Unique and non-redundant function of csf1r paralogues in regulation and evolution of post-embryonic development of the zebrafish

Joana Caetano-Lopes1,2, Katrin Henke1,2, Katia Urso3,4, Jeffrey Duryea5, Julia F. Charles3,4, Matthew L. Warman1,2 and Matthew P. Harris1,2,*

ABSTRACT
Evolution is replete with reuse of genes in different contexts, leading to multifunctional roles of signaling factors during development. Here, we explore osteoclast regulation during skeletal development through analysis of colony-stimulating factor 1 receptor (csf1r) function in the zebrafish. A primary role of Csf1r signaling is to regulate the proliferation, differentiation and function of myelomonocytic cells, including osteoclasts. We demonstrate the retention of two functional paralogues of csf1r in zebrafish. Mutant analysis indicates that the paralogues have shared, non-redundant roles in regulating osteoclast activity during the formation of the adult skeleton. csf1ra, however, has adopted unique roles in pigment cell patterning not seen in the second parologue. We identify a unique noncoding element within csf1ra of fishes that is sufficient for controlling gene expression in pigment cells during development. As a role for Csf1r signaling in pigmentation is not observed in mammals or birds, it is likely that the overlapping roles of the two paralogues released functional constraints on csf1ra, allowing the signaling capacity of Csf1r to serve a novel function in the evolution of pigment pattern in fishes.

KEY WORDS: Zebrafish, csf1r, Osteoclast, Bone, Pigment, Evolution

INTRODUCTION
Paralogous genes, or cases of close gene copies due to duplication, represent a unique case in which to study changes in gene function throughout evolution (Ohno, 1970). Related genes, arisen from a whole-genome duplication, or ohnologues, are often lost due to drift unless these genes evolve new functions (neo-functionalization) or, alternatively, partition previously performed roles (sub-functionalization) (Kassahn et al., 2009). Instances of gene duplication can relax constraint on individual gene function and allow compartmentalization as well as evolution of novel functions. However, overlap in function of retained genes can also lead to a bias in the expression of phenotypic variability due to buffering/redundancy and inter-dependent phenotypes. It is unclear how such dependent mechanisms shape the phenotypic response of organisms and contribute to diversification in development and evolution.

Ray-finned fishes represent the most specious lineage of vertebrates, harboring great diversity in skeletal shape, size and function. Resorption of bone during skeletal development as well as homeostasis is an essential aspect to the formation of shape, structure and function of the skeleton. The form of bones of all vertebrates arises through the interactions between two main cell types: osteoblasts, which lay down new bone, and osteoclasts, which serve to remove bone (Apschner et al., 2011; Weigele and Franz-Odendaal, 2016). The role of osteoclasts in bone remodeling is quite complex, depending on the local cellular environment as well as mechanical forces on the tissue. These interactions are not yet well defined, but it is clear that they are important in the response and adaptability of the skeleton, including the dentition, to functional demands. Osteoclast differentiation and function are dependent on signaling through colony-stimulating factor1 (csf1) (Felix et al., 1990; Hakeda et al., 1998; Mellis et al., 2011) and its receptor, colony-stimulating factor 1 receptor (csf1r), a receptor tyrosine kinase expressed by cells of the monocyte/macrophage lineage and necessary for their differentiation and function (Barreda et al., 2004; Chitu and Stanley, 2017; Dai et al., 2002; Herbomel et al., 2001; Oosterhof et al., 2018; Ryan et al., 2001). In addition, in teleost fishes, csf1r has been found to be expressed in neural crest progenitors and to act to regulate pigmentation cell differentiation and pattern (Parichy et al., 2000), a trait not related with CsF1 signaling in other vertebrates. It has been argued that selection on pigmentation phenotypes in fishes have contributed to their diversification (Irion et al., 2016; Salis et al., 2019). However, the additional role of csf1r function in regulating skeletal development may constrain the phenotypic spectrum of morphological variation in evolution.

Teleost fishes, which include the zebrafish, share an evolutionary whole-genome duplication event. Whereas previous analyses of csf1r gene evolution in marine fishes have uncovered two csf1r paralogues in many lineages (Braasch et al., 2006), in zebrafish or carp, members of a common teleost lineage Ostariophysi, a functional second parologue could not be identified. Rather, phylogenetic analysis supported the loss of this parologue due to drift (Braasch et al., 2006). Recently, a homologue of csf1ra, csf1rb, was identified in zebrafish regulating microglia density and distribution (Oosterhof et al., 2018). This csf1r parologue showed an additive effect in the regulation of microglial density, suggesting a shared role in microglial regulation (Oosterhof et al., 2018). It remains unclear whether both genes are functionally redundant or whether a sub-functionalization occurred as effects in other tissues were not analyzed.

Here, we describe phylogenetic and functional evidence for retention of a functional csf1rb parologue in zebrafish that has
overlapping and non-overlapping roles with csf1ra in skeletal remodeling. However, csf1ra has adopted a unique role in regulating pigmentation not seen in the retained paralogue. The functional buffering provided by the two paralogues may have provided context for the development of a novel function of csf1ra in development and important in the evolution of diversity.

