* highly expressed in the brain, eyes and pharyngeal arches in the
688 haplochromine cichlid fish, *Pundamilia nyererei*, but weakly expressed in the
689 cichlid fish, *Oreochromis niloticus*.
690
691 **Fig.3 *lg* gene knockout in zebrafish and comparisons between mutant**
692 **(MUT) and wild-type (WT) zebrafish.** A. *lg* is successfully disrupted using
693 CRISPR/Cas9. This is confirmed by both Sanger sequencing (up) and Western-
694 blot (down). B. Immune responses comparisons between MUT and WT using
695 whole mount *in situ* hybridization in the CHT (caudal hematopoietic tissue)

696 treated with LPS and PBS (n=10 for WT, and n=18 for MUT), respectively.
697 Quantification of the numbers of *lyz*$^{+}$ signals across eight somites were shown
698 in dots. n.s. not significant. Mean and SD error bar were presented. C.
699 phenotype comparisons between MUT and WT zebrafish. Brain and eye
700 phenotypes in mutants were examined with neuron marker HuC. No obvious
701 phenotype was found. There is no obvious mutant phenotype with ORO
702 staining under fasting condition, either, which detect energy metabolism [22].
703
704 **Fig.4 *lg* sequence evolution and regulatory function evolution in cichlid**
705 **fishes.** Aa. Both upstream region and coding region exhibited SNPs segregated
706 with the modern haplochromine (MH) lineage. i.e. three SNPs (GTA) in the
707 upstream region and one nonsynonymous amino acid change (A to S) located at
708 the 38th position which is under positive selection (Table S4). Transcription
709 factors binding sites prediction showed that these SNPs can set up the
710 opportunity to bind to glucocorticoid modulatory element binding protein 2
711 (GMEB2). The reference tree is from [24]. Ab. Three types of transgenic plasmids
712 construction. The pT2AL-On-eGFP line (On line) adopted the upstream region
713 of *O. niloticus* as the promoter to drive eGFP expression. The pT2AL-Pn-eGFP
714 line (Pn line) used the homologous upstream region of *P. nyererei* in the MH
715 lineage as the promoter. The pT2AL-Pn-no-SNPs-eGFP line (Pn-noSNPs line)
716 used the same promoter as the Pn line but deleted the three MH-segregated
717 SNPs (GTA). B. The fluorescence images in the Pn line. Merged figures of
718 double staining of eGFP and cell marker antibodies were shown if there are
719 available antibodies. Red is from the cell marker antibody. Green is from eGFP.
720 Blue is from DAPI. Pictures were taken from 4 dpf (days post fertilization)
721 zebrafish except the hatching gland which is from 24 hpf (hours post
722 fertilization) embryos. C. Different expression patterns of three transgenic lines
723 in zebrafish. D. Statistics of expression patterns comparisons among three
724 transgenic lines, focusing on cerebellum, hindbrain, and the first two pharyngeal

725 arches. The Pn line showed significantly stronger fluroscence than the other two
726 lines. With the three MH-segregated SNPs deletion, Pn-noSNPs line showed
727 weak fluroscent expressions, similar to the On line. Mean and SD error bars
728 were presented. n=6 for each measurement. Detail statistical analyses can be
729 found in Table S5.
730
731 **Fig.5 Hypothesis of *lg* sequence and function evolution with respect to evo-**
732 **devo.** *lg* is a fish-specific gene. Its ancestral function is not indispensable.
733 Interestingly, *lg* evolves a cichlid fish-specific upstream region, thus gained new
734 regulatory functions to drive gene expression in new tissues related to
735 adaptation, such as in pharyngeal cartilages and the hatching gland in cichlid
736 fishes. Especially in the most species-richness lineage, the modern
737 haplochromine (MH) lineage, the upstream of which not only drive stronger
738 eGFP expression patterns compared to the homologous region of tilapia, but
739 also evolved new profiles again, in epidermal serotonin-positive cells. And this
740 can be simply due to the three MH-segregated SNPs in the upstream region
741 which are predicted to bind to a hormone related transcription factor, the
742 glucocorticoid modulatory element binding protein 2 (GMEB2). If this is true, it
743 opens a new opportunity to response to the external environemnt. Besides, the
744 coding region of lg also possesses one MH- segregated amino acid mutation
745 which is under positive selection, suggesting that new function might also
746 evolved in the coding region which deserve further investigation. lg thus
747 provides a good example of new gene function emergence in a redundant
748 lineage-specific gene with respect to evo-devo.
749
750
751
752
753

