The evolutionary emergence of the *rac3b/rfng/s.gca* regulatory cluster refined mechanisms for hindbrain boundaries formation

Joaquin Letelier, Javier Terriente, Ivan Belzunce, Adria Voltes, Cristian Undurraga, Rocio Polvillo, Lucie Devos, Juan Tena, Ignacio Maeso, Sylvie RETAUX, José Luis Gómez-Skarmeta, Juan R. Martinez-Morales, Cristina Pujades

1 Centro Andaluz de Biología del Desarrollo, 2 UNIVERSITAT POMPEU FABRA, 3 CNRS UMR1917 Paris-Saclay Institute of Neuroscience, University Paris-Sud and Paris-Saclay, 4 CSIC Universidad Pablo de Olavide, 5 Universidad Pablo de Olavide

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Developmental programs often rely on parallel morphogenetic mechanisms that guarantee precise tissue architecture. While redundancy constitutes an obvious selective advantage, little is known on how novel morphogenetic mechanisms emerge during evolution. In zebrafish, rhombomeric boundaries behave as an elastic barrier, preventing cell intermingling between adjacent compartments. Here, we identify the fundamental role of the small-GTPase Rac3b in actomyosin cables assembly at hindbrain compartments. Here, we identify the fundamental role of the rac3b regulatory cluster expressed at boundaries emerged by the establishment of conserved regulatory blocks across species (9, 15). Despite this, there is little information on how the disruption of regulatory blocks during evolution may act as a positive force for the emergence of novel gene expression patterns and developmental mechanisms. Our report addresses the appearance of a novel gene regulatory landscape underlying hindbrain segmentation in zebrafish.

Embryonic segments are fundamental building blocks of the body plan. During animal development, proliferating cells are organized into segmental compartments with boundaries across which cells fail to intermingle, ensuring that their fates remain segregated as they proliferate and move. Thus, the establishment and maintenance of compartment boundaries is of critical importance in tissue segmentation and body plan organization (16). The vertebrate hindbrain is a model good to address the specification of segmental domains and the establishment of boundaries, since it is transitory segmented into rhombomeres (r1-r7) that constitute developmental units of gene expression and cell lineage compartments (17-19). The process of compartmentalization involves the sorting of cells from neighboring rhombomeres, which express Eph receptors or Ephrin ligands, and the formation of a cellular interface between adjacent segments named hindbrain boundary (20). This cell population not only displays a different morphology than its neighbors (21), but it serves to distinct functions as development proceeds. First, when morphological segments arise boundary cells work as an elastic mesh, preventing cell intermingling between adjacent compartments. We have previously shown that, in zebrafish, this is due to the enrichment of actomyosin cable-like structures in their apical side (Figure S1), whose formation requires Eph/Ephrin signaling and downstream small GTPase effectors (22). During neurogenesis, hindbrain boundaries behave as a node for signaling pathways –such as Notch, Wnts, or semaphorins- instructing the differentiation and organization of neurons in the neighboring rhombomeres (23-26).
enriched at the hindbrain boundaries, as shown by double staining with egr2a (red) and mouse monoclonal antibody against rhombomeres 3 and r5 (Figure 1f-i). Finally, sgca is also expressed at the hindbrain boundary regions (Figure 1j-m) and in the somites (32). The onset of expression of these three genes in hindbrain boundaries is very similar: expression starts around 17hpf, just when actomyosin cables have been assembled (22), and is maintained in boundary cells at least until 30hpf (Figure 1a, i, m). None of the additional genes within this chromosomal region - wdr64, lrc45, dcrx, and collagen 1b - are specifically enriched in boundaries and in turn present a non-spatially restricted pattern of expression (32). This common spatiotemporal expression led us to think that rac3b, rfng, and sgca may share cis-regulatory information responsible for driving their common expression to the hindbrain boundaries.

Interestingly, when comparing the chromosomal organization of rac3b, rfng, and sgca across different vertebrates, we found that this microsynteny block is conserved in zebrafish and Astyanax mexicanus (Figure 2a), both from the Ostariophysi superorder, as well as in all other osteichthyian species according to the available genomic information (Figure S2). In contrast, this microsynteny block is absent in tetrapods and non-ostariophysian fish, which show an alternative genomic organization, with sgca orthologs lying next to ppp1r9b and separated from the rac3b-rfng region (Figure 2a, Figure S3). This led us to test whether this specific chromosomal organization was relevant for gene expression at the boundaries. If this assumption is true, A. mexicanus embryos should display rac3b, rfng, and sgca expression in these same territories. To test this hypothesis, we cloned the corresponding Astyanax genes and performed in situ hybridization experiments.
Fig. 2. Expression of rac3b, rfng and sgca in hindbrain boundaries is related to their chromosomal organization. a) Phylogenetic tree displaying the microsynteny arrangements around rac3b, rfng and sgca in different vertebrate species; genes are represented by arrows showing their transcriptional orientation. Note that conservation of the rac3b/rfng/sgca cluster is present in zebrafish and A. mexicanus Pachón cavefish genomes. Orange boxes indicate the conservation of the synteny within the 5' region, and grey boxes the conservation within the 3' region. b-d) In situ hybridization experiments in A. mexicanus surface fish embryos with rac3b, rfng and sgca probes. Note that rac3b and rfng are expressed within the hindbrain boundary domains. Although the expression of sgca in the somites is maintained, no expression was observed in the boundaries. e-h) Expression of the orthologous rac3b, rfng and sgca genes in medaka embryos; note that there is no expression of any of these genes in the hindbrain boundaries, although of sgca expression within the somites is conserved. All pictures are dorsal views with anterior to the left.