RESULTS
Recent work established a potential paralogue of csf1r that is maintained in zebrafish (Oosterhof et al., 2018). In order to understand patterns of csf1r gene evolution, we assessed the phylogenetic relationship of the existing, annotated csf1r orthologues in vertebrates (Fig. S1). The ‘b’ orthologues did not strongly differentiate as a single clade (>70 bootstrap probability). However, following Braasch et al. (2006), analysis of synteny among csf1r in the genomes of teleost fishes supports the conclusions that the ‘b’ orthologues share a common genomic structure on the chromosome and are differentiated from ‘a’ orthologues (specifically loss of hmgxb3 and retention of camk2a). We further assessed clustering of the putative csf1rb paralogue with closely related Kit and Pdgf genes from zebrafish. An unrooted maximum likelihood tree shows strong support for the newly identified csf1r paralogue, csf1rb, as being in the csf1r gene family (Fig. S2A).

**csf1r paralogues are differentially regulated during early development**

To assess potential different roles of csf1r paralogues in development, we analyzed expression of both csf1ra and csf1rb during early development and in scales of the adult zebrafish. Consistent with what has been previously reported for csf1ra (Herbomel et al., 2001; Parichy et al., 2000; Thise et al., 2008), we detected expression in migrating neural crest (dorsal) and in early macrophages at 2 dpf (arrowheads and arrows, Fig. 1A). We saw no detected expression in migrating neural crest (dorsal) and in early (Herbomel et al., 2001; Parichy et al., 2000; Thisse et al., 2008), we capitalized on the specificity of this line to experimentally isolate osteoclast cell populations using fluorescence-activated cell sorting (FACS) to assess expression of csf1r paralogues in skeletal cells (Fig. S4B). As a comparison, we used a second transgenic line Tg[sp7:EGFP] (DeLaurier et al., 2010) to isolate pre-osteoblasts from comparable tissue sources (Fig. S4A). We confirmed isolation of osteoclasts and osteoblasts through analysis of the expression of known differentiation markers for these cell populations (Fig. 1D; Fig. S4C). Consistent with the known role of Csfr1 in osteoclast maturation (e.g. Chitu and Stanley, 2017), we were able to detect expression of both csf1ra and csf1rb in mature osteoclasts isolated from adult zebrafish (Fig. 1D). This expression was specific to osteoclasts, as sp7+ osteoblasts showed no expression of either csf1r paralogue. After this, we isolated RNA from individual skeletal elements (fin, teeth and vertebra) with detectable populations of Tg[ctsk:DsRed]-positive cells and we were able to determine that both csf1r paralogues were expressed in these tissues (Fig. 1E), complementing the detection of osteoclasts associating with the skeletal structures.

**Isolation of csf1ra and csf1rb zebrafish mutants**

To test function of Csfr1 receptors in zebrafish, we isolated mutants in both csf1r paralogues. In a mutagenesis screen, we isolated a mutant allele of csf1ra through expression of its recessive...
phenotype of altered adult pigmentation (Henke et al., 2017) (Fig. 2B). The loss of xanthophores and reduction of melanophores seen in mh5 mirrors the phenotype previously observed in the panther and salz und pfeffer mutants due to loss of function of csf1ra (Parichy et al., 2000). Sequencing of the coding region of csf1ra in the mh5 mutant revealed a G to T substitution at bp 1466, predicted to cause early truncation of the protein at position 454 (E454X; Fig. S2B). The mutation is tightly linked to the pigmentation phenotype and has not been observed in wild-type zebrafish populations. Given the phenotypic similarity to known csf1ra mutants, and the severity of the predicted change, the G to T substitution is thought to be the causative mutation underlying the allele of function alleles of csf1rb mutant phenotype.

Second, we used CRISPR/Cas9 genome editing to generate loss-of-function alleles of csf1rb. We induced small INDELs at the csf1rb locus by targeting exon 3 of the csf1rb gene with a pool of 5 guide RNAs. We isolated two mutant zebrafish lines, csf1rbmh108 and csf1rbmh112. csf1rbmh108 leads to an induced frame shift in exon 3 (c.479-482delinsTGAATTAT) predicted to cause an early truncation of the protein. The csf1rbmh112 deletion is also predicted to lead to early truncation (c.476-488del; Fig. S2C). In contrast to the csf1ra mutants, which are easily identifiable by their aberrant pigmentation, homozygous csf1rb mutants do not show a detectable morphological phenotype as adults (Fig. 2C). Double csf1ra;csf1rb mutants exhibit a pigmentation phenotype indistinguishable from the csf1ra single mutants (Fig. 2D).

