EVOLUTION & DEVELOPMENT 18:1, 7–18 (2016)
DOI: 10.1111/ede.12131

Natural bone fragmentation in the blind cave-dwelling fish, *Astyanax mexicanus*: candidate gene identification through integrative comparative genomics

Joshua B. Gross,* Bethany A. Stahl, Amanda K. Powers, and Brian M. Carlson

Department of Biological Sciences, University of Cincinnati, 312 Clifton Court, Cincinnati, Ohio 45221, USA

*Author for correspondence (e-mail: grossja@ucmail.uc.edu)

SUMMARY Animals that colonize dark and nutrient-poor subterranean environments evolve numerous extreme phenotypes. These include dramatic changes to the craniofacial complex, many of which are under genetic control. These phenotypes can demonstrate asymmetric genetic signals wherein a QTL is detected on one side of the face but not the other. The causative gene(s) underlying QTL are difficult to identify with limited genomic resources. We approached this task by searching for candidate genes mediating fragmentation of the third suborbital bone (SO3) directly inferior to the orbit of the eye. We integrated positional genomic information using emerging *Astyanax* resources, and linked these intervals to homologous (syntenic) regions of the *Danio rerio* genome. We identified a discrete, approximately 6 Mb, conserved region wherein the gene causing SO3 fragmentation likely resides. We interrogated this interval for genes demonstrating significant differential expression using mRNA-seq analysis of cave and surface morphs across life history. We then assessed genes with known roles in craniofacial evolution and development based on GO term annotation. Finally, we screened coding sequence alterations in this region, identifying two key genes: *transforming growth factor β3* (*tgfb3*) and *bone morphogenetic protein 4* (*bmp4*). Of these candidates, *tgfb3* is most promising as it demonstrates significant differential expression across multiple stages of development, maps close (<1 Mb) to the fragmentation critical locus, and is implicated in a variety of other animal systems (including humans) in non-syndromic clefting and malformations of the cranial sutures. Both abnormalities are analogous to the failure-to-fuse phenotype that we observe in SO3 fragmentation. This integrative approach will enable discovery of the causative genetic lesions leading to complex craniofacial features analogous to human craniofacial disorders. This work underscores the value of cave-dwelling fish as a powerful evolutionary model of craniofacial disease, and demonstrates the power of integrative system-level studies for informing the genetic basis of craniofacial aberrations in nature.

INTRODUCTION

Natural model systems have helped illuminate the genetic and developmental bases of a variety of widespread but poorly understood phenotypes (Woese et al. 1990). Novel insights into diverse traits—with direct human relevance—have been discovered using equally diverse natural model systems (Albertson et al. 2009). Such phenotypes, including senescence (naked mole rat; Buffenstein 2008), eye degeneration (blind cavefish; O'Quin et al. 2013), body size (domesticated dogs; Sutter et al. 2007), blood pressure regulation (giraffes; Paton et al. 2009), diabetes (dolphins; Venn-Watson and Ridgway 2007), and obsessive-compulsive disorder (deer mice; Korff et al. 2008) capitalize on extreme traits evolving under natural selective pressures.

Our understanding of the evolution of the craniofacial complex has similarly benefited from natural model system research. Genetic studies using Darwin's finches (Mallarino et al. 2011), teleost fish (Hulsey et al. 2007; Miller et al. 2014), and Malagasy primates (Viguier 2004) illustrate the vast diversity and lability of the vertebrate craniofacial skeleton. Most of these modified traits, however, arose under obvious selective pressures imposed by changes in the environment, feeding modes, and/or nutritional source (Kocher 2004; Lieberman 2008). Conversely, craniofacial defects evolving in the absence of obvious selective pressure have received little attention (Gross et al. 2014). This class of craniofacial phenotypes may target genetic loci, shared broadly among vertebrates, which are vulnerable to mutation (Stern and Orgogozo 2009; Martin and Orgogozo 2013). Therefore, investigations into the genetic basis for trait evolution in the wild can augment parallel studies of craniofacial biology using traditional model systems, and identify novel genetic loci mediating aberrant craniofacial phenotypes (Albertson et al. 2009).

We recently characterized a striking craniofacial malformation—asymmetric bone fragmentation—present in natural populations of the cave-dwelling fish *Astyanax mexicanus* (Gross et al. 2014). This species exhibits seemingly spontaneous fragmentations and fusions of the bony series encircling the orbit of the eye (Yamamoto et al. 2003). These aberrations were

identified shortly after discovery of the cave-dwelling forms in the 1930s (Alvarez 1947); but, their origin was long assumed to be indirectly linked to the complete absence of an eye in these troglomorphic fish (Mitchell et al. 1977). However, both developmental (Yamamoto et al. 2003) and genetic evidence (Protas et al., 2008; Gross et al. 2014) shows that fragmentation of the third bone in the circumorbital series (SO3) occurs independently of eye loss.

Despite the eye-independent basis of this fragmentation phenotype, it has recurrently evolved in numerous phylogenetically and geographically distinct cave populations (Jeffery 2009). Using a linkage mapping approach, we discovered that an asymmetric (right sided) genetic signal is linked to this phenotype in cavefish drawn from the Pachón cave locality (Gross et al. 2014). Pachón cavefish are geologically isolated (Mitchell et al. 1977; Bradic et al., 2012; Gross 2012) and harbor some of the most extreme troglomorphic characteristics, including albinism (Bilandzija et al. 2013), complete eye loss (Borowsky 2008), and dramatic expansion of the taste (Kowalko et al. 2013a) and lateral line systems (Yoshizawa et al. 2013). Therefore, this natural cave population is ideal for further analyses of this trait.

Until recently, genomic resources were limited for this model system. In the past two years, however, two transcriptome analyses (Gross et al. 2013; Hinaux et al. 2013) and a draft genome have become publicly available. Here, we leverage these enhanced genomic resources to nominate candidate genes mediating SO3 fragmentation in our system. We previously identified deeply conserved genomic synteny between *Astyanax* and zebrafish (Gross et al. 2008, 2013), despite approximately 140 My divergence between these species (Lundberg 1993). Here, we report a highly conserved approximately 6.1 Mb syntenic region of chromosome 17 in zebrafish encompassing the critical region associated with SO3 bone fragmentation. We identified candidate genes using a combination of gene ontology (GO) term studies and literature searches focused in the predicted genomic region near our peak association marker. Following this step, we analyzed transcriptomic data to evaluate the structure and expression of candidate genes. This high-resolution analysis identified two candidate genes, *transforming growth factor* **β***3* (*tgfb3*) and *bmp4* with previously appreciated roles in craniofacial development and evolution. One of these genes, *tgfb3*, produces analogous phenotypes in other traditional and emerging model systems. Using these combined approaches, we demonstrate the feasibility and utility of next-generation approaches for understanding craniofacial trait evolution in nature.

MATERIALS AND METHODS

Bone fragmentation phenotype, analysis, and QTL mapping

We previously characterized the genetic architecture of SO3 fragmentation number using an experimental F_2 pedigree ($n = 237$ individuals) generated from a surface fish × Pachón cavefish cross (Gross et al. 2014). Within this pedigree, the number of bony elements ranged from 1 to 5 separate elements. Cavefish demonstrate frequent fragmentation (Fig. 1, B, D, and F) which varies in severity (i.e., number of fragmented bony elements) and laterality (i.e., inconsistent fragment numbers on the left and right sides of the cranium; see Gross et al. 2014). However fragmentation is never observed in surface fish (Fig. 1, A, C, and E). We assessed this trait in both a continuous and binary fashion; both methods identified the identical genetic associations (Gross et al. 2014). A genomic scan utilizing a genetic linkage map comprising 175 microsatellite markers, assayed in 237 recombinant hybrid fish (Protas et al. 2008), revealed two QTL associated with fragment number variation at markers 206A and NYU53. The first marker, 206A, produced a significant association using both marker regression (LOD = 5.43) and Haley–Knott (LOD = 5.34) mapping methods, and therefore we decided to pursue this marker. A phenotypic effect study revealed that fragmentation was uniquely associated with the homozygous cave allelic condition, and this genetic effect was recessive to surface fish alleles (Gross et al. 2014). Moreover, the genetic signal was asymmetric. Cavefish frequently display asymmetries in nature; however, surface forms appear to be uniformly symmetric at the population level (Gross et al. 2014).