Fig. 3. Expression and function of Rac3b in the hindbrain boundary cells. a-b) Whole mount in situ hybridization with rac3b in Tg(ela:GFP) embryos injected at 1-cell stage with MO-CONTROL (a-a') or MO-EphA4 (b-b'), followed by anti-GFP staining. Note the expression of rac3b in boundary cells (white arrowheads), and how this expression diminishes upon EphA4 downregulation. (a-b) Dorsal views, and corresponding (a-a') sagittal views displaying only the red channel. c-h) Loss of function of Rac3b either by splicing blocking morpholino MO-Rac3b (d-g), or by CRISPR-Cas9 induced mutation, rac3b/- (e-h), results in the disruption of actomyosin cables (d-e) and cell mixing (g-h), compared to control embryos (c,f). White arrowheads in (g,h) point to ectopic r3/r5 rhombomeric cells. MO-Rac3b figures correspond to MO-rac3bSBI4E5 injected Mu4127 embryos (Figure S6c). Rhombomeric cell mixing in morphants was observed by expression of mCherry in r3 and r5 cells, upon injecting the Mu4127 transgenic line. Cell mixing in rac3b/- hindbrains was assessed by in situ hybridization with egr2a. i-j) Sagittal views of representative examples of Tg(myosinII:GFP) embryos injected with the corresponding Rac3b construct, displaying either the merge (i-j), or the separate channels (i'-j'). i') hs:DN-Rac3b-Myc clone (in magenta) hitting the rhombomeric boundary, with the subsequent disruption of actomyosin cables; j') hs:CA-Rac3b-Myc clone (in magenta) in rhombomer 5 generating ectopic actomyosin II structures. Yellow arrowheads in (i-j') point to disrupted cables, and white arrowheads in (i-j') indicate ectopic actomyosin structures. Anterior is to the left in all images.
Fig. 4. Identification of hindbrain boundary cells cis-regulatory elements by the analysis of the rac3b/rfng/sgca cluster regulatory landscape. a) Chromosomal localization of the rac3b/rfng/sgca cluster and chromatin interaction profile by 4C-seq at 24hpf (n=2; overlayed grey peaks correspond to both replicates). The viewpoint used for both replicates is showed with a black arrowhead. b) Zoom in from (a) of a 170Kb window of chromosome 12 where most of the chromatin interactions, unveiled by 4C-seq, occur. Epigenetic marks of putative promoters (H3K4me3, green peaks) and active enhancers (H3K27ac, magenta peaks) (40) are shown along with ATAC-seq profiles from dissected hindbrains at 24hpf (n=2; black and grey profiles). c) Magnification of the region framed in (b) showing the H3K27ac profile and hindbrain specific ATAC-seq signatures. The ≈5.6Kb region (grey shaded region) was divided in three fragments associated with hindbrain ATAC-seq peaks (Box A-C), which were cloned in an enhancer reporter vector to generate stable transgenic lines. d-i) Dorsal views of embryonic hindbrains from Tg[Box A:GFP] (d-e), Tg[Box B:GFP] (f-g) and Tg[Box C:GFP] (h-i) stable transgenic lines at indicated stages. Note that Box B and Box C sequences (but not Box A) are able to drive GFP expression to the hindbrain boundaries (white arrowheads in f-i). White asterisks in (f-g) show the expression of GFP in the somites. In all pictures anterior is to the top.

As shown in Figure 2b-c, both rac3b and rfng are expressed at the boundary cells in A. mexicanus surface fish embryos at 24hpf. In contrast, although the expression of sgca in the somites is maintained, no expression was observed in boundaries (Figure 2d).
Interestingly, the onset of rac3b and rfg expression in hindbrain boundaries is at 18hpf, which is the equivalent developmental stage to zebrafish (Figure 1). It is important to note that rac3b, rfg and sgea orthologous genes are not expressed in the hindbrain boundaries in vertebrate species that do not share this specific chromosomal organization, including Actinopterygii fish such as medaka (Figure 2e-g), and tetrapods such as chick and mice (33). Moreover, rac3a, the sister paralog of rac3b is not expressed in the zebrafish hindbrain boundaries (32), and rather shows an expression pattern highly similar to that of the medaka rac3b (Figure S4).

These results suggest that thanks to the tight genomic linkage of rac3b-rfg with sgea, regulatory elements within this region are specifically combined to drive rac3b expression to the rhombomeric boundaries, allowing actomyosin assembly in the boundary cells. This prompted us to compare the presence of hindbrain actomyosin structures in species displaying the same genomic organization than zebrafish, such as A. mexicanus, and with different genomic organization such as medaka. Indeed, anti-PMLC immunostainings revealed enrichment of these structures within the hindbrain in A. mexicanus but not in medaka at the equivalent embryonic stages (18-19s, Figure S5). This suggests that Rac3b function in the hindbrain boundaries is conserved across the Ostariophysi superorder.

**Rac3b induces the formation of actomyosin structures in hindbrain boundaries**

We observed that the enriched expression of rac3b in boundary cells (Figure 1; Figure 3a–j) was diminished upon EphA4 downregulation (Figure 3b–h), as previously shown for RhoA (22). In order to confirm the putative role of Rac3b in assembling apical actomyosin cables within the boundary cells to prevent cell intermingling, we undertook distinct functional approaches. Loss-of-function of rac3b was assessed either by its downregulation using splice-blocking morpholinos (Figure 3d; Figure S6a,c), by generating a mutant allele using CRISPR-Cas9 genome editing technology (Figure 3e,h; Figure S6b,d–i), or by the clonal mosaic expression of a dominant negative form of Rac3b (DN-Rac3b-Myc; Figure 3i, j–q”). MO-Rac3b injected embryos displayed a disruption of the actomyosin cables when compared to control embryos (Figure 3c–d; control n=0/10 vs. MO-rac3bSB42E5 n=21/25 MO-rac3bSB4E1H4 n=10/13). According this disruption resulted in rhombomeric cell intermingling at 20hpf (Figure 3f–g; control n=1/16 vs. MO-rac3bSB4E5 n=10/15). rac3b/- embryos displayed defective actomyosin structures within the boundaries, although phenotypes were milder than in rac3b morphants (Figure 3e; control n=1/25 vs. rac3b/- n=3/7). Furthermore, rhombomeric cell sorting was also compromised at 24hpf (Figure 3b; control n=1/20 vs. rac3b/- n=11/22). The mutation of the rac3b locus by CRISPR-Cas9 did not compromise embryo survival, resulting only in mild phenotypic abnormalities at the standard temperature of 28°C (Figure S6d–e). However, curled body axis and increased lethality were observed when mutant embryos were incubated at 36°C and higher temperatures (Figure S6f–i). The mild phenotypes displayed by rac3b loss-of-function in actomyosin disruption could be explained either by genetic compensation induced by deleterious mutations (34, 35), or by the functional redundancy of the ubiquitously expressed RhoA (22).