Figures

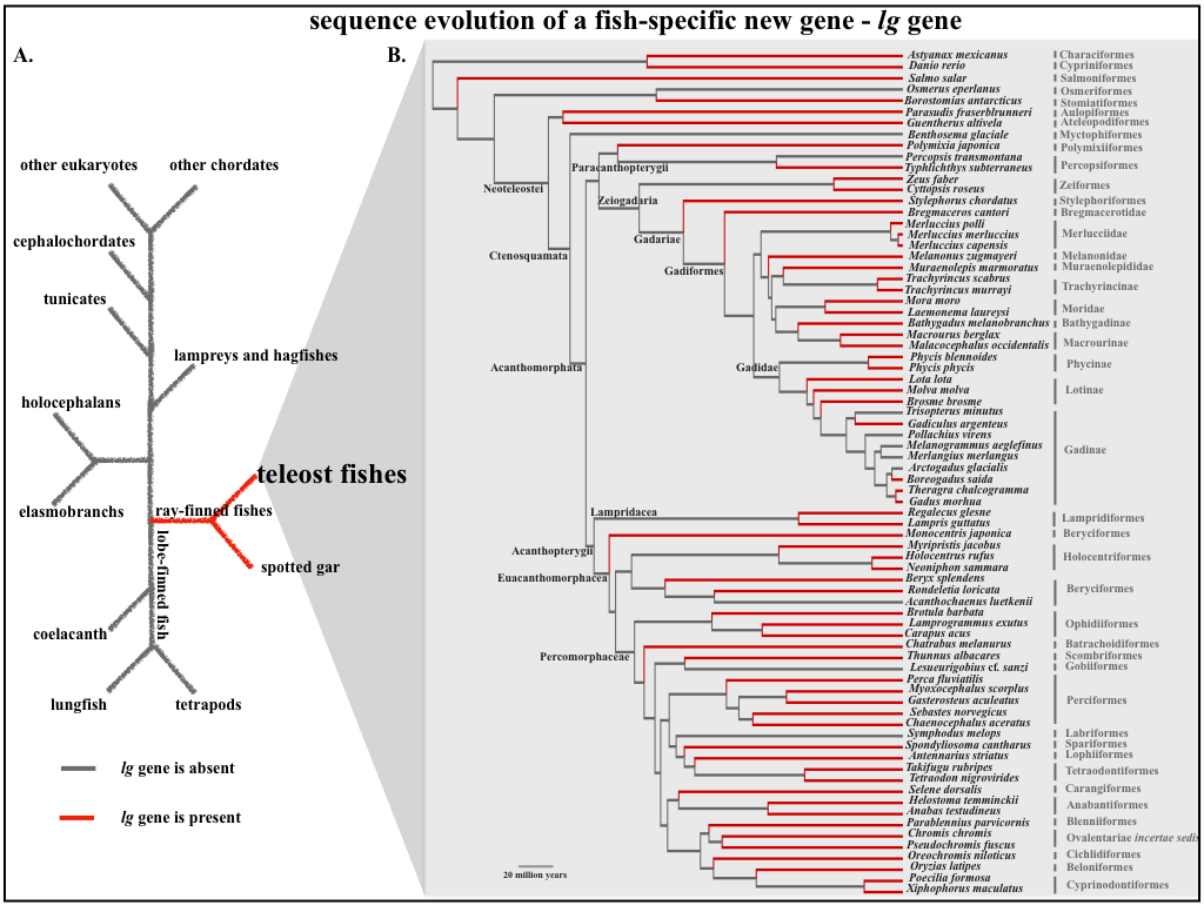

Fig.1 Sequence evolution of *lg*.