To determine the effect of these mutations on Csf1r activity, the expression of csf1ra (Fig. 2E) and csf1rb (Fig. 2F) was analyzed by qRT-PCR on scales of mutants and wild-type siblings. In each csf1ra or csf1rb single mutant, we were unable to detect significant expression of the affected gene; thus, it is likely that both mutants represent null alleles. It is unlikely that the effect on expression is solely due to reduction in the number of osteoclasts, as both csf1ra and csf1rb are seen expressed in the mutant of the other paralogue. csf1rb expression is significantly reduced in csf1ra mutants, however, suggesting that the loss of csf1ra may limit the population of cells normally expressing csf1rb or, alternatively, that signaling through csf1ra is necessary to maintain csf1rb expression. In csf1ra;csf1rb double mutants, there was no detectable expression of either gene (Fig. 2E,F).

**csf1r paralogues and remodeling of the post-embryonic zebrafish skeleton**

Maturation and remodeling of the zebrafish skeleton occurs during juvenile development and throughout adulthood. Zebrafish have both mononucleated and multinucleated osteoclasts that participate in normal bone remodeling (Witten et al., 2001; Witten and Huysseune, 2009). Mononucleated osteoclasts are numerous and participate in laminar bone resorption without the presence of lacunar pits, whereas multinucleated osteoclasts are large and fewer in number but have more differentiated resorption pits. We confirmed the broad spectrum of nucleation in zebrafish osteoclasts in adult tissue. Hoechst staining of isolated ctsk+ cells in adult zebrafish skeletal tissue demonstrated the recovery of a wide diversity of single and multinucleated cells (Fig. S4E,F). The presence and activity of both populations of cells in tissues in situ can be detected through analysis of the activity of tartrate-resistant acid phosphatase (TRAP) using histochemical methods. We addressed the role of each csf1r parologue in regulating the number and function of osteoclasts in late development.

---

**Fig. 2. Adult phenotypes of csf1ra and csf1rb mutants.** (A-D) Representative photomicrographs of adult (A) wild-type zebrafish (wt), (B) homozygous csf1ra<sup>mh5</sup>, (C) csf1rb<sup>mh108</sup> and (D) csf1ra<sup>mh5</sup>;csf1rb<sup>mh112</sup> mutant zebrafish. (E,F) Quantitative RT-PCR of csf1ra and csf1rb in isolated scales of wild type, or mutants singly homozygous for csf1ra (a<sup>−/−</sup>) or csf1rb (b<sup>−/−</sup>), and of double mutants (a<sup>−/−</sup>b<sup>−/−</sup>). (G,H) Representative photomicrographs of scales collected from 10 wpf wild-type fish stained using TRAP (G, arrow) and (H) modified Von Kossa staining (arrow). (I,J) Quantitation of TRAP and Von Kossa staining in wild-type and mutant scales. For TRAP activity, each data point represents the percentage of scales stained by TRAP obtained by evaluating 10 scales per fish. For Von Kossa analysis, each data point represents the percentage of resorbed area per individual scale. Data are mean±s.d. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.
We took advantage of the dynamic homeostasis of scale mineralization in fish to assess osteoclast activity (de Vrieze et al., 2014). The scales that form across the flank of a fish are uniquely suited for analysis of matrix resorption and provide individual test cases for osteoclast function. As an indirect measure of the presence as well as the localization of osteoclasts, we performed an enzymatic assay for TRAP on scales from csf1ra and csf1rb mutants, csf1ra;csf1rb double mutants, and their respective wild-type siblings (Fig. 2G-I). On average, 75% of scales from wild-type fish had TRAP staining, with localized areas of staining flanking the center and in broad rings at the more distal aspect of the scale (Fig. 2G). In contrast, in homozygous csf1ra mutant fish, on average ~10% of scales have TRAP staining, suggesting that some osteoclasts can still be generated in the absence of csf1ra (Fig. 2I). TRAP staining was not drastically affected in scales from csf1rb mutants (~50%); however, double mutants show little TRAP activity (<5% Fig. 2I). Scales containing lateral line canals were not selected, as these have been shown to have high levels of reworking and osteoclast numbers (Wada et al., 2014).

As osteoclasts function to resorb mineralized tissue, we quantified matrix resorption in fish scales by staining mineralized tissue with von Kossa staining and measuring the area of unstained material at the focus of the scale (Fig. 2H). Unstained pits or tracks in the scale, which correspond to common sites of TRAP staining in wild-type fish, are assumed to represent sites of osteoclast activity (de Vrieze et al., 2014). Evidence of resorption was detected in 99.7% of wild-type scales [coefficient of variation (CV) 5.5%]. Relative to each scale’s overall size, an average of 1.2% of the area was resorbed in wild-type scales (Fig. 2J). Using this method, we detected significantly reduced resorption in csf1ra and csf1ra; csf1rb double mutant scales to less than 0.2%; csf1rb mutants also showed a significant, but more subtle, decrease in reworking (Fig. 2J).