To further confirm the role of Rac3b in actomyosin contractile structures in boundary cells, we conditionally and clonally modulated Rac3b expression in Tg[mysosinILGF] embryos at 14hpf and scored the phenotype at 18hpf. DN-Rac3b-Myc clones result in disruption of actomyosin structures when the clone hits the inter-rhombomeric actomyosin cable (n=44/54; see yellow arrowheads in sagittal views Figure3i–j”) but not in control clones. (n=14/81). Interestingly, this enrichment of myosin II structures can be observed in all rhombomeres, independently of their odd/even-identity, and in non-apical locations, suggesting that constitutive Rac3b activity is able to recruit myosin II to these sites, as previously described for RhoA (22).

These results support the role of Rac3b as a new small-GTase player in boundary cells inducing the assembly of actomyosin structures to prevent cell mixing.

**Analysis of the rac3b/rfg/sgea regulatory landscape and identification of hindbrain boundaries’ enhancers**

To gain insight into the evolutionary emergence of the rac3b/rfg/sgea cluster present in zebrafish and A. mexicanus, we performed a comparative analysis of the chromosome landscapes containing the ancestral gene blocks rac3/col1a1/col1a1 (Figure S3a) and ppp1r9/gho/col1a1 (Figure S3b) along the fish evolutionary tree. We observed that these conserved syntenic regions present in bony vertebrates (Figure 2) are also conserved in spotted gar, whose lineage diverged from other teleost fish before their specific whole genome duplication (36, 37). Whole genome analyses of chromatin interactions by Hi-C support the notion that these two blocks are each contained within their respective TADs in mammals (38), and thus it is very likely that they constitute distinct regulatory domains. In contrast, after the genome duplication in teleosts the architecture of these ancestral blocks became more flexible and, for each of the two generated paralogs, gene losses and chromosomal rearrangements are frequently observed in the different lineages (Figure S3). In particular, we observed that the rac3b/rfg/sgea conformation found only in Ostariophysi (i.e. zebrafish and Astyanax) entailed the fragmentation and fusion of the two ancestral blocks: rac3/col1a1/gho (Figure 2a; Figure S3) at the fish evolutionary tree. The presence of these two blocks in the same chromosome in zebrafish as well as in other vertebrates, such as medaka (Figure S3), northern pike and human suggests that intra-chromosomal rearrangements were the causative events that gave rise to this new gene organization in Ostariophysi.

The conservation of the syntenic arrangement rac3b/rfg/sgea in Ostariophysi (Figure S2), together with the shared expression patterns of the genes within this locus in zebrafish, suggest the emergence of a new common regulatory landscape. To investigate chromatin interactions in this region, we performed Chromosome Conformation Capture combined with high-throughput sequencing (39) (4C-seq) in 24hpf zebrafish embryos using the rfg/sgea

**Table 1. Coordinates of the tested cis-genomic regions within the Chromosome 12 Corresponding coordinates in the zebrafish genome annotation Zv9/danReR7 of the different regulatory regions (Box A–F) and sub-boxes.**

<table>
<thead>
<tr>
<th>Chr 12 coordinates Zv9</th>
<th>Box A</th>
<th>Box B</th>
<th>Box B1</th>
<th>Box B2</th>
<th>Box C</th>
<th>Box C1</th>
<th>Box C2</th>
<th>Box C3</th>
<th>Box C4</th>
<th>Box C5</th>
<th>Box D</th>
<th>Box E</th>
<th>Box F</th>
</tr>
</thead>
</table>
CRISPR-Cas9 deletion of the hindbrain boundaries cis-regulatory elements revealed the existence of redundant enhancers. 

(a) Scheme depicting the 2.5Kb deletion induced by CRISPR-Cas9 technology (Δ2.5Kb, grey shadowed stretch) containing Box B2 and Box C2/3, along with ATAC-seq profiles from 24hpf dissected hindbrains (black and grey profiles correspond to two different replicates). Position of the sgRNAs (in blue), used to generate the 2.5Kb deletion, and two pairs of primers (in magenta) used for genotyping the mutant line are displayed. 

(b-g) In situ hybridization analyses of rac3b, rfng and sgca in: wild type siblings (b-d), and CRISPR-Δ2.5Kb homozygous mutant embryos (e-g) at 22hpf. Note that rac3b and rfng expression does not considerably change within the hindbrain boundaries between wild type and homozygous mutant embryos. sgca expression in the somites is abolished in mutant embryos (black asterisk in g), due to the deletion of the main sgca promoter. All pictures are dorsal views of flat-mounted hindbrains with anterior to the left.

(h) Epigenetic profiles of putative promoters (H3K4me3, green peaks) and active enhancers (H3K27ac, magenta peaks) are shown within the chromosomal region containing rac3b along with ATAC-seq signatures (black and grey profiles) from dissected hindbrains at 24hpf. Three regions associated with most prominent ATAC-seq peaks were selected (shadowed in grey, Box D-F) and each of the fragments was cloned in an enhancer reporter vector to generate a stable transgenic line. 

(i-n) Dorsal views of embryonic hindbrains from Box D (i-j), Box E (k-l) and Box F (m-n) stable transgenic lines at indicated stages. Note that Tg[Box D:GFP] embryos display GFP expression in the hindbrain boundaries starting at 48hpf and that Box F is able to drive GFP expression to the hindbrain boundaries before 28hpf (white arrowheads in j, m, n). Box E did not drive GFP to the boundaries. In all pictures anterior is to the top. Overall, the expression of the rac3b/rfng/sgca microsyntenic group at the hindbrain boundaries is regulated by multiple enhancers. We have identified at least two early-activated (Box C and F), and two late-activated regulatory elements (Box B and D).

fusion border as reference bait. This analysis showed that rac3b, rfng and sgca genes are all part of the same interaction domain, with most of the chromatin interactions occurring within a region of 170Kb, between the genes wdr64 and col1a1b (Figure 4a). Next,
we focused on this region to identify the putative cis-regulatory elements responsible for the expression of the syntenic genes in hindbrain boundaries. To this end, we analyzed the distribution of predictive promoter and active enhancer epigenetic marks available for whole zebrafish embryos at 24hpf (40) in the previously described locus (green/magenta peaks in Figure 4b).

To further refine these analyses, we performed a genome-wide comparative study of open-chromatin domains by ATAC-seq (41), using zebrafish 24hpf dissected hindbrains as starting material.