Light micrographs of adult (>2 years) surface and Pachòn cavefish were captured using a Leica M205FA series microscope at 7.8× magnification (Fig. 1, A and B). Montage z-stacks were compiled using Leica software (Leica Application Suite v3.8). To visualize bony phenotypes, adult surface and Pachòn were subjected to high-resolution micro-computed tomography (μCT; CCHMC Imaging Research Center) using an Imtek MicroCAT II scanner (Siemens). Each high-resolution scan captured a 19–20 micron isotropic slice from the most anterior skull to the caudal-most aspect of the operculum of each specimen. For each specimen, the number of raw slices ranged from 800 to 1000, which were reconstructed into DICOM files using COBRA software (v6.9.42; Exxim Corporation). The DICOM files were then imported into Amira software (v5.5) where they were constructed into 3D surface renderings (Fig. 1, C–F) using the VolRen feature.

GO term analyses

To identify prospective genes associated with craniofacial development, we evaluated all genes within the syntenic regions of the *Astyanax* draft genome and in *Danio rerio* using gene ontology (GO) terms. Based on our comparative genomic mapping, we utilized BioMart (v0.8; Kasprzyk 2011) to collect all genes and their associated GO terms within these regions of interest. GO analyses included the 101 genes present on *Astyanax* Scaffold KB882233.1 (approx. 6.5 Mb), and also the approximately 5 Mb (477094444–53930577 bp) and approximately

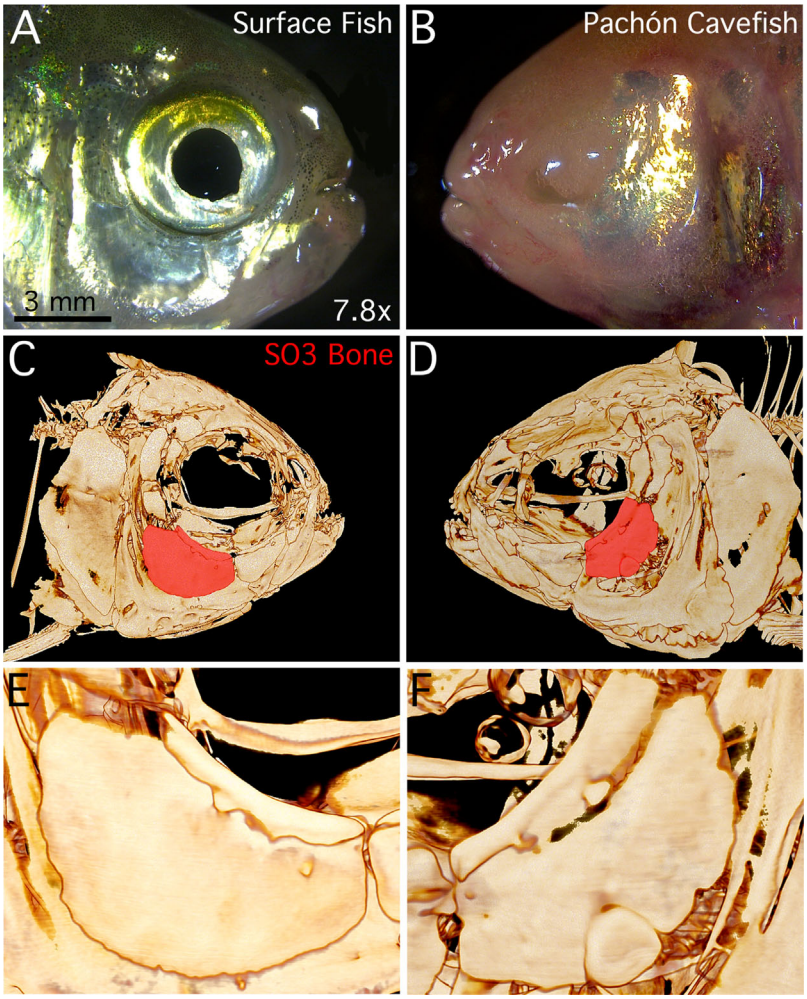

Fig. 1. Surface- and cave-dwelling forms of *Astyanax mexicanus* harbor distinct craniofacial morphologies. Surface morphs (A) demonstrate typical cranial features including a large, spherical orbit housing the eye (C). Cavefish from the limestone Pachón cave in NE Mexico (B) show a number of extreme features including eye loss, albinism, and a sloping dorsal skull (D). Among the many craniofacial features altered in this cave-dwelling fish is fragmentation of the third suborbital bone (red in D; close-up in F). This bone is unfragmented, i.e., exists as a singular intramembranous bone, in surface dwelling fish (red in C; close up in E). A and B are whole mount light micrographs of two representative individuals, collected with a Leica M205FA stereoscope at 7.8× magnification. C–F are surface-rendered 3D skull created using Amira software (v.5.5) from μCT scans of each specimen (CCHMC Imaging Research Center). Scale = 3 mm.

1.5 Mb (23846666–3857276 bp) syntenic intervals on *Danio* chromosome 17, comprises 145 and 22 genes, respectively. This approach yielded a total of 267 terms in *Astyanax* and 579 terms from *Danio*. Within these gene sets, we mined for craniofacial- and bone-related GO terms, including: "ossification, skull, face, cranial, cranio, craniofacial, odontogenesis, dentin, bone, bone mineralization, bone reabsorption, bone development, osteoblast, and osteoclast." The *Astyanax* genome is in an early draft form, and therefore is limited in the depth of available annotation. Our approach of evaluating the GO terms representing both *Astyanax* and *Danio* genomes permitted the highest sensitivity in detecting genes related to craniofacial bone fragmentation.

Developmental mRNA-seq and qPCR analyses

To interrogate developmental gene expression differences, mRNA sequencing was carried out on stage-matched Pachón cave- and surface-dwelling embryos at each of four developmental time points—10 hpf, 24 hpf, 1.5 dpf, and 3 dpf—and one juvenile stage. mRNA pools were extracted from pooled sets of individuals ($n = 50$) at each developmental stage using the RNeasy Mini Kit (Qiagen). RNA from juveniles was collected from three surface and cavefish at approximately 4 months of age. Highly purified RNA was reconstituted to a final volume of 50 μl at a concentration of 40 μg/ml. Library generation was carried out at the Cincinnati Children's Hospital Sequencing Core using Illumina TruSeq (v.2) kit. Each sample was sequenced in triplicate for both sets of stage-matched morphotypes, across all four developmental stages and in duplicate for one juvenile stage (total = 28 sequencing runs). Sequencing was performed to a depth of up to 10 million, 50 bp long, unpaired reads.

To determine gene expression profiles for each candidate gene, we mapped Illumina surface and cavefish reads to the draft genome template, inclusive of all craniofacial-related genes, using the QSeq module within the program ArrayStar (DNASTAR.v.11.0). This template file comprises the complete set of cDNA transcript sequences from the current *Astyanax* draft genome build (AstMex102v75) (McGaugh et al. 2014). Within this program, surface and cave sequencing sets were

subdivided using the replicates at each developmental stage (10 hpf, 24 hpf, 1.5 dpf, 3 dpf, and juvenile). Reads were then quantified as a measure of gene expression utilizing the RPKM normalization strategy to calculate the normalized read count based upon the replicate mean (Mortazavi et al. 2008). Additionally, significant levels of differential expression between morphotypes, at each stage of development, were determined using a Student's *t*-test with FDR (Benjamini and Hochberg 1995) using ArrayStar software (DNASTAR.v.11.0).