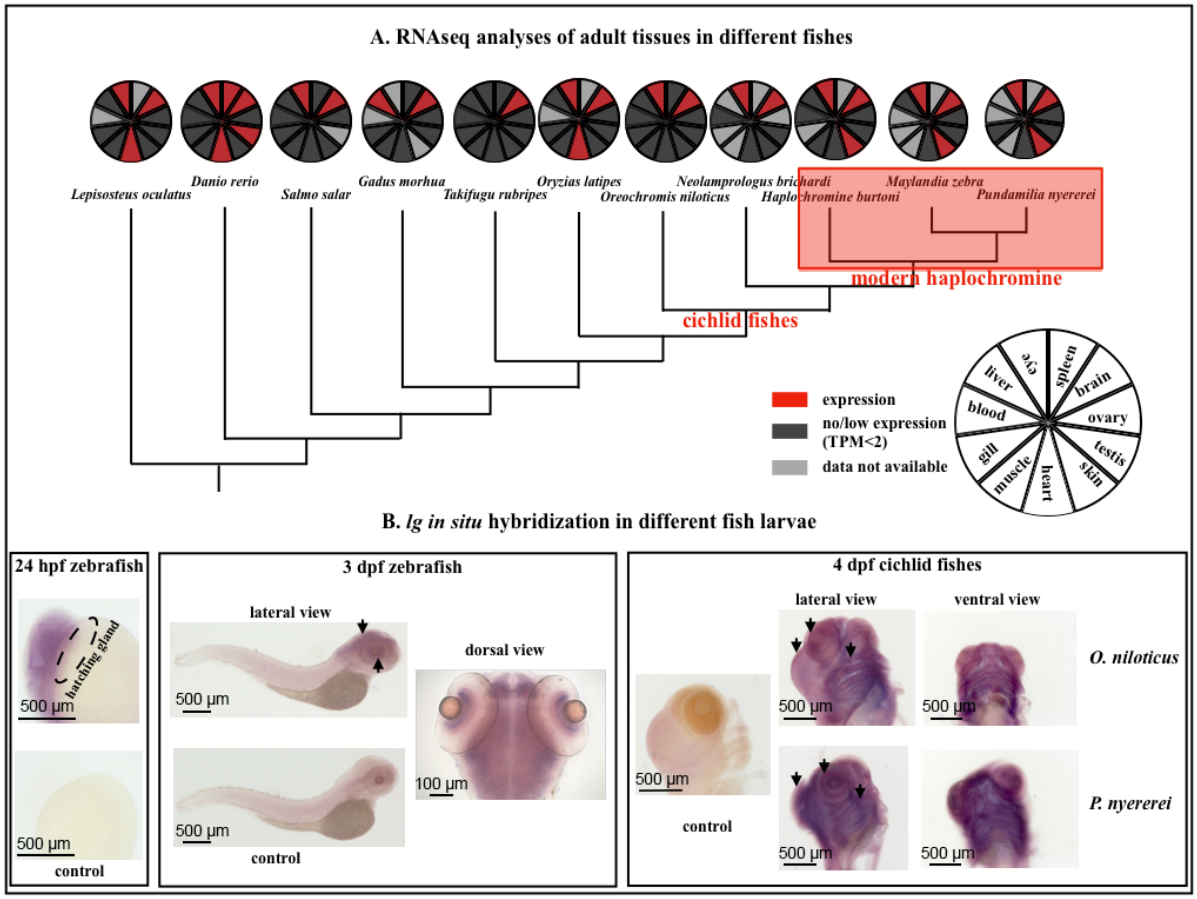

Fig.2 Expression profiles of *lg* in fishes.