A small collection of ATAC-seq peaks was identified through this procedure within the 170Kb window previously defined (grey and black peaks in Figure 4b). Interestingly, some of the most prominent ATAC-seq peaks were located at the junction between the two ancestral syntenic blocks, at the *rfng*/*sgca* border, partially overlapping with predictive marks for active enhancers (Figure 4b). Thus, to pinpoint cis-regulatory regions driving expression to boundary cells we focused our attention on this border using a classical transgenesis approach. To dissect the region, we generated stable transgenic lines harboring mainly non-overlapping genomic fragments (see grey boxes in Figure 4c). Three regions: Box A (2Kb), Box B (1.1Kb) and Box C (2.5Kb) (see Table 1 for chromosomal coordinates) were cloned into a vector carrying the GFP reporter and an internal (midbrain) transgenesis control (42), and injected into 1 cell-stage zebrafish embryos. Embryos deriving from F1 crosses showed no enhancer activity in the hindbrain boundaries for Box A (Figure 4d-e), whereas Box B and Box C were able to drive GFP expression to the boundaries (Figure 4f-i). In addition, Box B, which contains the proximal promoter region of *sgca*, was also able to drive expression to the developing somites (see asterisks in Figure 4f-g). Box B and C display the same spatial but different temporal activity at the hindbrain boundaries: Box B is active only from 24hpf onwards (Figure 4f-g, Figure S7a-c), and Box C is already active at 19hpf (Figure 4h-i, Figure S7d-f). This suggests that initiation and maintenance of gene expression might be under the control of different enhancer elements. Additional transgenesis experiments allowed further dissection of Box B and Box C (Figure S7g-n), narrowing the enhancer activity to Box B2 and Boxes C2/C3 (Table 1), consistent with the distribution of the ATAC-seq signal in this area.

---

**Fig. 6. Evolutionary model for the emergence of a novel hindbrain boundaries mechanism in Ostariophysi.**

a) Diagram showing the arrangement of the two ancestral blocks *rac3b*/*dcxr*/*rfng*/*gps1* (shaded in orange) and *ppp1r5b*/*sgca*/*coll1a1* (shaded in grey), in the common ancestor of all bony vertebrates. The new *rac3b*/*rfng*/*sgca*/*coll1a1* cluster emerged by intra-chromosomal rearrangement in the Ostariophysi lineage [1]. This new regulatory space allowed the gain of long-range regulatory contacts [2], and the acquisition of shadow enhancers to provide regulatory robustness [3].

b) The new topology allowed the emergence of a novel gene expression domain, and subsequently the formation of actomyosin cables at the rhombic boundaries. Pictures in b) are dorsal views except for the actomyosin cables image that is a sagittal view. Anterior is always to the left.

---

**Unveiling enhancers with highly correlated spatial activity**

In order to investigate the functional contribution of the identified enhancers to the regulation of expression of the neighboring genes, we deleted a 2.5Kb genomic region comprised between 3,669,916 and 3,672,466 by CRISPR-Cas9 genome editing (CRISPR- △ 2.5Kb; Figure 5a). This deletion contains both Box B2 and Box C2/C3 (Figure 5a, Figure S7g). Founders were screened for this 2.5Kb deletion and the F1 was generated and crossed to obtain embryos homozygous for the deletion. When the expression of *rac3b*, *ring* and *sgca* was compared between mutants and wild type siblings, no main changes in their spatiotemporal profile were observed at the hindbrain boundaries (Figure 5b-g). Since the 2.5Kb CRISPR-deletion contained the promoter of *sgca*, which acts as a somite proximal enhancer (Figure 4c), the expression of this gene in the somites was abolished at 24hpf and, thus, we could use it as an internal control of the deletion (compare Figure 5d and g). In agreement with the mild expression changes observed, the CRISPR- △ 2.5Kb mutant line is viable in between 3,669,916 and 3,672,466 by CRISPR-Cas9 genome editing (CRISPR- △ 2.5Kb; Figure 5a). This deletion contains both Box B2 and Box C2/C3 (Figure 5a, Figure S7g). Founders were screened for this 2.5Kb deletion and the F1 was generated and crossed to obtain embryos homozygous for the deletion. When the expression of *rac3b*, *ring* and *sgca* was compared between mutants and wild type siblings, no main changes in their spatiotemporal profile were observed at the hindbrain boundaries (Figure 5b-g). Since the 2.5Kb CRISPR-deletion contained the promoter of *sgca*, which acts as a somite proximal enhancer (Figure 4c), the expression of this gene in the somites was abolished at 24hpf and, thus, we could use it as an internal control of the deletion (compare Figure 5d and g). In agreement with the mild expression changes observed, the CRISPR- △ 2.5Kb mutant line is viable in homozygosis. However, the survival of the mutants is compromised, as only a sub-mendelian proportion of adults obtained from a heterozygous cross are mutants (4/43 = 9.3%). These data strongly suggest that other shadow enhancer regions might act redundantly, either totally or partially, to define the precise spatial and temporal activity of the interrogated genes.

Regulatory elements with overlapping functions may provide robustness to gene expression during embryonic development. Thus, we explored whether other putative regulatory elements identified by our ATAC-seq analysis may function as redundant additional shadow enhancers driving gene expression at the hindbrain boundaries. We cloned in our enhancer reporter vector some of the most prominent peaks within the contact region unveiled by 4C-seq on zebrafish hindbrains (Figure 4a), here termed Box D, E, F (Figure 5h; Table 1). After the generation of the corresponding stable zebrafish lines, embryos were assayed for GFP expression; indeed, two of the analyzed peaks drive GFP expression to the hindbrain boundaries, acting as partially redundant enhancers (Figure 5i-n). These zebrafish enhancer regions display different temporal activity: Box D activates only late in the hindbrain boundaries (Figure 5i), whereas Box F is already active at 28hpf (Figure 5m-n). Finally, we failed to detect...
The adaptive reasons behind the emergence of a cell sorting backup mechanism are currently unclear. A possible explanation may come from the faster embryonic development in Ostariophysi when compared to other fish (Figure S10, www.fishbase.org). The short hatching period in zebrafish (~48hpf) and Astyanax (~28hpf) (49), is shared by other orders of the Ostariophysi clade. This is in contrast to other teleost species such as members of the Euteleostei super-group, Ostariophysi sister orders, or even basal actinopterygian lineages, for which embryos take a week longer to hatch (Figure S10). Zebrafish embryos are able to swim and display a full hindbrain-wired scape response as early as 48hpf (50). In agreement, the heterochronic development of the muscles and the nervous system is particularly noticeable when the zebrafish transcriptome is compared with that of medaka (29). Thus, a possibility is that the machinery for cell sorting had evolved under the requirements imposed by the accelerated development of the hindbrain. An alternative hypothesis comes from the presence of a particular adaptation in Ostariophysi, the WeberianGland. The WeberianGland, representing an organ connecting the swim bladder with the auditory system that plays a role as sound amplifier (51). The acquisition of novel sensory inputs into the hindbrain vestibular nuclei may have contributed to an increased selective pressure for rhombomeric cell segregation in Ostariophysi. Importantly, these two scenarios are not incompatible. A rapid development and novel hearing capabilities may have both contributed to the preeminence of the ostariophysian lineage in freshwater environments, and both would have set specific evolutionary pressures on hindbrain developmental processes. In agreement with this, it has been recently shown that the zebrafish hindbrain has experienced an intense process of recent genetic innovation, being the only brain region with a highly significant recruitment of zebrafish- and ostariophysian-specific genes (52).