Quantitative PCR analyses were performed as previously described (Gross and Wilkens 2013) using a Bio-Rad MiniOpticon light cycler (Bio-Rad). Each cDNA template was generated using the Transcriptor RT kit (Roche). Total RNA was extracted from 3 dpf Pachón cavefish ($n = 50$) and surface fish ($n = 50$) embryos using the RNeasy RNA extraction kit (Qiagen). Isolated RNA (1 μg) was hybridized to Oligo dT primers (Invitrogen) at 65°C, cooled on ice for 5 min, after which were added 4 μl 5 × RT buffer, 0.5 μl Protector RNase Inhibitor, 2 μl dNTP mixture, and 0.5 μl Transcriptor RT per reaction. This reaction mixture was incubated at 50°C for 1 h and then inactivated at 85°C for 5 min. Each template was tested in sextuplet across three separate experiments. The cycling threshold (C[t]) values were collected using cDNA pools as templates and normalized to the expression of the housekeeping gene *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase; primers: forward=5′-TGTGTCCGTGGTGGATCTTA-3′, reverse=5′-TGTCGCCAATGAAGTCAGAG-3′). The primers used to measure *tgfb3* expression were the following: *tgfb3* forward primer: 5′-AAAGCTGTGGCCAGGATAAC-3′; *tgfb3* reverse primer: 5′-TAGGGCAGGTCGTTGTTTTC-3′. These primers amplified a 100 bp fragment in surface and Pachón cavefish cDNA. The primers used to measure *bmp4* expression were the following: *bmp4* forward primer: 5′-AGCCGAGCCAACACTGTAAG-3′; *bmp4* reverse primer: 5′-AAAGATGAAGCGCAGTGGAG-3′. These primers amplified a 102 bp fragment in surface and Pachón cavefish cDNA. Cycling parameters were as follows: step 1—95°C for 30 s, step 2—95°C for 5 s, step 3—55.1°C for 10 s, plate read, go to step 2 for 39 additional cycles. Experiments were carried out using the EvaGreen supermix dye (Bio-Rad). All data were collected and analyzed using the Bio-Rad CFX Manager v.1.6 software program (Bio-Rad) as previously described (Gross and Wilkens 2013; Gross et al. 2013). Qualitative expression measurements were carried out (Fig. 3F) utilizing the same RNA pools used for RNA-seq studies. Briefly, 1 μg of total RNA was reverse transcribed to cDNA using the Transcriptor RT kit (Roche), PCR amplified using primers amplifying a fragment of *GAPDH* (680 bp), *tgfb3* (640 bp), and *bmp4* (612 bp) in both 3 dpf and 4-month-old juveniles. The primers used to amplify the *GAPDH* fragment were the following: *GAPDH* forward primer: 5′-TCCATTCCAAGAAGGTGGAG-3′ and *GAPDH* reverse primer: 5′-CAAGCGGACAGTCAAGTCAA-3′. The primers used to amplify the *tgfb3* fragment were the following: *tgfb3* forward primer: 5′-GCTGTGGCCAGGATAACACT-3′ and *tgfb3* reverse primer: 5′-TTGGTGTCCAGTGCCCTTTT-3′. The primers used to amplify the *bmp4* fragment were the following: *bmp4* forward primer: 5′-CGGCTTCTGGACACAAGACT-3′ and *bmp4* reverse primer: 5′-GCAGCCCTCCACTACCATTT-3′. Using standard electrophoresis, PCR fragments were visualized in a 2% agarose gel.

Transcript structural analyses

For sequence analyses, the reads generated from both surface and cavefish pools were aligned to a template composed of the full set of sequences (approx. 10,000 scaffolds) from the current draft genome (McGaugh et al. 2014) using the SeqMan NGen assembly software package with default parameters for Illumina reads <50 bp (DNASTAR.v.11.0). Using the "Multi-Plex" option, sequences from each fastq file were tagged with the appropriate "Surface" or "Cave" prefix to denote the origin of each read. The gene contigs for four potential candidate genes, including *tgfb3, bmp4, bmp2a*, and *col12a1a*, were each interrogated for any prospective sequence alterations segregating between the surface- and cave-dwelling fishes with SeqMan Pro software (DNASTAR.v.11.2a). Although the genes *bmp2a* and *col12a1a* both reside on chromosome 17 in *Danio rerio*, they were discarded from further analyses because they reside outside the critical region of the fragmentation locus. We mapped our complete RNA-seq data (approx. 280 million sequencing reads) enabling direct analyses of SNP variants and indels, using the "SNP Report" tool in SeqMan Pro. If a mutation was identified, it was then evaluated to determine the nature of the sequence alteration (SNP, indel), the consequence of this mutation on the encoded gene product (synonymous, nonsynonymous, and frameshift), and the extent to which this mutation was shared within and among morphotypes.

Biological materials

Specimens used in this study were derived from natural collections of both cave- and surface-dwelling forms of *A. mexicanus*. Pachón cave embryos represented the offspring of a full sibling Pachón cavefish cross. These siblings were members of a large family bred from wild-caught individuals, collected from the Pachón cave (near Tamaulipas, Mexico) in 2001. Surface fish embryos used in this study were similarly the offspring of a full sibling cross. These siblings were offspring of a female parent who was two generations removed from the wild (i.e., her grandparents were collected at Arroyo Sarco, at the Río Sabinos drainage, Mexico). The male parent of these siblings was one generation removed from the wild, whose parents were collected in the Río Valles drainage. All specimens were generously provided to us by Dr. Richard Borowsky (New York University).

RESULTS

Fragmentation of the third suborbital bone harbors a genetic basis

Developmental and genetic studies affirmed dermal bone fragmentation in cavefish evolved independently from visual system loss (Yamamoto et al. 2003; Gross et al. 2014). A QTL analysis identified two linked microsatellite markers, 206A and NYU53, residing on linkage groups 23 and 26, respectively (Gross et al. 2014). The 206A locus, which accounts for 12.9% of the phenotypic variance of this trait (Protas 2005), was identified with three scan-one mapping methods: marker regression (Kearsey and Hyne 1994), expectation maximum (Xu 2010), and Haley–Knott (Haley and Knott 1992) using the mapping software program, R/qtl (Broman et al. 2003). Furthermore, this trait is significantly associated with both quantitative (number of SO3 bony elements) and qualitative (presence absence of fragmentation) scoring on the right side of the skull (LOD 5.6; Gross et al. 2014). Interestingly, all associations disappeared when scoring for fragmentation on the left side of the skull suggesting that this trait demonstrates variable/reduced penetrance across the left-right axis (Gross et al. 2014), as has been reported for a variety of craniofacial disorders (Mooney et al., 1994; Muenke 1994; Bourgeois et al., 1998; Wilkie and Morriss-Kay 2001).

Astyanax genomic scaffold KB882233.1 anchors strongly to *Danio rerio* chromosome 17

To identify putative candidate genes mediating bone fragmentation, we performed comparative genomic mapping of the linked microsatellite marker, 206A, through alignment to the *Astyanax* and *Danio* genomes. Our attempts to align flanking sequence to this microsatellite (653 bp) directly to the *Danio rerio* genome were unsuccessful, implying substantial sequence differences have accumulated at this locus as the two species diverged. However, a blast analysis of this sequence to the draft assembly of the *Astyanax* genome revealed a single, strong match (E-value = 0) to genome scaffold KB882233.1.

A prior analysis suggested the linkage group on which 206A resides (integrated linkage group 23) anchored strongly to *Danio rerio* chromosome 17. Here, we report six markers, distributed along approximately 37 cM, comprising an approximately 32 Mb syntenic block in *Danio*. To reduce the genomic interval in which to search for candidate genes, we focused only on those genes present on the same scaffold as marker 206A. Sixty-five of the genes present on this approximately 6.5 Mb scaffold in *Astyanax* are orthologous to genes on *Danio rerio* chromosome 17. Although 10 markers were distributed within a small (approx. 1.5 Mb) proximal region, the vast majority ($n = 54$) were distributed within a stretch of approximately 6.2 Mb on distal chromosome 17 (Fig. 2).

This *Astyanax* genomic scaffold is occupied by genes covering roughly the same genomic distance in the *Danio* genome. We believe this indicates a highly conserved syntenic "block" between scaffold KB882233.1 and the distal portion of chromosome 17 in *Danio rerio*. Thus, the genes residing on this scaffold represent potential causative genes linked to the QTL for SO3 bone fragmentation. Two potential candidates genes, based on syntenic positional information, include *tgfb3* and *bmp4*. The *Astyanax* sequence for *tgfb3* blasts to *Danio* chromosome 17: 52587615 (E-value = 6.6e–170), *bmp4* blasts to chromosome 17: 51114273 (E-value = 2.7e–290). Two additional genes, *bmp2a* which blasts to chromosome 17: 4004629 (E-value = 5.7e–17) and *col12a1a* which blasts to chromosome 17: 50356833 (E-value = 2.7e–95), were explored but not further pursued since neither gene resides on the same genomic scaffold harboring the genetic marker (206A) associated with fragmentation, and both reside outside the critical region of the fragmentation locus.

Comparative genomic mapping

Prior work demonstrated substantial genomic similarity between linkage group 23 in *Astyanax* (upon which our peak genomic marker 206A resides) and *Danio rerio* chromosome 17 (Gross et al. 2008). We sought to further clarify this relationship by performing additional blast identity searches using the draft *Astyanax* genome (Assembly: AstMex102; GenBank Assembly ID GCA_000372685.1) and the current zebrafish genome assembly (Assembly: Zv8; Ensembl). We employed a standard blast protocol wherein query sequences were subjected to local alignment using either *Astyanax* or *Danio* as the DNA database, using the BlastN search tool. In all cases, the search sensitivity was set to default ("near-exact matches").