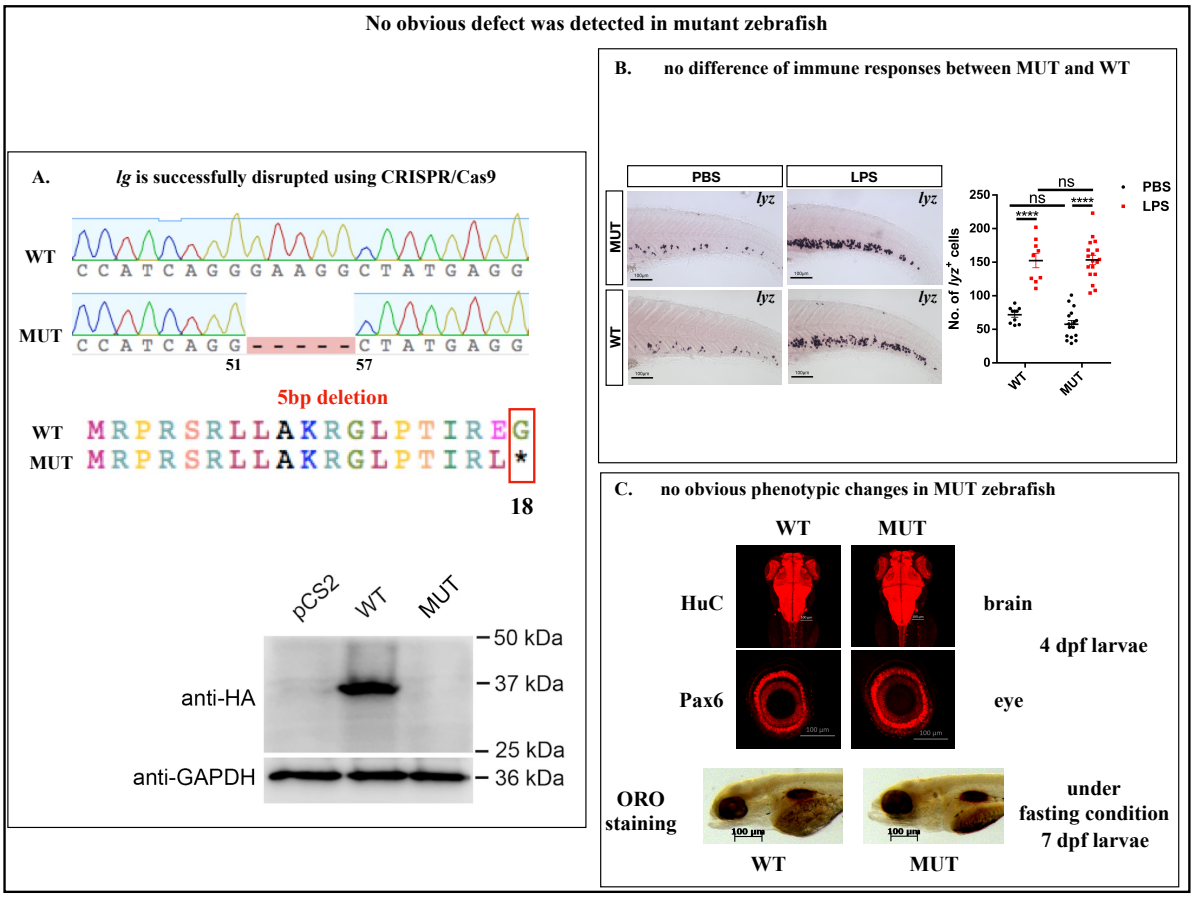

Fig.3 *lg* gene knockout in zebrafish and comparisons between mutant (MUT) and wild-type (WT) zebrafish.