Fig. 3. Loss of csf1r parologue function leads to unique and shared phenotypes in the skeleton. MicroCT imaging and quantification of skeletal metrics of the spine of age-matched csf1r paralogues deficient mutant zebrafish noted alterations in the formation of the neural and hemal arches in adult vertebrae (Charles et al., 2017; Chatani et al., 2011). The analysis of osteoclast distribution with our Tg[ctsk:DsRed] transgenic line shows localization of mature, resorbing osteoclasts to the forming vertebral neural and hemal arches (Fig. S5). Isolated vertebrae indicate that both csf1r paralogues are expressed in these structures (Fig. 1D,E). The proportion of DsRed-expressing cells in csf1ra mutant vertebrae was reduced in contrast to the csf1rb mutant vertebrae which still retained wild-type levels of ctsk+ cells. This is easiest seen in juveniles (2 weeks post-fertilization, Fig. S5A-C), however in 4-month-old individuals the number of osteoclasts on the formed centra and arches was extremely variable (Fig. S5D-F). This is consistent with what has been found in Csfr1 mutant knock-out mice, where residual macrophage populations were found, indicating that other growth factors can partially compensate for loss of Csfr1 (Dai et al., 2002).

Using microcomputed tomography (µCT), we quantified metrics of vertebral shape and density in adult mutant fish (Charles et al., 2017) (Fig. 3). As shown previously (Charles et al., 2017; Chatani et al., 2011), csf1ra mutants show an observable alteration in the shape of the vertebrae and the pitch of the arches (Fig. 3B). In contrast, csf1rb mutants show vertebrae and arches that are morphologically similar to those of wild-type siblings (Fig. 3A, C). Further analysis showed that the angles of the arches in csf1rb mutants were significantly increased (Fig. 3J). In contrast, the arch angles in csf1ra and double mutant fish were decreased. Although arch angle and area of the arches are not completely independent measures, the area of arches remains unaffected in either single mutant (Fig. 3H, red, K). However, the combined action of csf1ra and csf1rb has a significant effect on the arch area, suggesting synergistic effects of deficiencies of csf1r paralogues on this
phenotype (Fig. 3I,K). These data suggest that csf1r paralogues may have nuanced effects in regulating vertebral morphology.

Increased bone density can occur in the context of decreased osteoclastic activity. In this regard, we found that csf1ra and csf1rb deficiency results in a reduction in bone volume, not seen in csf1ra mutants (Fig. 3F). Furthermore, loss of csf1ra leads to an increase in the radius of the vertebral body, while csf1rb mutants are unaffected (Fig. 3G). Thus, both paralogues are required for the regulation of shape and structure of the skeleton, with shared and unique structural components that are sensitive to alterations in their function.

Synergistic effects of csf1r paralogues on zebrafish dentition

Teeth are a component of the vertebrate dermal skeleton. Zebrafish develop teeth in a predictable pattern along the ventral/peripheral aspect of the last ceratobranchial arch (Fig. S6A) (Huysseune and Witten, 2006). With growth, the arch increases in size and teeth are replaced as new teeth are formed in succession (Fig. S6B). In zebrafish, resorption by osteoclasts at the tooth base is necessary for tooth loss (Huysseune et al., 1998; Witten et al., 2001; Witten and Huysseune, 2009). We reasoned that the dentition might be a sensitive skeletal structure with which to assess the role of csf1r in osteoclastic bone remodeling.

In staining for osteoclasts present in the mature dentition of adult fish, we observed intense TRAP staining in the arch up to the base of individual teeth (Fig. S6C). The presence of osteoclasts in this region was confirmed using the Tg[ctsk:DsRed] transgenic line, where osteoclasts were localized to areas of high bone deposition, as shown by calcein staining (Fig. 5D-F). Despite the abundance of osteoclasts in this region and variation in numbers in single mutants (Fig. 4G-J), whole-mount skeletal staining showed little change in the number or shape of teeth in csf1ra or csf1rb single mutants (Fig. 4B,C,E). However, changes in the retention of attachment bone can be seen at the base of the forming teeth (Fig. 4B,C, arrows). Quite strikingly, csf1ra;csf1rb double mutants show extreme polyodontia, with up to 20 teeth on a single arch (Fig. 4D-F). The teeth within these extended dentitions remain normal in shape and pattern, forming bouquets of teeth aligned in ventral-dorsal tracts (Fig. 4F, line). Thus, these findings suggest that the normal homeostasis of the zebrafish dentition is quite sensitive to loss of csf1r receptor activity, showing shared function of both receptors in remodeling of bone surrounding and ankylosing to the formed teeth, having synergistic functions in regulating tooth number in mature fish.