The study of the regulation of the small GTPase rac3b brought us to the identification of the block rac3b/fgn/scca as a microsyntenic group expressed at the hindbrain boundaries. Our functional and comparative genomic analyses outline a possible evolutionary scenario for the emergence of this regulatory block (Figure 6). The two ancestral gene blocks rac3b/fgn/scca and ppp1r3b/gca/col1a1 would be initially present in the same chromosome in the last common ancestor of all bony vertebrates (e.g. they are separated approximately 32Mb in the human chromosome 17). The specific whole genome duplication in fish has been postulated as a positive force for teleost's diversity and evolutionary success (53). This duplication may have led to a relaxation of some evolutionary constraints, releasing enough evolutionary pressure to allow the breakage of the two ancestral blocks. Then, the new block rac3b/fgn/scca/col1a1 would have emerged by intra-chromosomal rearrangement specifically in the Ostariophysi lineage (Figure 6). Through this rearrangement, scca could have brought pre-existing regulatory elements into the proximity of the rac3b locus, allowing the establishment of new long-range cis-regulatory interactions and the emergence of novel expression domains from the faster embryonic development we identified, Box B2, lies precisely at junction of this genomic rearrangement, at the scca promoter, and this enhancer element contains a conserved sequence present in all teleost species examined, including all Ostariophysi outgroups. This indicates that the evolutionary origin of the Box B2 enhancer predates the chromosomal rearrangement of Ostariophysi, the emergence of rac3b expression in hindbrain boundaries and the appearance of the actomyosin cables (Figure S11). Thus, it is likely that the evolution of these pioneer regulatory interactions with the...
ancient Box B2 element facilitated the emergence of additional redundant enhancers within the locus (Figure 6). Consistent with this scenario, the sequences of the two additional boundary enhancers located outside the junction genomic region with gca, Box D and Box F, are conserved among cypriniform species (Figure S12), suggesting that these two shadow enhancers evolved long after the origin of Ostariophy and the emergence of the rac3b/frng/gca regulatory block.

The partial overlapping activity of redundant enhancers appears to be a common regulatory theme, as they ensure robust development by suppressing transcriptional noise (54-56). It is the case for two neighboring genes due to the reliance of their promoter on other tethering elements, as enhancers and their target core promoters are not necessarily nest. It happens often that genes within a genomic domain are coor-

wdr64, lrrc45, they are contained within the 4C-delineated domain, some genes, selectively with genes within the

mentioned that their enhancer of gca locus is 3' of ftz and the enhancer of ftz lies between two genes; here, the selectivity of the enhancer for the Scr promoter depends on a proximal-promoter tethering element (60). For the rac3b/frng/gca locus, we show a differential behavior for gca as it is recruited to the boundaries in zebrafish but not in Astyanax.

Several examples of changes in regulatory sequences that provide the basis for evolving species-specific traits have been unveiled (3, 4, 61). In this study we reveal a specific example linking the appearance of a new regulatory domain to the emergence of a novel morphogenetic mechanism in the Ostariophy superclass: the formation of actomyosin cables at the rhombomeric boundaries. We have shown that this mechanism relies on the cooption of rac3b to the boundaries. However, the appearance of the rac3b/frng/gca regulatory space may have facilitated the evolution of more than one mechanism for hindbrain morphogenesis. In addition to rac3b, frng, which is not specifically expressed in this territory in other vertebrates (62, 63) was also coopted to this rhombomeric domain in Ostariophy. It has been suggested that activation of Notch in boundary cells promotes cell segregation, in a process dependent on frng (23). Moreover, frng plays a signaling role in hindbrain boundaries in the patterning of neurogenesis (26). Overall, our observation suggests that the umbrella of the novel rac3b/frng/gca regulatory block allowed the emergence of an entire morphogenetic program providing robustness to the hindbrain boundary cells to unfold their different functions upon morphogenesis. Thus, this seemingly serendipitous mutational event may have been crucial for adapting hindbrain morphogenes-

s to novel developmental traits of the ostariophysian lineage.

MATERIALS AND METHODS

Fish samples

Animals are treated according to the Spanish/French and European regulations for handling of animals in research. All protocols have been approved by the Institutional Animal Care and Use Ethics Committees and implemented according to national (Spanish/French) and European regulations. All experiments were carried out in accordance with the principles of the 3Rs.

Zebrafish strains

Zebrafish (Danio rerio) embryos were obtained by mating of adult fish using standard methods. All zebrafish strains were maintained individually as inbred lines. MGI4127 is an enhancer trap line in which the trap cassette containing a modified version of Gal4 (KaiT4) and mCherry (KaiT4A-UAS-mCherry cassette) were inserted in the 1.5Kb downstream of egr2a/krr20 gene (64). Tg[gal4-GFP] line is a stable reporter line where chicken element A from egr2a was cloned upstream of the gfp reporter in a modified pTol2 vector (65). Tg[mynsion1-HmCherry/GFP] visualizes myosin II (non-cardiac Myosin associated to Actin filaments) (66, 67). Embryos homozygous for rac3b/- and CRISPR-/-SK mutations in the Tg[mynosin1/GFP] genetic background were obtained by incross of heterozygous carriers.