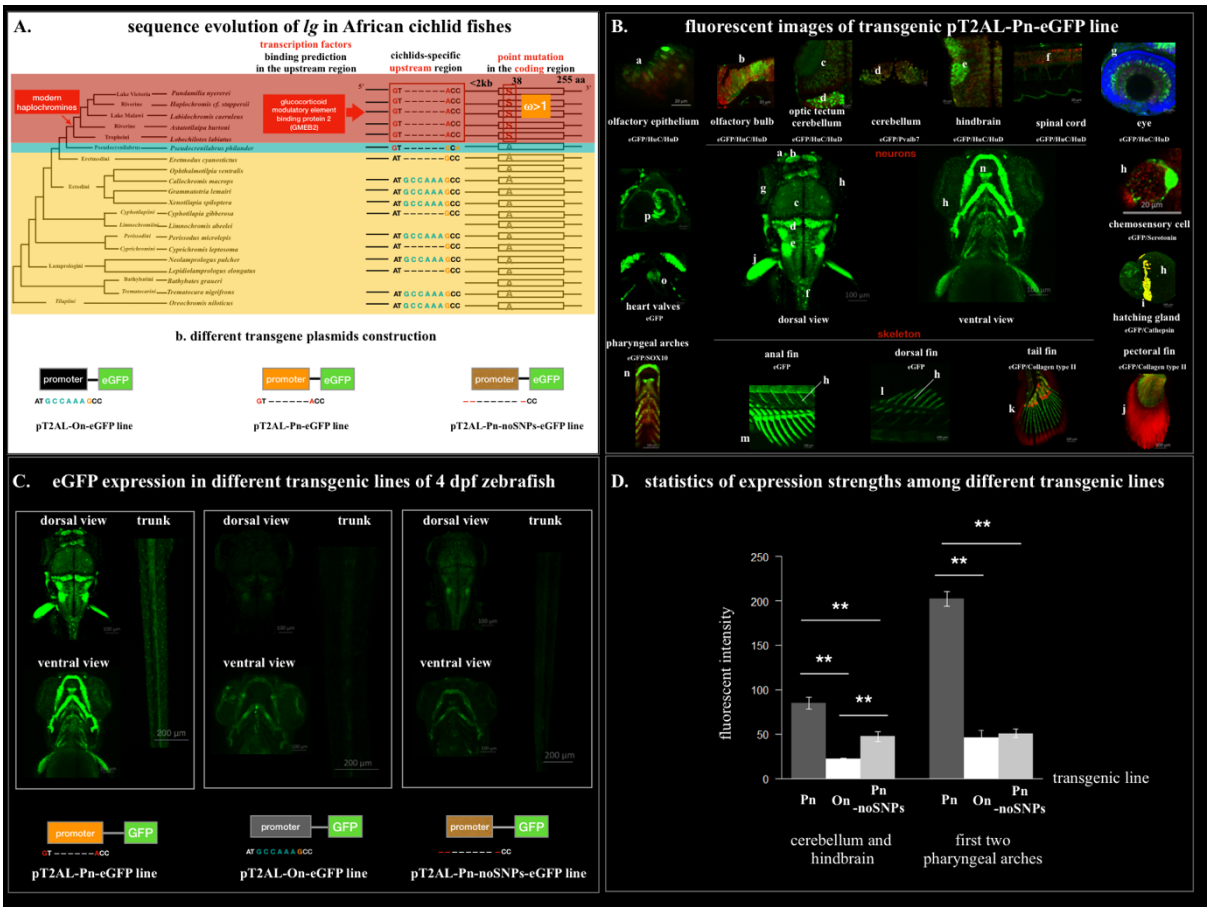

Fig.4 *lg* sequence evolution and regulatory function evolution in cichlid fishes.