Identification of a unique regulatory enhancer of csf1ra driving novel gene function

The shared activity observed in regulation of the skeleton is in contrast to the early differences in expression observed, as well as to the overall pigmentation phenotype between the mutants analyzed (Fig. 2B,C). The required function for csf1ra in osteoclast formation as well as microglia distribution (Oosterhof et al., 2018), we hypothesized that there would be a specific and evolutionary unique gain of cis-regulatory control in csf1ra, shared among all teleost fishes, but unique from other vertebrates (Fig. 5A). Making use of the 11-way genomic evolutionary rate profiling (GERP) track in Ensembl, we identified a 51 bp conserved non-coding element...
(CNE) within csf1ra that is not seen in csf1rb (Fig. 5B; Fig. S7). We did not find evidence of this sequence in human or other amniote genomes, suggesting it is unique to teleosts (Fig. 5B); this is consistent with a gain of regulation in csf1ra (neo-functionalization, Fig. 5A). The identified CNE is conserved throughout evolution of teleost fishes with a broad order-specific region in Danio species flanking a highly conserved core region shared by all teleosts (Fig. 5C, Fig. S7). When the 50 bp CNE is placed upstream of an E1b minimal promoter (Fig. 5D), we find that it is sufficient to drive reporter expression within stellate pigment cells in clones of injected wild-type fish (P0; Fig. 5E) as well as in identified transgenic lines (F1, Fig. 5F,G).

**DISCUSSION**

Diversity in skeletal shape among vertebrates provides an opportunity to define shared and dissimilar mechanisms through which skeletal remodeling occurs. Previous genetic analyses in the zebrafish and medaka have shown an essential role for csf1ra in regulating osteoclast development and activity (Chatani et al., 2011; Mantoku et al., 2016). In zebrafish, however, the effect on skeletal development and remodeling was slight (Chatani et al., 2011). Such a limited skeletal phenotype raises questions about the importance of Csfr1 signaling in regulating skeletal form of the zebrafish, as well as the applicability of teleost fish as a model for studying osteoclast-mediated skeletal remodeling in mammals. Here, we identify broad retention of a second csf1r gene within teleost lineages. We identify mutants in both csf1r paralogues in the zebrafish, and define specific and shared functions of these genes in regulating post-embryonic development.

**Functional analysis of a csf1rb paralogue and the broader role of csf1r function in skeletal remodeling**

We confirmed the known phenotype of csf1ra-deficient zebrafish: having reduced numbers of osteoclasts resulting in altered vertebral arch shape and increased bone mass (Charles et al., 2017; Chatani et al., 2011). However, for the most part, the skeletal shape does not seem overly affected in this mutant in comparison with the osteopetrotic phenotypes observed in rodent models of csf1 or csf1r deficiencies (Cotton and Gaines, 1974; Dobbins et al., 2002; Van Wesenbeeck et al., 2002). On finding a potential csf1rb paralogue in zebrafish, we asked whether loss of both copies of csf1r could better reflect shared roles in osteoclast function in vertebrate post-embryonic development. Each paralogue has quite distinct expression patterns early in development, suggesting diversification of regulation (Fig. 1). Although we do not find evidence for csf1rb being expressed during early development of hematopoietic cells, functionally both paralogues are essential for osteoclast function and expression is seen in mature osteoclasts.

![Fig. 5. Evolution of a unique enhancer controlling csf1ra expression in pigmentation and its variation.](image-url)
suggesting *csf1rb* is active in these lineages at later time points in juvenile development.

Overall, *csf1rb* shares functions with *csf1ra* in remodeling of the skeleton. Often, their activities are combinatorial with additive and synergistic effects. Patterning of the spine is particularly interesting as the effects of each paralogue can be detailed by variation in vertebrae characteristics. These phenotypes are consistent with the localization of osteoclasts on the arches during development (Fig. S5). We notice a difference in effect of *csf1ra* in maintaining resting numbers of osteoclasts compared with *csf1rb*. However, we find that the number of osteoclasts is highly variable in wild-type adults; thus, it is difficult to assign specific roles to each paralogue in terms of effect on number or localization of osteoclast populations on bone (Fig. S5). An analysis of the differential role of *csf1r* paralogues on microglia during development suggest regulation of cell migration may play a large role in the observed phenotype (Herbomel et al., 2001; Oosterhof et al., 2018); however, this is not specifically tested here. The genetic dissociation between neural arch and vertebral phenotypes has also been observed in a medaka osteopetrosis model (To et al., 2015). The modular effect of *csf1r* genes is intriguing as it would allow different regulation of these areas during development through functional variation of specific Csf1R receptors.

### The essential role of osteoclasts in shaping the dentition and tooth replacement

Of the skeletal defects observed in *csf1ra;csf1rb* deficient zebrafish, the effects on tooth number are quite striking. Either *csf1ra* or *csf1rb* function is sufficient to complement the loss of the other paralogue to permit adequate removal of older teeth as new teeth are formed. The bouquet formation of the dentition in the double knockout is suggestive that osteoclasts are essential for clearance of successive teeth in a series. Similarly in medaka, a broad inhibition of osteoclast function leads to the formation of supernumerary teeth (Mantoku et al., 2016; To et al., 2015). In teleost fishes, this manifestation appears to be a failure of tooth shedding. Mice do not show tooth replacement, therefore comparable effects on dental patterning cannot be addressed. However, an important manifestation of osteoclast deficiency in mice is the failure of tooth eruption (Cotton and Gaines, 1974; Dobbins et al., 2002; Tiffée et al., 1999; Van Wesenbeeck et al., 2002). Our data detail the essential role of osteoclast function for normal tooth replacement mechanisms in the zebrafish. The synergistic effects of *csf1r* paralogue functions on the retention of teeth reveal an intricate skeletal patterning process yet to be uncovered in analyses of natural variants or in previously identified experimental mutants.