The original genomic marker associated with right-sided fragmentation of the SO3 bone, 206A, was identified through a microsatellite screen for CA_n dinucleotide repeats probed from a genomic DNA library (Protas et al. 2006). The insert identified alongside this marker harbored 653 bp flanking the microsatellite. Attempts to blast this sequence directly to the zebrafish genome were unsuccessful. However, five flanking markers from the same linkage group as 206A (integrated linkage group 23; Gross et al. 2008) aligned to zebrafish chromosome 17 (Fig. 2), including markers 234C, 207E, and polymorphisms near the genes *pomc*,*otx1*, and *bmp4*. This finding was further strengthened by analysis of the first genomic draft release (>10,000 genomic scaffolds) for *A. mexicanus*.

Blast analyses using the ensembl search engine (ensembl.org/Danio_rerio/blastview) identified Scaffold KB882233.1, upon which the flanking sequence surrounding 206A resides. We then pulled off every predicted gene sequence from this scaffold and blasted them all to the zebrafish genome (Fig. 2, gray lines). The majority of known genes distributed on this scaffold anchored strongly to either a region of approximately

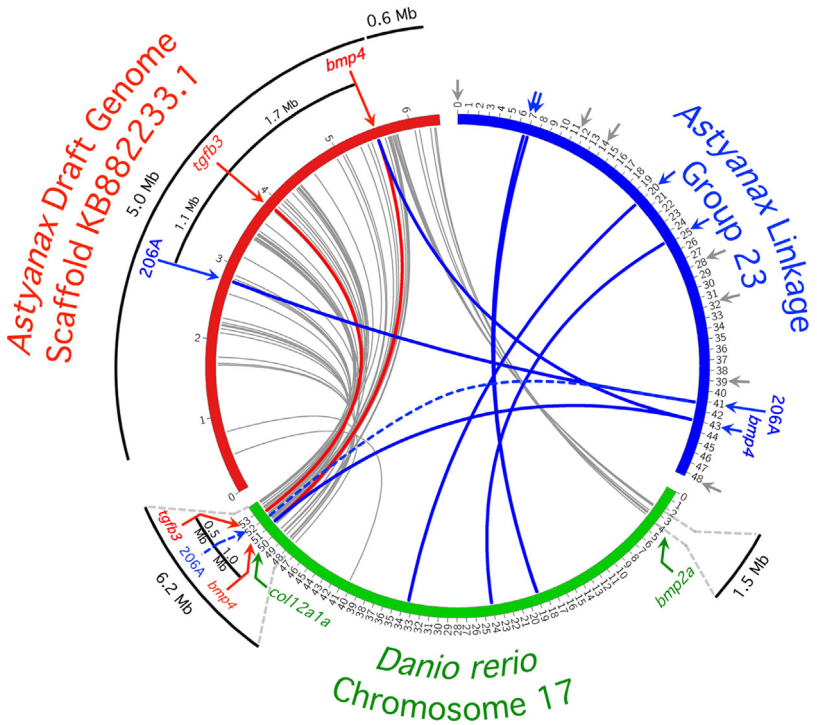

Fig. 2. Emerging genomic resources reveal a syntenic block informing the genomic positions of candidate genes mediating bone fragmentation. A genomic marker associated with SO3 fragmentation in cavefish, 206A, resides on integrated linkage group 23 (Gross et al. 2008). Five of the first-generation map microsatellite markers from this group map directly to *Danio rerio* chromosome 17. Two markers from the critical fragmentation locus (including 206A) are present on *Astyanax* draft genomic scaffold KB882233.1. This approximately 6.5 Mb scaffold anchors strongly to the distal region of chromosome 17 in zebrafish. This syntenic region of chromosome 17 in *Danio rerio* harbors two candidate genes (*tgfb3* and *bmp4*) with previously characterized roles in craniofacial biology. Arrows represent the position of markers along *Astyanax* integrated linkage group 23.

6.2 Mb distal chromosome 17, or approximately 1.5 Mb region of proximal chromosome 17 (Fig. 2). Scaffold KB882233.1, which is approximately 6.5 Mb in length, has therefore retained significant syntenic identity between these two genomes. The incongruous positioning of the two principal syntenic blocks we identified may be attributable to either translocation events as divergence of the two species and/or errors in genome assembly and alignment for either one or both of these systems.

GO term analyses identify genes associated with craniofacial processes

To identify genes involved with craniofacial processes, we investigated the syntenic intervals previously identified with our comparative genomic mapping using GO term analyses. For the most comprehensive analyses, we collected the annotations for genes that reside on *Astyanax* Scaffold KB882233.1 and within the two approximately 6.2 Mb and 1.5 Mb conserved genomic intervals on *Danio rerio* chromosome 17 with BioMart (v.0.8; Kasprzyk 2011). With this approach, we identified 267 GO terms associated with the 101 genes present on *Astyanax* Scaffold KB882233.1, and 579 terms for the 167 genes comprising the two syntenic blocks on chromosome 17 in *Danio*. Here, we interrogated these gene ontology annotations for any terms related to craniofacial processes and bone formation. Interestingly, within this entire data set we only detected one specific craniofacial-related GO term called "regulation of ossification" (GO ID:0030278) associated with the gene *tgfb3* in for both the *Astyanax mexicanus* and *Danio rerio* orthologues. The candidate gene *bmp4* is likely involved in facial bone formation; however, it is not specifically associated with terms related to bone and craniofacial formation in *Danio rerio* or *A. mexicanus*. Instead, the assigned GO terms for *bmp4* are quite broad, e.g., "growth" (GO ID:0040007), "growth factor activity" (GO ID:0008083), and "protein binding" (GO ID:0005515).

Gene expression pattern differences between cave and surface forms

After identifying two candidate genes, we then evaluated dynamic gene expression differences between morphotypes. *Tgfb3* demonstrated one of the most intriguing expression patterns. We discovered significant differential expression between surface (red, Fig. 3A) and cavefish (blue, Fig. 3A) at every developmental stage. Expression differences ranged from approximately 2-fold (at 24 hpf) to approximately 6-fold (at 10 hpf) through the embryonic stages we evaluated. Morphotype-specific expression of *tgfb3* was significantly different at three of the stages, including: 10 hpf ($P = 0.0479$), 24 hpf ($P = 0.0486$), and 3 dpf ($P = 0.0365$). The expression level of *tgfb3* was consistently lower for cavefish with the exception of the first stage analyzed (10 hpf). The relevance of this "switched polarity" of developmental expression is unclear.

Tgfb3 and *bmp4* both demonstrated modest differences in expression between cave and surface fish based on our RNA-seq