Astyanax mexicanus samples

Adult Astyanax mexicanus surface fish (origin: San Solomon Spring, Balmorhea State Park, Texas) were maintained and bred at 26°C on a 12:12 hours light/dark cycle in tap water (68). Embryos and larvae were collected after natural spawning and staged according to the developmental staging table (49).

Medaka samples

Medaka (Oryzias latipes) wild type strain was kept as closed stock in freshwater under artificial reproductive conditions (10:14 hours light/dark cycle) at 26-28°C. Embryos were staged as previously described (69).

Comparative genomics

Genes included within the two ancestral gene blocks, Rac3/Dcxr/Rfng/Gps1 and Ppl196/Sgea/Col1a1, were searched across the different studied species using a range of BLAST algorithms both genome-wide and within syntenic regions. Gene annotations were generated and manually curated using information from cross-species comparisons and available expression data. We compared the relative orientations and positions of these genes by browsing the genomes of the studied species through the NCBI (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi), UCSC (http://genome.ucsc.edu), Ensembl Metazoa (http://www.ensembl.org/index.html), and EFISH Genomics (http://efishgenomics.integrebiology.msu.edu/) webpages, using the following genome versions: Astyanax mexicanus (Pachón cavefish) 1.0.2 and 2.0, Clupea harengus (Atlantic herring) AS965633v1, Cyprinus carpio (common carp) 000951615, Danio rerio (zebrafish) Zv9, Electrophorus electricus (electric eel) v.1.0, Esox lucius (Northern pike) EsoLuc1.0, Gadas morhua (Atlantic cod) gadMor1, Gallas gallus (chicken) galGal4, Gasterosteus aculeatus (stickleback) v1.0, Homo sapiens (human) h38, Latimeria chalumnae (coelacanth) latChA1, Oryzias latipes (medaka) v1.0 and ASM223467v1, Pimephales promelas (fathead minnow) FHM_SOAPdenovo, Sinocyclochilus rhinoceros SAMN0332009v1.1. Due to incompleteness and gaps in some of the previous genome assemblies, to recover the Box B2 conserved non-coding sequence (Figure S11) we also used alternative assembly versions for C. carpio (ASM127010v1), and the genome assembly of A. rhinocerus sister species, S. grahami (SAMN03320097,WGSv1.1). The phylogenetic relationships depicted in Figure S3 and S11 are according to the most recently published phylogenetic trees of these ostariophysian lineages (46, 70).

Sequence conservation of zebrafish shadow enhancers within the rac3b-dcr region was characterized using VISTA (71) and zebrafish Zv9 as a reference sequence, SLAGAN as the alignment program and the following parameters: 100bp window and 70% identity in 70bp.

cDNA cloning
Total RNA from zebrafish embryos at 24hpf was reverse-transcribed using oligo-dT primers (SuperscriptIII kit, Invitrogen). Complete or partial sequences for wdr64, lrc45, rac3b, dcxr, and sgca were amplified by PCR using the following primers: wdr64-Fw: 5'-GGA GCT TTC CTT CAG GAC AA-3', wdr64-Rv: 5'-TTG GTG GAG ACC GAG TCA A-T-3'; lrc45-Fw: 5'-TTC TGT CAG AAA CTT CCT GAA CAT TAA-3', lrc45-Rv: 5'-AGC AGA CTT ACC CTG GAA TTC TTG-3'; sgca-Fw: 5'-GGG TAA AAC CTG CCT GCT GAT-3', sgca-Rv: 5'-AGT GGC CTC TGG GAT GAT GAT G-3'; rac3b-Fw: 5'-GGG TAA AAC CTG CCT GCT GAT-3', rac3b-Rv: 5'-AGT GGC CTC TGG GAT GAT GAT G-3'; dcxr-Fw: 5'-TTC ATG AAC GGA TGG TTA TCT-3', dcxr-Rv: 5'-ATG AAC GGA TGG TTA TCT-3'. PCR products were sub-cloned in pGEM®-T Easy (Promega) and sequenced. rac3b (FO250597) and sgca (FO250544) partial cDNAs originate from our cDNA library (73).

Total RNA from medaka embryos was reverse-transcribed using oligo-dT primers (SuperscriptIII kit, Invitrogen) and partial sequences for rac3b, dcxr, and sgca were amplified by PCR. Primers were the following: rac3b-Fw: 5'-CTC GTC GGT TTC ATC CCA GAC CC-3' and rac3b-Rv: 5'-ATG GCC TCT TGC AAG ACC GAA CAT CAC T3' and 10ng/μl MO-rac3bSBI4E5 5'-AAA AAA AAG ACA GAA CAT CAC ACC TTT GTA CCG A-3' (Figure S6a). Efficiency of the MO-rac3b was checked by PCR using the following primers: primary-Fw: 5'-ATG GTA AAC CAC TGA ACC TGG G-3', primer-Rv: 5'-GGT AAG TGA TGG GAT GCA TT-3'. Since MO-rac3bSBI4E5 was more efficient and had less toxic effects for the embryos, we did most of the experiments with this splicing-defective morpholino (Figure S6c).

Antisense morpholinos

For morpholino knockdowns, embryos were co-injected at 1-cell stage with translation-blocking or splicing defective morpholino oligomers (MOs) obtained from GeneTools, Inc. MOs were as follows: 5ng/μl MO-EphA4a (25), 5'-AACC ACA AGC GCA GCC ATT GGT TTC-3', 5'-GAGC CCA TTT GCT TGC AAG AAT TG-3', 2.5ng/μl MO-rac3bSBI4E5 5'-TAC CAC TGC AGC ACA GAA CAT CAC T3' and 10ng/μl MO-rac3bSBE4I4 5'-AAA AAA AAG ACA GAA CAT CAC ACC TTT GTA CCG A-3' (Figure S6a). Efficiency of the MO-rac3b was checked by PCR using the following primers: primary-Fw: 5'-ATG GTA AAC CAC TGA ACC TGG G-3', primer-Rv: 5'-GGT AAG TGA TGG GAT GCA TT-3'. Since MO-rac3bSBI4E5 was more efficient and had less toxic effects for the embryos, we did most of the experiments with this splicing-defective morpholino (Figure S6c).