### The evolution of a key role for *csf1ra* in pigment pattern diversification in teleost fishes

Teleost fishes express a broad array of pigmented varieties and pigmentation patterns are often under intense selection and are key components of fish speciation (Irion et al., 2016; Salis et al., 2019). Csf1ra function in the zebrafish was first identified through its action in regulating pigment pattern of the adult zebrafish (Haffter et al., 1996; Odenthal et al., 1996; Parichy et al., 2000). However, alteration of Csf1r activity in mammals does not lead to pigment deficiencies. Whereas we find both *csf1r* paralogues are necessary for skeletal remodeling with a combined phenotype comparable with that seen in mammals, *csf1ra* has a unique role in xanthophore differentiation and stripe formation (Parichy et al., 2000). We find that *csf1rb* mutants do not show comparable pigmentation phenotypes to *csf1ra* mutants and do not show any increased effect in combination with *csf1ra* deficiency. In support of a separate role for *csf1ra* in regulating pigmentation, we find that *csf1ra* is uniquely expressed in early neural crest lineages, whereas *csf1rb* is not detected in these cells (Fig. 1). Thus, the function of *csf1ra* in neural crest differentiation and pigment pattern formation in fishes is a derived function for this gene. Concordantly, we identify a CNE within zebrafish *csf1ra* that is unique to that paralogue and highly conserved among teleosts. The CNE is not found among non-teleost vertebrates, suggesting it arose after the whole-genome duplication. The conserved CNE is sufficient to regulate expression in pigment cell during zebrafish development, suggesting this is a component of the specialization of *csf1ra*.

In diverse teleost fishes, *csf1ra* has been implicated in functional variation in pigmentation (Parichy and Johnson, 2001; Quigley et al., 2005). The pivotal role during pigment pattern formation is thought to drive diversification of the *csf1ra* gene in these groups (Salzburger et al., 2007). However, the essential role of *csf1ra* in regulating skeletogenesis would be expected to constrain the capacity of this gene in driving pigment diversity in evolution. Our finding of an essential shared role for *csf1rb* in skeletal remodeling in zebrafish suggests that sub-functionalization of *csf1r* paralogues provides buffering of function and thus the means for *csf1ra* to evolve new roles in neural crest and pigment pattern differentiation (Fig. 6), consistent with the broad retention of pigment-related gene paralogues in evolution of teleosts (Braasch...
et al., 2009; Lorin et al., 2018). We provide here the first molecular and functional evidence of this neo-functionalization.

**MATERIALS AND METHODS**

**Phylogenetic analysis**

The phylogenetic relationship between zebrafish csf1ra and csf1rb paralogues and their corresponding vertebrate orthologues, the amino acid sequences of Astyanax mexicanus, Clupea harengus, Chlamyphorogdon idellie, Cyprinodon variegatus, Danio rerio, Exos lucius, Homo sapiens, Latimeria chalumnae, Mus musculus, Oreochromis niloticus, Orzyias latipes, Poecilia reticulata, Pygocentrus nattereri, Salmo salar, Sinocyclocheilus anshuensis, Sinocyclocheilus grahami, Sinocyclocheilus rhinoceros, Takifugu rubripes and Xenopus tropicalis were retrieved from NCBI and ENSEMBL databases. The same strategy described before was used to construct a phylogenetic tree. Accession numbers of protein sequences used in the analysis are listed in Table S1.

**Generation of cathepsin K osteoclast reporter line**

We used a 3 kb upstream region of the medaka (Oryzias latipes) cathepsin K gene, previously shown to drive expression in osteoclasts in zebrafish (Chatani et al., 2011), and cloned it into the p5E-MCS plasmid (228, Tol2kit; Kwan et al., 2007) in front of the DsRed sequence flanked by FRT sites. To enable simultaneous generation of an osteoclast reporter and an osteoclast-specific line driving Cre expression, sequence encoding Cre (from the plasmid pOG231; O’Gorman et al., 1997) was cloned in a middle entry gateway plasmid (pENTR/D-TOPO, Life Technologies). For generation of an expression reporter construct, we used the multisite Gateway system in which the 3’ entry plasmid (302, p3E-polYA) and the destination vector (395, pDestTol2CG2) were obtained from the Tol2kit (Kwan et al., 2007). As a marker for transgenesis, the pDestTol2CG2 destination vector contains a cmlc2 promoter driving GFP expression in the heart. The construct was made via recombination from gateway-based clones using LR clonase II (Life Technologies).