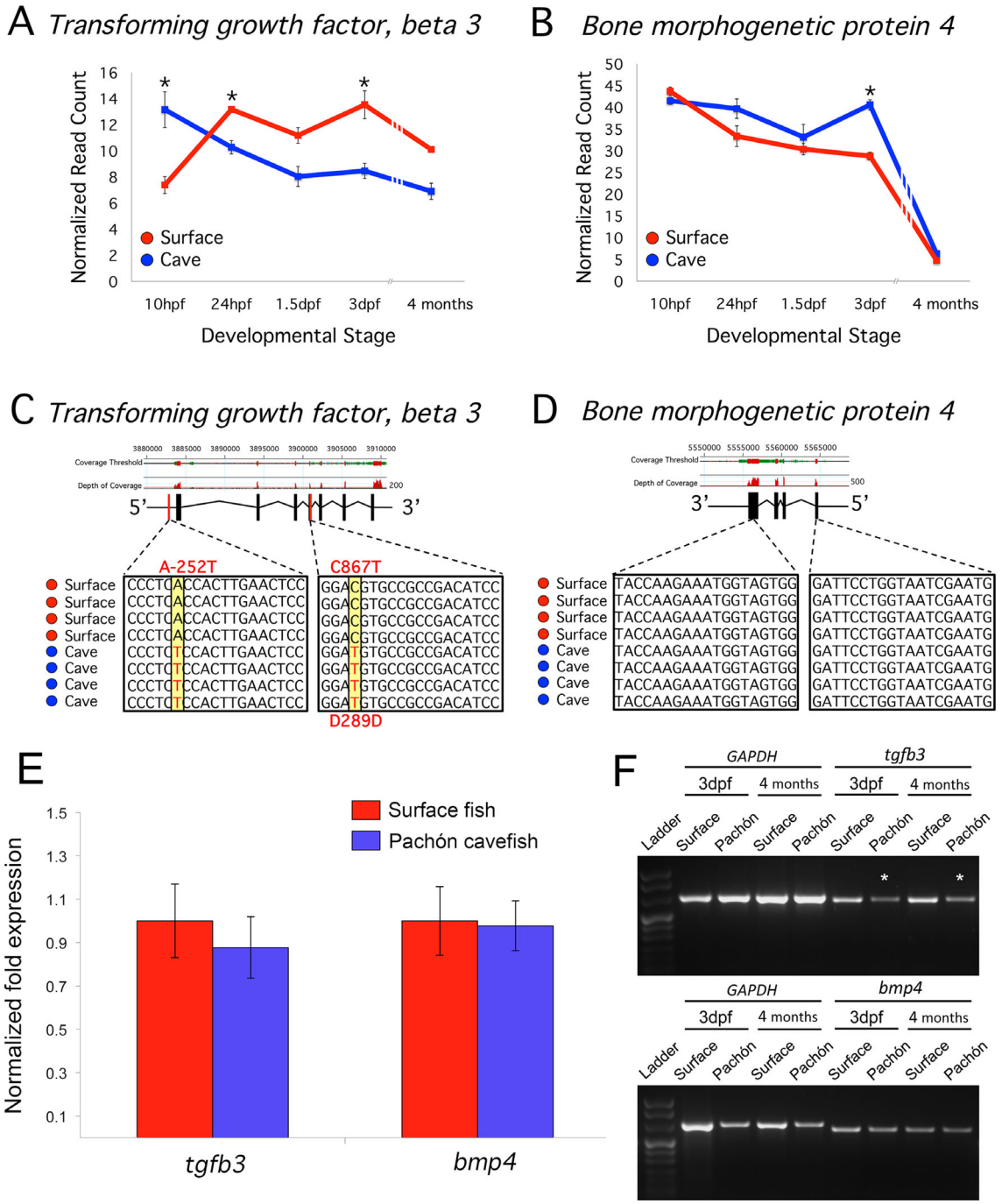

Fig. 3. Expression level and structural analyses of the candidate genes *tgfb3* and *bmp4*. Gene expression analyses demonstrated that *tgfb3* is generally overexpressed in surface fish (red) compared to cave individuals (blue) throughout most of early development (A). In contrast, the candidate gene *bmp4* (B) is expressed at similar levels in both morphotypes. Sequence analysis of *tgfb3* demonstrated a SNP present in the 5′ region upstream of the translation start codon at −252 (SNP depth = 126; SNP% = 74.5%) that segregates between the surface-dwelling (red circles) and Pachón cavefish tetra (blue circles; C). Interestingly, a single nucleotide polymorphism at C867T in exon 5 results in a synonymous change in the D289D amino acid residue (SNP depth = 201; SNP% = 74.4%). The contig coverage threshold and depth of coverage lend strong support to the observed polymorphisms (green = above threshold, 100; red = exceeds threshold, four coverage threshold + two minimum on each strand; thin green). Interestingly, the *bmp4* coding sequence appears to be entirely conserved across both surface and cave-dwelling individuals (D). qPCR analysis of *tgfb3* and *bmp4* at 3 dpf failed to identify statistically significant differences in gene expression (E), however the polarity of differential *tgfb3* expression (lower in Pachón cavefish compared to surface fish) was observed with both qPCR (E) and qualitative PCR analysis (asterisks, F). In A and B, $^{*}P < 0.05$.

studies. For instance, the pattern of *bmp4* expression (Fig. 3B) was largely conserved between morphotypes across each of the four stages analyzed. At 3 dpf, we observed the largest difference in expression of approximately 10-fold higher expression in cavefish compared to surface fish ($P = 0.00748$).

As both genes demonstrated significant differential expression based on RNA-seq at 3 dpf (Fig. 3, A and B), we evaluated mRNA pools from this stage of development using quantitative PCR (qPCR). Interestingly, although the polarity of gene expression difference was validated for *tgfb3*, neither gene demonstrated statistically significant differential expression (Fig. 3E). This could be explained by any of a number of factors. For instance, our RNA-seq data may be more sensitive, and therefore accurate, identifying differences in expression of this gene that cannot be validated using a (perhaps) less-sensitive qPCR approach. Alternatively, the RNA-seq data may be spurious, and the qPCR evidence may be reporting the absence of a significant biological difference in expression at this stage between morphs. Finally, these incongruent results may be attributed to a technical error (e.g., slight differences in developmental staging of 3 dpf embryos for our RNA-seq and qPCR experiments), extremely low levels of expression (as suggested by our RNA-seq results), or variable expression in diverse parts of the developing embryo.

Finally, we performed a qualitative analysis of expression using routine PCR and observed a weaker band in 3 dpf embryos and 4-month-old juveniles (Fig. 3F). We performed the same analysis for *bmp4*, but did not identify a difference at either stage (Fig. 3F). Future expression analyses using in situ hybridization will help clarify its cellular of expression, and determine if *tgfb3* is indeed reduced in developing craniofacial tissues in cave-dwelling forms.

Sequence analyses reveal coding sequence alterations in craniofacial candidate genes

To discover putative coding sequence alterations in our candidate genes, we aligned our sequencing read libraries to the *Astyanax* draft genome containing our two candidate genes. The gene *tgfb3*, which was identified based on close physical genomic location to the peak LOD marker (206A) for fragmentation, demonstrated two SNPs throughout the predicted transcript (Fig. 3C). The first SNP was an A-to-T substitution 252 bp upstream from the translational start codon (SNP depth = 126, SNP% = 74.5%). This modification to the 5′UTR is intriguing as a similar mutation in the 5′UTR has also been identified in a cardiomyopathic phenotype in humans (Beffagna et al. 2005). We note, however, that this mutation is associated with higher levels of *tgfb3* expression compared to controls (Beffagna et al. 2005). We also identified a silent mutation in the coding sequence of exon 5 (D-to-D at position 289 in cavefish; SNP depth = 201, SNP% = 74.4%). We tested whether this SNP may introduce a cryptic splice site, and consequently alter the length of the resulting transcript; however, we did not observe a difference in amplicon lengths between cave and surface fish (data not shown). We therefore assume this SNP does not introduce a cryptic splice site and is inconsequential for the encoded protein. The gene *bmp4*, which has been previously implicated in craniofacial alterations in several natural systems, revealed no structural differences between cave and surface fish (Fig. 3D).

DISCUSSION

Genomic analyses identify two putative candidate genes mediating natural bone fragmentation in cavefish

One of the most promising candidate genes identified from this analysis is *tgfb3*. This gene encodes a member of the TGF-β family of secreted proteins (Yang et al. 2008), which are involved in diverse developmental processes (Oshima et al. 1996; Azhar et al. 2009). Therefore, this encoded product likely plays a critical role, shared deeply among vertebrates, in craniofacial morphogenesis (Yang and Kaartinen 2007; Aluwihare et al. 2009; Brugmann et al. 2010). *Tgfb3* null mouse studies demonstrate a cleft palate phenotype (Proetzel et al., 1995; Kaartinen et al. 1997; Yang and Kaartinen 2007; Juriloff and Harris 2008). Further, in chickens deficient for *tgfb3*, palatal shelves fail to fuse, and the application of exogenous *tgfb3* can rescue clefting by stimulating palatal fusion (Sun et al. 1998). In humans, several population association studies have revealed a putative role for *tgfb3* in non-syndromic oral clefting phenotypes (Lidral et al. 1997, 1998; Vieira et al. 2003). The fragmentation phenotype we observed affecting the third suborbital bone may indicate a "failure-to-fuse" phenotype. Therefore, a similar aberration involving the molecular signaling of *tgfb3* may cause the persistence or incomplete fusion of dermal bone condensations associated with the cavefish SO3 bone. Alternatively, the fragmentation phenotype may be the result of an imbalance in bone resorption or mineralization later in life history.

We previously demonstrated that the SO3 fragmentation phenotype is genetically asymmetric (Gross et al. 2014). It remains unknown if this asymmetry reflects variable penetrance of fragmentation across the left-right axis, a cryptic "handedness" (Chase et al. 2004) that manifests itself through aberrant ossification, or another unknown explanation. In any case, *tgfb3* has previously been implicated in gross asymmetrical differences (Antunes et al. 2012). Thus, this candidate gene is particularly attractive as it impacts analogous structural defects, in addition to asymmetrical phenotypes associated with craniofacial morphogenesis.