Conditional overexpression

CA-Rac3 (Q61L-mutation) and DN-Rac3 (T71N-mutation) constructs were generated by site-directed mutagenesis (QuickChange Site-Directed Mutagenesis Kit, Stratagene #200518), and cloned into the MCS of a Tol2-based custom vector containing a heat shock (HS) promoter and a Myc-tag. Tg[mysinIIImCherry/GFP] embryos were injected at 1-cell stage, grown at 28.5°C, and heat-shocked at 14hpf. All embryos were fixed at 18-20hpf, co-immunostained for Myc and GFP, and imaged for further analysis. For the phenotypic analysis: i) the integrity of the actomyosin cable was assessed in Myc or DN-Rac3b-Myc clones hitting the boundary; and ii) ectopic actomyosin structures were scored observed in Myc or CA-Rac3b-Myc clones located within the rhombomere.

Whole mount in situ hybridization

Zebrafish whole-mount in situ hybridization was adapted from (75). The following riboprobes were generated by in vitro transcription from cloned cDNAs: egr2a/lrc20 (76), rfgf (23), and ggrf (22), rac3b, and sgca. For arhgap29b and rasgef1ba, in vitro transcription was performed by RT-PCR using the following primers: arhgap29bFw: 5'-GGT GAG CAG CTC ATG AAA CA-3', arhgap29bRv: 5'-TTA TAC GAC TCA CTA TAG GGT GTT TTC GAC AGC AAC TAA G-3'; rasgef1baFw: 5'-CTC AGG TCG CTC TTT CT-3' and rasgef1baRv: 5'-TTA TAC GAC TCA CTA TAG GGT GTT TTC GAC AGC AAC TAA G-3'. The chromogenic in situ hybridizations were developed with NBT/BCIP (blue) and FastRed (red) substrates. For fluorescent in situ hybridization, DIG-labeled riboprobes were developed with fluorescein-tyramide substrate (TSA system). After staining, embryos were either flat-mounted and imaged under Leica DM6000B fluorescence microscope or whole-mounted in agarose and imaged under SP5 or SP8 Leica confocal microscopes.

In toto embryo immunostainings

For immunostainings, embryos were blocked in 5%GS/PBT 1h at RT and incubated O/N at 4°C with primary antibody. Primary pAbs were the following: anti-DesRed (1:500, Clontech), anti-b-tubulin (1:200, Torrey Pines), and anti-Myc (1:200, Clontech). After extensive washings with PBS, embryos were incubated with secondary Ab conjugated with Alexa Fluor®488 or Alexa Fluor®555 (1:500, Invitrogen). Embryos were flat-mounted or whole-mounted in agarose, and imaged under a Leica SP5 or SP8 confocal microscope.

Immunostaining for phospho-myosin light chain (anti-PMLC) was as follows: embryos were blocked in 10%GS/0.8% Triton-X100/PBS 3h at RT and incubated O/N at 4°C with phospho-Myosin Light Chain 2 antibody (1:500, Cell Signaling). After extensive washings with 0.8%Triton-X100/PBS, embryos were incubated with secondary Ab conjugated with Alexa Fluor®488 (1:500, Invitrogen). Embryos were whole-mounted in agarose and imaged under SP5 or SP8 confocal microscope.

Assessment of actomyosin cable-like structures

Live embryos from Tg[ggtag/flac:GFP] lines were anesthetized with 0.4% Tricaine (Sigma #A-5040) and mounted as previously described (22). In some cases, cables were imaged in fixed embryos immunostained for the reporter protein or anti-PMLC. For actomyosin cable-structure analysis, whole-mounted embryos were imaged in the 5'6 confocal microscope, and 0.6um z stacks were acquired in dorsal view and re-sliced to generate YZ confocal cross-sections. Images were re-sliced in XZ, and finally, a maximal projection of the XZ sections corresponding to the apical side of cells in the neural tube was generated (22). Animations of cable-like structures were generated using ImageJ (Figure S1). For scoring the disruption of actomyosin cable-like structures, the integrity of these structures in each embryo was analyzed and when at least two of the cables were disrupted embryos were considered affected.

Analysis of cell mixing

Footline Author
Confocal images of life or fixed embryos were acquired in dorsoventral view covering the r3-r5 region with 1 μm z distance. Images were re-sliced in Z planes and analyzed along the whole DV axis for presence of red ectopic cells either in M24127 or embryos in situ hybridized with egr2a. These stacks were then projected into a single dorsal view image for display (22). When ectopic cells were observed the embryo was scored as positive, regardless the number of ectopic cells observed. Ectopic cells are the result of cells undergoing mitosis, which incurred into the neighbouring territory, and due to the disruption of cables cannot be brought back to the territory of origin (22). Therefore, the number of ectopic cells usually varies between one and three, since within this time period cells do not undergo more than two cell cycles. This makes difficult to use the number of total ectopic cells to measure the expressivity of the phenotype. Thus, we like better to score for ectopic cells vs. non-ectopic cells.

Transgenesis

For the analysis of regulatory DNA elements, the PCR fragments were subcloned into PCR8/GW/TOPO vector and, using Gateway technology (Life Technologies), they were shuttled into the cassette for expression in transgenic zebrafish (42). Zebrafish transgenic embryos were generated using the Tol2 transposon/transposase method, with minor modifications. One-cell stage embryos were co-injected with 2 nl of a solution containing 250ng/μl Cas9 protein and 40ng/μl sgRNAs. For screening of the edited genome, gDNA was obtained by incubating the samples in TE buffer supplemented with 5% Chelex-100 (BioRad) and 10 μg/ml Proteinase K (Roche) for 4h at 55°C and 10min at 95°C and then stored at 4°C. 1 μl of the supernatant was used as template in a standard 25 μl PCR reaction.

Injection and genotyping

One-cell stage zebrafish embryos were co-injected with 2-3nl of a solution containing 250ng/μl Cas9 protein and 40ng/μl sgRNAs. For screening of the edited genome, gDNA was obtained by incubating the samples in TE buffer supplemented with 5% Chelex-100 (BioRad) and 10 μg/ml Proteinase K (Roche) for 4h at 55°C and 10min at 95°C and then stored at 4°C. 1 μl of the supernatant was used as template in a standard 25 μl PCR reaction.