To generate transgenic fish, plasmid DNA (10 ng/μl) and transposase RNA (10 ng/μl) were injected into one-cell stage zebrafish embryos. These animals were raised to adulthood and outcrossed to Tuebingen (Tue) wild-type fish. Progeny were screened for cardiac GFP expression at 48 hpf and raised to establish the Tg[Ola.ctsk:FRT-DsRed-STOP-FRT,Cre, cmlc2: GFP] transgenic line (referred to here as Tg[ctsk:DsRed]). Tg[ctsk:DsRed] transgenic fish were crossed with csf1ra<sup>−/−</sup> and csf1rb<sup>−/−</sup> to visualize the cathepsin K-expressing cell mutants. Progeny were stained overnight with 100 μg/ml calcein (Sigma) solution in water as counterstain for mineralized matrix.

**Zebrafish husbandry and identification of mutant lines**

Zebrafish strains were maintained under standard conditions (Westerfield, 1993) in compliance with internal regulatory review at Boston Children’s Hospital. At the indicated age, fish were euthanized with 20% MS-222 and scales were collected from the anterior area of either side of the trunk.

The csf1ra<sup>−/−</sup> mutant line was isolated in the Harris lab in an ENU mutagenesis screen. The mutation in csf1ra was identified by a candidate gene approach due to similarity to panther and salz und pfeiffer mutants. A sense mutation in csf1ra was detected in cDNA from homozygous 4 dpf csf1ra<sup>−/−</sup> mutant larvae amplified using two primer combinations [csf1ra-f1 (ACGGTCAGCTGACTTCTTC) with csf1ra-r2 (CTGTCAGGCTGTCGCTGC)] and csf1ra-r2 (CGACCTCTCAACTTCTGC) to cover a large portion of the csf1ra-coding region (NM_131672.1). For genotyping of individuals from genomic DNA, we used primers csf1ra<sub>g</sub>-f1 (CGGGTCTCTGGAATTTGT) and csf1ra<sub>g</sub>-r1 (TGGCTAATTACCCAAACTTG). The csf1ra mutation is consistently linked to the pigment phenotype.

CRISPR-Cas9 technology and specific guide RNAs (gRNAs) designed against csf1rb were used for the generation of mutants. Specifically, five gRNAs were designed using CHOPCHOP (Montague et al., 2014) (chopchop.cbu.ubc.no) targeting csf1rb exon 3 (gRNA1-GGTCTCTGAGAGAGAAGACAG, gRNA2-GACGTATACTTACCTCC-AGG, gRNA3-GCCCTGGATGATGATATGCTG, gRNA4-GATACCTCGG-GATCCCGAGGACG and gRNA5-GTCGTGATGATGGTGACATGG). Guide RNAs were synthesized as described by Gagnon et al. (2014). Briefly, templates for gRNA transcription were generated by annealing gene-specific oligonucleotides containing the T7 promoter sequence (5’-TAATAC-GACTCACTATAGGG-3’), the target site without the protospacer adjacent motif (PAM) sequence and a complementary region to a constant oligonucleotide encoding the reverse complement of the tracrRNA tail. Guide RNAs were transcribed using the MEGAtscript T7 kit (Ambion). Cas9 mRNA was synthesized from a linearized pcDNA3-nCas9 plasmid (Jao et al., 2013) using the mMESSAGE mMACHINE SP6 kit (Ambion). To generate mutant fish, a pool of the five gRNAs (100 ng/μl) and Cas9 mRNA (500 ng/μl) were injected into one-cell stage zebrafish AB wild-type embryos.

CRISPR-induced mutations were identified through amplification of a 743 bp genomic region containing csf1rb exon 3 using primers 5’- AACAGATCCAGACAAAGGC-3’ and 5’-CGTCCGATTGACCAGCGC-3’ followed by T7 endonuclease digestion. To establish mutant lines, injected fish were grown to adulthood and outcrossed to AB wild-type fish.

The nature of induced INDELS created by CRISPR/Cas9 was identified in the F1 progeny by sequencing of subcloned PCR products.

**In situ hybridization**

The transgenic lines Tg[ctsk:DsRed] and Tg[sp7:EGFP] (DeLauroir et al., 2010) were used to isolate osteoclasts and osteoblasts, respectively. Protocol was based on methodologies outlined by Mael et al. (2018) and Buettner et al. (2015). Fish were euthanized and tissues samples harvested for isolation of cells. Tissues were digested by vigorous shaking at 30°C in dissociation solution [0.21% trypsin-EDTA (Invitrogen), 16 mg/ml collagenase from Clostridium histolyticum, type IA (Sigma)] for 15 min. The cell suspension was washed with Dulbecco’s Modified Eagle Media (DMEM) supplemented with 10% of fetal bovine serum (FBS, Invitrogen) and filtered on a 40 μm nylon mesh. The filtered cell suspension was layered on Histopaque-1077 (Sigma) at a 2:1 ratio and centrifuged at 400 g, without acceleration or break, for 30 min in a swinging bucket rotor. Both the ring and the pellet were collected in PBS+2%FBS. Immediately prior to sorting, cells were stained with 1X SYTOX Blue (Invitrogen). After doublet exclusion, live cells were identified by the absence of SYTOX Blue staining. EGF or DsRed-positive cells were separated on an FACS Aria II cell sorter (BD Biosciences) using an 85 μm nozzle at a flow rate of ~2000 events/s. For RNA studies, sorted cells were collected directly in RNA lysis buffer for subsequent RNA extraction.