Our coding sequence studies indicate *tgfb3* does not harbor loss-of-function or hypomorphic coding sequence errors (Fig. 3C), but potentially reduced expression at key stages of development (Fig. 3, A and F). Although our RNA-seq studies provided a highly sensitive method to detect expression

differences between surface- and cave-dwelling fish at 3 dpf, we were unable to validate this modest variation in expression with qPCR analyses for either *tgfb3* ($P = 0.509024$) or *bmp4* ($P = 0.745742$; Fig. 3E). We did, however, validate the polarity of expression difference using both quantitative (Fig. 3E) and qualitative (Fig. 3F) approaches.

Tgfb3 is widely expressed during development (Doetschman et al. 2012), and harbors numerous pleiotropic roles in physiological processes such as cardiac development (Azhar et al. 2003) and wound healing (Galatz et al. 2006). Transgenic mice that harbor mutated forms of this gene demonstrate multiple phenotypic effects and die at birth. Therefore, it is not surprising that we do not observe destructive coding sequence alterations in this gene. In the context of craniofacial development, normal *tgfb3* signaling promotes mesenchymal confluence during late phase palatogenesis (Jalali et al. 2012). Thus, the reduced expression we may be observing in cavefish could contribute to insufficient fusion of CNC-derived mesenchymal cells, and lead to the fragmented SO3 bone phenotype observed in Pachón cavefish.

The gene *bmp4* has a well-established role in craniofacial morphology (Helms et al. 2005) and evolution (Trainor et al. 2003). In addition to the elucidated roles of this gene in traditional model systems (Golden et al. 1999; Liu et al. 2005), a wealth of research in natural systems (Streelman et al. 2003; Albertson and Kocher 2006) indicates a deeply conserved set of processes similarly mediated by this gene. For example, prior studies in Darwin's finches demonstrated increased expression of *bmp4* during critical stages of development impacts beak depth and breadth (Abzhanov et al. 2004). This increased *bmp4* expression experimentally induces similar changes when mis-expressed in the cranial mesenchyme of chicken embryos (Abzhanov et al., 2004). Similar to Darwin's finches, we did not observe coding sequence errors in Pachón cavefish (Fig. 3D). This was unsurprising given the pleiotropic roles *bmp4* plays in numerous critical developmental processes. However, our RNA-seq and qPCR analyses similarly did not indicate substantial expression level differences between morphotypes across embryogenesis (Fig. 3B), with the exception of 3 dpf. At this stage, however the expression of *bmp4* was counter-intuitively higher in cavefish compared to surface fish, and this expression pattern could not be validated using qPCR (Fig. 3E). Taken together, these results may suggest this gene product is not implicated in the craniofacial phenotype we present here. Overall, our approach illustrates the ability to integrate multiple sources of information to inform the nomination of the most promising candidate genes for diverse processes evolving in the natural world.

Astyanax as an evolutionary mutant model for craniofacial abnormalities

As a consequence of life in complete darkness, *Astyanax* cavefish have evolved a series of striking phenotypes (Gross 2012; Kowalko et al. 2013b; Rohner et al. 2013; Elipot et al. 2014). Among the most salient of these features are eye loss (Pottin et al. 2011; Hinaux et al. 2013; O'Quin et al. 2013) and albinism (McCauley et al. 2004; Protas et al. 2006; Bilandzija et al. 2013; Gross and Wilkens 2013). As a consequence, these traits have received a great deal of attention in the scientific literature. However, craniofacial phenotypes were among the first traits characterized in this model system (Alvarez 1946, 1947). Alvarez (1946) first characterized bony fragmentation, which was subsequently used as a diagnostic feature for particular cave populations by Mitchell et al. (1977). For several decades, the basis for this fragmentation phenotype remained unknown. Yamamoto et al. (2003) demonstrated, using experimental embryonic lentectomy, that several craniofacial features likely evolved as an indirect consequence of eye loss. However, certain features—including suborbital bone fragmentation—occur independently of eye loss (Yamamoto et al. 2003). This feature was discovered to be present in nearly every member of particular cave populations, further implying that it is genetically controlled (Mitchell et al. 1977). The first formal evaluation of this trait indeed demonstrated a robust association with two genomic regions (Gross et al. 2014). Furthermore, this genetic association was asymmetric—only appearing when scored on the right side in an experimental F_2 pedigree (Gross et al. 2014).

At the morphological level, this phenotype provides a powerful opportunity to explore fundamental problems in contemporary craniofacial biology. For instance, dermal bone fragmentation resembles, at the gross morphological level, fusion abnormalities affecting sutural regions (Vij and Mao 2006; Wang et al. 2006). This failure-to-fuse phenotype may serve as a model for understanding analogous issues affecting common disorders such as non-syndromic cleft palate, which arises from failed fusion of the palatal shelves during embryogenesis (Proetzel et al. 1995; Miettinen et al. 1999). At the cellular level, this phenotype offers the opportunity to explore how alterations to embryonic tissues can impact adult bone patterning. The craniofacial complex is a highly integrated organ system arising from multiple embryonic tissue sources (Noden and Schneider 2006; Brugmann et al., 2010). The facial bones of the dermal skeleton, however, arise from the cranial neural crest (Gross and Hanken 2008). The presence of this natural phenotype therefore enables us to evaluate aberrations at all stages of cranial neural crest cell development, including specification, delamination, migration, and differentiation. Moreover, this aberrant phenotype is manifested in an asymmetric fashion. Future studies will determine the genetic, molecular, and cellular underpinnings of this breakage of laterality. This remarkable model system evolved these phenotypes following millions of generations in the complete darkness of the cave (Ornelas-Garcia et al. 2008), allowing us to determine how and why analogous features may have evolved in humans. With the increasing availability of numerous genomic

and transcriptomic resources, we will have unprecedented ability to shed light on the biological basis for abnormal craniofacial processes evolving in this natural model system.

Future directions

In this report, we integrated data from QTL analyses, genomic positional information, gene ontology terms, and RNA-sequencing expression data to identify candidate genes mediating bone fragmentation. The analysis we present in this report centers on the selection of candidate genes, specifically using comparative genomics. However, it remains unclear precisely when during development (or juvenilehood) the fragmentation defect arises. Because of the nature of the phenotype, it is unclear if fragmentation is rooted in an early developmental aberration affecting specification of neural crest progenitors, a neural crest migration defect, a defect at the level of the bone differentiation (condensation, fusion failures of ossification centers) or some combination of these processes. Our RNA-seq studies were designed to capture differential expression patterns between cave and surface fish at critical stages of development during which extreme morphological changes are first evident (e.g., eye and pigmentation regression). However, this timing is prior to circumorbital bone ossification. The SO3 bone commences ossification at the roughly 22 mm body size stage of development from multiple primary foci, and completes ossification over a period of approximately 3–4 weeks, by the approximately 40 mm+ stage of development (Yamamoto et al. 2003). Future transcriptional profiling across these latter stages will identify expression level differences between morphotypes during the ossification process and possibly inform the molecular basis for the fragmentation defect. Further, if the candidate genes *tgfb3* and/or *bmp4* mediate asymmetric fragmentation, then left-right asymmetric expression of both genes may be predicted across early stages of development. Future studies will focus on validating in situ hybridization expression patterns across the protracted period between 3 dpf and juvenile/adult stages. Further, in vivo and/or in vitro functional validation studies are expected to elevate these genes from "candidate" to "causative" genes. This work will aim to clarify the roles of these genes, and the embryonic cranial neural crest, in mediating aberrant bone differentiation (e.g., fragmentation) using a natural model system for understanding the cellular, developmental, and genetic bases for asymmetric craniofacial defects.

CONCLUSIONS

As a consequence of life in total darkness, cavefish have evolved several craniofacial abnormalities. Unlike many cave-associated phenotypes, there seems to be a loss of left-right sided constraint on certain aberrations such as dermal bone fragmentation and fusions. In contrast, other craniofacial traits such as bone area size remain constrained—exhibiting identical genetic bases on both left and right sides of the head. Here, we leverage emerging genomic resources in *Astyanax* to nominate candidate genes mediating bony fragmentation. We identified two key genes that may participate in incomplete ossification (*tgfb3* and *bmp4*) by integrating genomic positional information, GO term annotation, RNA-seq expression, and transcript structural studies. These genes are particularly attractive as they have been implicated in similar failure-to-fuse phenotypes, such as cleft palate, in other systems. Both genes have also been implicated in asymmetric craniofacial abnormalities in other systems. We propose that the blind Mexican cavefish is a powerful emerging genomic model for understanding complex trait evolution. Future studies, aimed at functional validation of these candidate genes, will help inform how and why craniofacial abnormalities evolve in the context of the natural world. By understanding the fundamental mechanisms through which these traits arise, we will inform how and why analogous clinical phenotypes are present in humans.