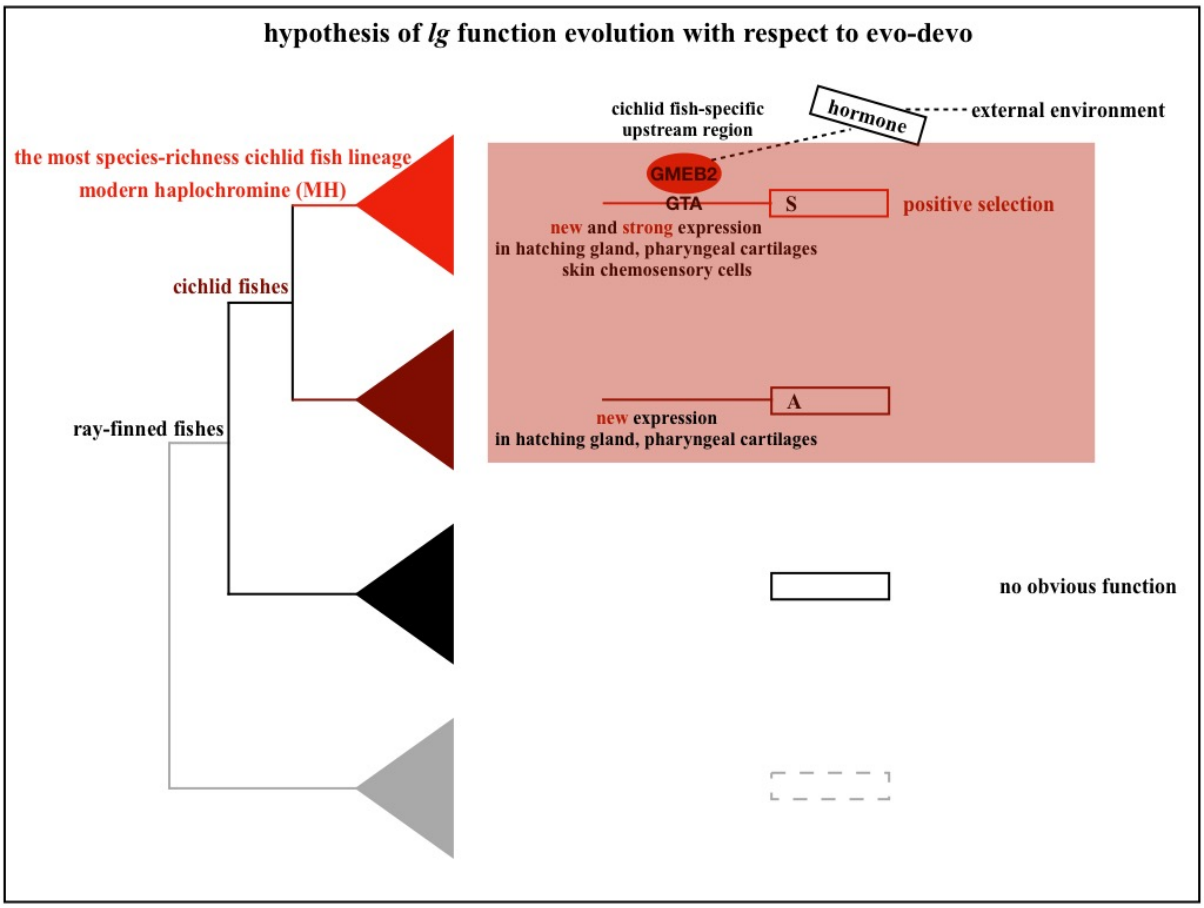

Fig.5 Hypothesis of *lg* sequence and function evolution with respect to evo-devo.

Supplementary information

Table S1 *lg* appearance in species with available genomes in Ensembl

Table S2 RNAseq analyses of *lg* in different fishes with available data

Table S3 Top differential expressed transcripts between MUT and WT zebrafish

Table S4 Positive selection detection of *lg* in cichlid fishes within codeml in PAML

Table S5 eGFP fluroscent strengths comparisons among different transgenic lines

Table S6 PCR primers and annealing temperatures

780 Table S7 The numbers of LPS immune response signals comparisons between
781 WT and MUT zebrafish
782 Supplementary Figures