Screening of the CRISPR-rac3b mutants was performed by PCR amplification using the following primers: rac3b-Fw: 5'-GAACCTCCAAAAATATGTGAATG-3', rac3b-Rv: 5'-TGGACTGTTGGTTATAGTG-3'. The wild type PCR product of 289bp undergoes digestion with BstXI leading to two bands, and a single band is detected after digestion in the mutant by loss of the BstXI restriction site.

Screening of CRISPR-rac3b deletion was performed by PCR amplification with the following primers: P1-Fw: 5'-TACTTGTCAGCAAAAG-3'; P2-Fw: 5'-GAAAATAAGCAGAATTAGTGG-3'. The wild type PCR product of 2.9Kb band and heterozygous fish amplify the 2.9Kb band and a 535bp band associated to the deletion. P2 primers are designed to be inside and outside the deletion, therefore mutant fish do not display the 764bp band (Figure 5a).

Pharmacological treatments

Treatment with para-Nitroblebbistatin was applied once the neural tube was already formed to avoid interfering with its early morphogenesis (22). Thus, in all experiments embryos at 14hpf were dechorionated and treated until 20hpf at 25°C with: i) 1× Para-Nitroblebbistatin (myosin II inhibitor (80)); ii) 5μM para-Nitroblebbistatin; and iii) 5μM para-Nitroblebbistatin plus an insertion of 205bp; sgRNA2: 5'-AGGCTCGTGAATTCG-3'. The wild type CRISPR -Cas9 genomic edition generated an insertion of 11bp and a deletion of 60bp in exon 2 resulting in a truncated protein of 22AA (Figure S6a-b). ii) CRISPR-rac3b 2.5kb genomic deletion covering the chromosomal region chr12 3,669,901-3,672,546 (from Zs9/danRe7r coordinates), resulted in a deletion of 2564bp plus an insertion of 205bp; sgRNA1: 5'-GGGATCTCTGTGATGCTGGC-3', sgRNA2: 5'-AGGCTCTGATCTGGTTAGG-3'. (Figure 5a).

Injection and genotyping

One-cell stage zebrafish embryos were co-injected with 2-3nl of a solution containing 250ng/μl Cas9 protein and 40ng/μl sgRNAs. For screening of the edited genome, gDNA was obtained by incubating the samples in TE buffer supplemented with 5% Chelex-100 (BioRad) and 10 μg/ml Proteinase K (Roche) for 4h at 55°C and 10min at 95°C and then stored at 4°C. 1 μl of the supernatant was used as template in a standard 25 μl PCR reaction.

Screening of the CRISPR-rac3b mutants was performed by PCR amplification using the following primers: rac3b-Fw: 5'-GAACCTCCAAAAATATGTGAATG-3', rac3b-Rv: 5'-TGGACTGTTGGTTATAGTG-3'. The wild type PCR product of 289bp undergoes digestion with BstXI leading to two bands, and a single band is detected after digestion in the mutant by loss of the BstXI restriction site.

Screening of CRISPR-rac3b deletion was performed by PCR amplification with the following primers: P1-Fw: 5'-TACTTGTCAGCAAAAG-3'; P2-Fw: 5'-GAAAATAAGCAGAATTAGTGG-3'. The wild type PCR product of 2.9Kb band and heterozygous fish amplify the 2.9Kb band and a 535bp band associated to the deletion. P2 primers are designed to be inside and outside the deletion, therefore mutant fish do not display the 764bp band (Figure 5a).

Pharmacological treatments

Treatment with para-Nitroblebbistatin was applied once the neural tube was already formed to avoid interfering with its early morphogenesis (22). Thus, in all experiments embryos at 14hpf were dechorionated and treated until 20hpf at 25°C with: i) 1× Para-Nitroblebbistatin (myosin II inhibitor (80)); ii) 5μM para-Nitroblebbistatin; and iii) 5μM para-Nitroblebbistatin plus an insertion of 205bp; sgRNA2: 5'-AGGCTCGTGAATTCG-3'. The wild type CRISPR -Cas9 genomic edition generated an insertion of 11bp and a deletion of 60bp in exon 2 resulting in a truncated protein of 22AA (Figure S6a-b). ii) CRISPR-rac3b 2.5kb genomic deletion covering the chromosomal region chr12 3,669,901-3,672,546 (from Zs9/danRe7r coordinates), resulted in a deletion of 2564bp plus an insertion of 205bp; sgRNA1: 5'-GGGATCTCTGTGATGCTGGC-3', sgRNA2: 5'-AGGCTCTGATCTGGTTAGG-3'. (Figure 5a).

Phenotype analyses of rac3b/- embryos (Figure S6)

Eye area (a=πnr²) was quantified using ImageJ software by measuring the external perimeter (P=2πr) around each eye. Data from rac3b+/+ and rac3b/- embryos were analyzed using an unpaired t-test. The body angle was measured from the eye lens to the tip of the tail using the center of the yolk as vertex. Angles obtained from three different clutches of rac3b+/+ and rac3b/- larvae were analyzed applying an unpaired t-test. Lethality data were analyzed applying a two-way ANOVA test.

ACKNOWLEDGEMENTS

We thank people who kindly provided us transgenic fish lines and reagents, especially P Charnay, CP Heisenberg, R Koster, A Málináš-Czimidza and D Wilkinson, and those who critically read the manuscript such as E Jimenez-Guri and C Cortes-Campos. We would like to thank S Calatayud for technical assistance, and the Martinez-Morales and Pujades labs for critical insights. This work was supported by La Marató TV3 grant 348/2014 to CP, "Equipue RMF" to SR, and Spanish Ministry of Economy and Competitiveness (MINECO-FEDER): Grants BFU2012-32199 and BFU2015-67400-P to CP, Unidad de Excelencia María de Maeztu MDM-2014-0370 to DCEXS-UPF, BFU2011-22916, P11-CV-7256, BFU2014-53785 and BFU2014-55738-RD to DCEXS-UPF.
Chile). JT was a recipient of a postdoctoral Beatriu de Pinos fellowship (AGAUR, Generalitat de Catalunya). CP. JL and CU were recipients of a postdoctoral Beatriu de Pinos fellowship (AGAUR, Generalitat de Catalunya). AV is recipient of a predoctoral fellowship from LaCaixa, and IB holds a predoctoral fellowship (AGAUR, Generalitat de Catalunya). CP is recipient of an ICREA Academy award (Institució Catalana per la Recerca i Estudis Avançats, Generalitat de Catalunya).