Acknowledgements

We are grateful to Dr. Thomas Diekwisch and the organizers of the 2013 Oral Biology Centennial Conference at the University of Illinois Chicago Dental School. The authors wish to thank two anonymous reviewers and Dr. Michael Coates for helpful comments on an earlier draft of this manuscript. We are also grateful to Drs. Wesley Warren and Suzanne McGaugh for generously providing access to the *Astyanax* draft genome assembly (funded by NIH grant no. R24 RR032658-01 to Dr. Wesley Warren, The Genome Institute at Washington University School of Medicine; Bioproject PRJNA89115 NCBI accession no. APWO00000000). This project was supported by NIH/NIDCR Award DE022403 to J. B. G.

References

Abzhanov, A., Protas, M., Grant, B. R., Grant, P. R., and Tabin, C. J. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305: 1462–1465.

Albertson, R. C. and Kocher, T. D. 2006. Genetic and developmental basis of cichlid trophic diversity. *Heredity* 97: 211–221.

Albertson, R. C., Cresko, W., Detrich, H. W., III, and Postlethwait, J. H. 2009. Evolutionary mutant models for human disease. *Trends Genet.* 25: 74–81.

Aluwihare, P., et al. 2009. Mice that lack activity of avß6- and avß8-integrins reproduce the abnormalities of Tgfb1- and Tgfb3-null mice. *J. Cell Sci.* 122: 227–232.

Alvarez, J. 1947. Descripcion de *Anoptichthys hubbsi* caracindo ciego de la Cueva de los Sabinos, S. L. P. *Rev. Soc. Mexicana Hist. Nat.* 8: 215–219.

Alvarez, J. 1946. Revisión del género *Anoptichthys* con descipción de una especie nueva (Pisces, Characidae). *An. Esc. Nac. Cien. Biol. Méx.* 4: 263–282.

Antunes, L. S., et al. 2012. *TGFB3* and *BMP4* polymorphism are associated with isolated tooth agenesis. *Acta. Odontol. Scand.* 70: 202–206.

Azhar, M., et al. 2003. Transforming growth factor ß in cardiovascular development and function. *Cytokine Growth Factor Rev.* 14: 391–407.

Azhar, M., et al. 2009. Ligand-specific function of transforming growth factor beta in epithelial-mesenchymal transition in heart development. *Dev. Dyn.* 238: 431–442.

Beffagna, G., et al. 2005. Regulatory mutations in transforming growth factor-ß3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. *Cardiovasc Res.* 65: 366–373.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.

Bilandzija, H., Ma, L., Parkhurst, A., and Jeffery, W. R. 2013. A potential benefit of albinism in *Astyanax* cavefish: downregulation of the *oca2* gene increases tyrosine and catecholamine levels as an alternative to melanin synthesis. *PLoS ONE* 8: e80823.

Borowsky, R. 2008. Restoring sight in blind cavefish. *Curr. Biol.* 18: R23–R24.

Bourgeois, P., et al. 1998. The variable expressivity and incomplete penetrance of the twist-null heterozygous mouse phenotype resemble those of human Saethre–Chotzen syndrome. *Hum. Mol. Genet.* 7: 945–957.

Bradic, M., Beerli, P., Garcia-de Leon, F. J., Esquivel-Bobadilla, S., and Borowsky, R. L., 2012. Gene flow and population structure in the mexican blind cavefish complex (*Astyanax mexicanus*). *BMC Evol. Biol.* 12: 9.

Broman, K. W., Wu, H., Sen, Ś., and Churchill, G. A. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890.

Brugmann, S. A., et al. 2010. Comparative gene expression analysis of avian embryonic facial structures reveals new candidates for human craniofacial disorders. *Hum. Mol. Genet.* 19: 920–930.

Buffenstein, R. 2008. Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. *J. Comp. Physiol. B* 178: 439–445.

Chase, K., Lawler, D. F., Adler, F. R., Ostrander, E. A., and Lark, K. G. 2004. Bilaterally asymmetric effects of quantitative trait loci (QTLs): QTLs that affect laxity in the right versus left coxofemoral (hip) joints of the dog (*Canis familiaris*). *Am. J. Med. Genet. A* 124: 239–247.

Doetschman, T., et al. 2012. Generation of mice with a conditional allele for the *transforming growth factor beta 3* gene. *Genesis* 50: 59–66.

Elipot, Y., Hinaux, H., Callebert, J., Launay, J. M., Blin, M., and Retaux, S. 2014. A mutation in the enzyme monoamine oxidase explains part of the *Astyanax* cavefish behavioural syndrome. *Nat. Commun.* 5: 3647.

Galatz, L. M., et al. 2006. Characteristics of the rat supraspinatus tendon during tendon-to-bone healing after acute injury. *J. Orthop. Res.* 24: 541–550.

Golden, J. A., Bracilovic, A., McFadden, K. A., Beesley, J. S., Rubenstein, J. L. R., and Grinspan, J. B. 1999. Ectopic bone morphogenetic proteins 5 and 4 in the chicken forebrain lead to cyclopia and holoprosencephaly. *Proc. Natl. Acad. Sci. USA* 96: 2439–2444.

Gross, J. B., Furterer, A., Carlson, B. C., and Stahl, B. A. 2013. An integrated transcriptome-wide analysis of cave and surface dwelling *Astyanax mexicanus*. *PloS ONE* 8: e55659.

Gross, J. B. and Wilkens, H. 2013. Albinism in phylogenetically and geographically distinct populations of *Astyanax* cavefish arises through the same loss-of-function *Oca2* allele. *Heredity* 111: 122–130.

Gross, J. B. 2012. The complex origin of *Astyanax* cavefish. *BMC Evol. Biol.* 12: 105.

Gross, J. B. and Hanken, J. 2008. Review of fate-mapping studies of osteogenic cranial neural crest in vertebrates. *Dev. Biol.* 317: 389–400.

Gross, J. B., Krutzler, A. J., and Carlson, B. M. 2014. Complex craniofacial changes in blind cave-dwelling fish are mediated by genetically symmetric and asymmetric loci. *Genetics* 196: 1303–1319.

Gross, J. B., et al. 2008. Synteny and candidate gene prediction using an anchored linkage map of *Astyanax mexicanus*. *Proc. Natl. Acad. Sci. USA* 105: 20106–20111.

Haley, C. S. and Knott, S. A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69: 315–324.

Helms, J. A., Cordero, D., and Tapadia, M. D. 2005. New insights into craniofacial morphogenesis. *Development* 132: 851–861.

Hinaux, H., et al. 2011. A developmental staging table for *Astyanax mexicanus* surface fish and Pachón cavefish. *Zebrafish* 8: 155–165.

Hinaux, H., et al. 2013. De novo sequencing of *Astyanax mexicanus* surface fish and Pachón cavefish transcriptomes reveals enrichment of mutations in cavefish putative eye genes. *PLoS ONE* 8: e53553.

Hulsey, C. D., Mims, M. C., and Streelman, J. T. 2007. Do constructional constraints influence cichlid craniofacial diversification? *Proc. R. Soc.* 274: 1867–1875.

Jalali, A., Zhu, X., Liu, C., and Nawshad, A. 2012. Induction of palate epithelial mesenchymal transition by transforming growth factor β3 signaling. *Dev. Growth Differ.* 54: 633–648.

Jeffery, W. R. 2009. Regressive evolution in *Astyanax* cavefish. *Annu. Rev. Genet.* 43: 25–47.

Juriloff, D. M. and Harris, M. J. 2008. Mouse genetic models of cleft lip with or without cleft palate. *Birth Defect Res. A Clin. Mol. Teratol.* 82: 63–77.

Kaartinen, V., Cui, X., Heisterkamp, N., Groffen, J., and Shuler, C. F. 1997. Transforming growth factor-b3 regulates transdifferentiation of medial edge epithelium during palatal fusion and associated degradation of the basement membrane. *Dev. Dyn.* 209: 255–260.

Kasprzyk, A. 2011. BioMart: driving a paradigm change in biological data management. *Database (Oxford)*, bar049.

Kearsey, M. and Hyne, V. 1994. QTL analysis: a simple 'marker-regression' approach. *Theor. Appl. Genet.* 89: 698–702.

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* 5: 288–298.

Korff, S., Stein, D. J., and Harvey, B. H. 2008. Stereotypic behaviour in the deer mouse: pharmacological validation and relevance for obsessive compulsive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 32: 348–355.