For analysis of DNA content, after isolation cells were counted and DNA content was determined by FACS Aria II cell sorter.
Sakamoto, 2001). Cells were washed in PBS+2% FBS prior to FACS analysis. After removal of doublings, DsRed+ or EGFP+ cells we identified and analyzed according to Hoechst 33342 expression. A Hoechst 33342 histogram was plotted and sp7:EGFP+ cells were used to determine gating strategy for mononuclear cells.

**RNA isolation and quantitative RT-PCR**

Scales from adult fish were collected in Trizol reagent (Life Technology), the tissue was mechanically homogenized and RNA was extracted using the Direct-zol RNA kit (Zymo Research). Sorted cells were collected directly in RNeasy plus micro kit (Qiagen) lysis buffer and RNA was extracted according to the manufacturer’s recommendation.

Reverse transcription was performed with the Protoscript first strand cDNA synthesis kit (New England Biolabs). Quantitative real-time PCR (qRT-PCR) was performed using Sybr green reagent (Life Technologies) in duplicate. Expression levels were normalized to tubulin. RT-PCR primers used in the analysis are: csf1ra (f) AATCAGAACACCCTGTCG; csf1ra (r) CGACCGAGCTAAGGCACCA; csf1rb (f) TAAAGGCAATGCA-CCGGTG; csf1rb (r) GCAGAAATCTGGTTGACCA; csf1rh (f) CGCCT-TCACATACGTCAAG; csf1rh (r) GGTCAGTGCCCTCTCGTTAC; tubulin (f) CGACCAGGTTAAAGGCCACA; tubulin (r) TCAGATACGTCAGCA, sp7 (f) CGTCGCGTTTTTAGCCGCTTG; sp7 (r) GTACGTTGGTAGATGAACCC.

Teeth from adult (4mpf) and bone mineral density, as performed previously (Charles et al., 2017). Two sets of three seed points were placed manually on the anterior and bone mineral density, as performed previously (Charles et al., 2017). Two sets of three seed points were placed manually on the anterior and posterior faces of the vertebral body. The radius was determined from fitting each set to a circle and calculating the average. The length was defined as the distance between the centers of the fitted circles. The volume was total number of voxels defined as bone by the thresholding algorithm and the angle was measured between the ray connecting the averaged circumference and a second line defined by seed points placed on the arch. With the exception of the angle, all parameters were normalized with the standard length (SL) of each fish. Thresholding was performed equally through all samples. Analysis of variant measures at different thresholds showed comparable data between wild-type and mutant fish. 3D reconstruction was made in Amira 6.0 (FEI) software.

**Generation of csf1ra-CNE overexpression reporter lines**

Using the 11-way genomic evolutionary rate profiling (GERP) track in ENSEMBL, we identified a conserved non-coding element (CNE) within csf1ra. To examine sequence similarities between Danios, Gadus morhua, Takifugu rubripes, Tetraodon nigroviridus, Oreochromis niloticus, Gasterosteus aculeatus, Orzyas latipes, Poecilia formosa, Xiphophorus maculatus and Astyanax mexicanus, genomic csf1ra sequences were retrieved from NCBI (O’Leary et al., 2016) and ENSEMBL (Yates et al., 2016) databases. CNE sequences were aligned using Clustal Omega multiple sequence alignment tool (Madeira et al., 2019). TFbind was used for transcription factors binding site prediction (Tsunoda and Takagi, 1999). Accession numbers of genomic DNA sequences used in the analysis are listed in Table S2.

Enhancer sequences (50 bp) from D. rerio were cloned into the E1b-GFP Tol2 plasmid upstream of the E1b minimal promoter (Birnbaum et al., 2012). To generate transgenic fish, plasmid DNA (10 ng/μl) and transposase RNA (10 ng/μl) were injected into one-cell stage Tue wild-type zebrafish embryos. These animals were raised to adulthood and outcrossed to Tue wild-type fish. Progeny were screened for GFP expression at 24 hpf and raised to establish the transgenic line.

**Statistical analysis**

Statistical analysis was performed using SPSS Statistics 23.0 (IBM). Continuous variables were expressed as mean±s.d. Groups were compared using analysis of variance (ANOVA) test followed by Bonferroni’s multiple comparison test. Statistical significance was set at P<0.05.

**Acknowledgements**

The authors thank Mrs. Althea James for assistance in zebrafish husbandry, the Flow Cytometry Research Facility at Boston Children’s Hospital, and David Parichy in sharing unpublished whole genome data from select Danio species.

**Competing interests**

The authors declare no competing or financial interests.

**Author contributions**


**Funding**

This work was partially supported by the National Institutes of Health (U01 DE024434 to M.P.H., KO8 AR62590 to J.F.C. and GM122471 to D.P.). Deposited in PMC for release after 12 months.

**Supplementary information**

Supplementary information available online at http://dev.biologists.org/lookup/doi/10.1242/dev.181834.supplemental