Kowalko, J. E., et al. 2013a. Convergence in feeding posture occurs through different genetic loci in independently evolved cave populations of *Astyanax mexicanus*. *Proc. Natl. Acad. Sci. USA* 110: 16933–16938.

Kowalko, J., et al. 2013b. Loss of schooling behavior in cavefish through sight-dependent and sight-independent mechanisms. *Curr. Biol.* 23: 1874–1883.

Lidral, A. C., et al. 1997. Studies of the candidate genes *TGFB2*, *MSX1*, *TGFA*, and *TGFB3* in the etiology of cleft lip and palate in the Philippines. *Cleft Palate Craniofac. J.* 34: 1–6.

Lidral, A. C., et al. 1998. Association of *MSX1* and *TGFB3* with nonsyndromic clefting in humans. *Am. J. Hum. Genet.* 63: 557–568.

Lieberman, D. E. 2008. Speculations about the selective basis for modern human craniofacial form. *Evol. Anthropol.* 17: 55–68.

Liu, W., et al. 2005. Distinct functions for bmp signaling in lip and palate fusion in mice. *Development* 132: 1453–1461.

Lundberg, J. G.(1993) African-South American freshwater clades and continental drift: Problems with a paradigm. *Biological Relationships Between Africa and South America*. In P. Goldblatted (ed.). New Haven: Yale University Press, pp. 156–199.

Mallarino, R., Grant, P. R., Grant, B. R., Herrel, A., Kuo, W. P., and Abzhanov, A. 2011. Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc. Natl. Acad. Sci. USA* 108: 4057–4062.

Martin, A. and Orgogozo, V. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67: 1235–1250.

McCauley, D. W., Hixon, E., and Jeffery, W. R. 2004. Evolution of pigment cell regression in the cavefish *Astyanax*: a late step in melanogenesis. *Evol. Dev.* 6: 209–218.

McGaugh, S. E., et al. 2014. The cavefish genome reveals candidate genes for eye loss. *Nat. Comm.* 5: 5307.

Miettinen, P. J., et al. 1999. Epidermal growth factor receptor function is necessary for normal craniofacial development and palate closure. *Nat. Genet.* 22: 69–73.

Miller, C. T., et al. 2014. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics*, In press.

Mitchell, R. W., Russell, W. H., and Elliott, W. R. 1977. *Mexican Eyeless Characin Fishes, Genus Astyanax: Environment, Distribution, and Evolution*. Lubbock, Texas: Texas Tech Press, p. 89.

Mooney, M. P., et al. 1994. Development of a strain of rabbits with congenital simple nonsyndromic coronal suture synostosis part I: breeding demographics, inheritance pattern, and craniofacial anomalies. *Cleft Palate Craniofac. J.* 31: 1–7.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* 5: 621–628.

Muenke, M. 1994. Holoprosencephaly as a genetic model for normal craniofacial development. *Semin. Dev. Biol.* 5: 293–301.

Noden, D. M. and Schneider, R. A. 2006. Neural crest cells and the community of plan for craniofacial development: historical debates and current perspectives. *Adv. Exp. Med. Biol.* 589: 1–23.

Nogai, H., et al. 2008. Follistatin antagonizes transforming growth factor-β3-induced epithelial-mesenchymal transition in vitro: implications for murine palatal development supported by microarray analysis. *Differentiation* 76: 404–416.

O'Quin, K. E., Yoshizawa, M., Doshi, P., and Jeffery, W. R. 2013a. Quantitative genetic analysis of retinal degeneration in the blind cavefish *Astyanax mexicanus*. *PLoS ONE* 8: e57281.

Ornelas-Garcia, C. P., Dominguez-Dominguez, O., and Doadrio, I. 2008. Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in mesoamerica reveals multiple morphological homoplasies. *BMC Evol Biol* 8: 340.

Oshima, M., Oshima, H., and Taketo, M, M. 1996. TGF-β receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. *Dev Biol* 179: 297–302.

Paton, J. F., Dickinson, C. J., and Mitchell, G. 2009. Harvey cushing and the regulation of blood pressure in giraffe, rat and man: introducing 'cushing's mechanism'. *Exp. Physiol.* 94: 11–17.

Pottin, K., Hinaux, H., and Retaux, S. 2011. Restoring eye size in *Astyanax mexicanus* blind cavefish embryos through modulation of the *shh* and *Fgf8* forebrain organising centres. *Development* 138: 2467–2476.

Proetzel, G., et al. 1995. Transforming growth factor-beta 3 is required for secondary palate fusion. *Nat. Genet.* 11: 409–414.

Protas, M. E.(2005) *The genetic basis of morphological evolution in the mexican cave tetra, Astyanax mexicanus*. Harvard University: Dissertation.

Protas, M., et al. 2008. Multi-trait evolution in a cave fish, *Astyanax mexicanus*. *Evol. Dev.* 10: 196–209.

Protas, M. E., et al. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat. Genet.* 38: 107–111.

Rohner, N., et al. 2013. Cryptic variation in morphological evolution: HSP90 as a capacitor for loss of eyes in cavefish. *Science* 342: 1372–1375.

Stern, D. L. and Orgogozo, V. 2009. Is genetic evolution predictable? *Science* 323: 746–751.

Streelman, J. T., Webb, J. F., Albertson, R. C., and Kocher, T. D. 2003. The cusp of evolution and development: a model of cichlid tooth shape diversity. *Evol. Dev.* 5: 600–608.

Sun, D., Vanderburg, C. R., Odierna, G. S., and Hay, E. D. 1998. *TGFbeta3* promotes transformation of chicken palate medial edge epithelium to mesenchyme in vitro. *Development* 125: 95–105.

Sutter, N. B., et al. 2007. A single *IGF1* allele is a major determinant of small size in dogs. *Science* 316: 112–115.

Trainor, P. A., Melton, K. R., and Manzanares, M. 2003. Origins and plasticity of neural crest cells and their roles in jaw and craniofacial evolution. *Int. J. Dev. Biol.* 47: 541–553.

Venn-Watson, S. K. and Ridgway, S. H. 2007. Big brains and blood glucose: common ground for diabetes mellitus in humans and healthy dolphins. *Comp. Med.* 57: 390–395.

Vieira, A. R., Orioli, I. M., Castilla, E. E., Cooper, M. E., Marazita, M. L., and Murray, J. C. 2003. MSX1 and TGFB3 contribute to clefting in south america. *J. Dent. Res.* 82: 289–292.

Viguier, B. 2004. Functional adaptations in the craniofacial morphology of Malagasy primates: shape variations associated with gummivory in the family Cheirogaleidae. *Ann. Anat.* 186: 495–501.

Vij, K. and Mao, J. J. 2006. Geometry and cell density of rat craniofacial sutures during early postnatal development and upon *in vivo* cyclic loading. *Bone* 38: 722–730.

Wang, Q., Strait, D. S., and Dechow, P. C. 2006. Fusion patterns of craniofacial sutures in rhesus monkey skulls of known age and sex from Cayo Santiago. *Am. J. Phys. Anthropol.* 131: 469–485.

Wilkie, A. O. and Morriss-Kay, G. M. 2001. Genetics of craniofacial development and malformation. *Nat. Rev. Genet.* 2: 458–468.

Woese, C. R., Kandler, O., and Wheelis, M. L. 1990. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. *Proc. Natl. Acad. Sci. USA* 87: 4576–4579.

Xu, S. 2010. An expectation-maximization algorithm for the lasso estimation of quantitative trait locus effects. *Heredity* 105: 483–494.

Yamamoto, Y., Espinasa, L., Stock, D. W., and Jeffery, W. R. 2003. Development and evolution of craniofacial patterning is mediated by eye-dependent and -independent processes in the cavefish *Astyanax*. *Evol. Dev.* 5: 435–446.

Yang, L. and Kaartinen, V. 2007. *Tgfb1* expressed in the *Tgfb3* locus partially rescues the cleft palate phenotype of *Tgfb3* null mutants. *Dev. Biol.* 312: 384–395.

Yang, L., Li, W., and Kaartinen, V. 2008. Tissue-specific expression of *cre* recombinase from the *Tgfb3* locus. *Genesis* 46: 112–118.

Yoshizawa, M., Jeffery, W., Van Netten, S., and McHenry, M. 2013. The sensitivity of lateral line receptors and their role in the behavior of Mexican blind cavefish (*Astyanax mexicanus*). *J Exp Biol* 217: 886–895.