METHODS & TECHNIQUES

Computational anatomy and geometric shape analysis enables analysis of complex craniofacial phenotypes in zebrafish

Kelly M. Diamond[1,*], Sara M. Rolfe[1,2], Ronald Y. Kwon[3,4] and A. Murat Maga[1,5]

ABSTRACT

Due to the complexity of fish skulls, previous attempts to classify craniofacial phenotypes have relied on qualitative features or sparce 2D landmarks. In this work we aim to identify previously unknown 3D craniofacial phenotypes with a semiautomated pipeline in adult zebrafish mutants. We first estimate a synthetic 'normative' zebrafish template using MicroCT scans from a sample pool of wild-type animals using the Advanced Normalization Tools (ANTs). We apply a computational anatomy (CA) approach to quantify the phenotype of zebrafish with disruptions in *bmp1a*, a gene implicated in later skeletal development and whose human ortholog when disrupted is associated with Osteogenesis Imperfecta. Compared to controls, the *bmp1a* fish have larger otoliths, larger normalized centroid sizes, and exhibit shape differences concentrated around the operculum, anterior frontal, and posterior parietal bones. Moreover, *bmp1a* fish differ in the degree of asymmetry. Our CA approach offers a potential pipeline for high-throughput screening of complex fish craniofacial shape to discover novel phenotypes for which traditional landmarks are too sparce to detect. The current pipeline successfully identifies areas of variation in zebrafish mutants, which are an important model system for testing genome to phenome relationships in the study of development, evolution, and human diseases.

This article has an associated First Person interview with the first author of the paper.

INTRODUCTION

The fish craniofacial skeleton is a useful system for elucidating genetic and environmental contribution to phenotype in vertebrates. In the context of development, studies have focused on the genetic mechanisms that shape the cranial skeleton (Kimmel et al., 2020; Miller et al., 2007). Craniofacial analyses have been used to understand the pathways that have enabled morphological evolution

[1]Center for Developmental Biology and Regenerative Medicine, Seattle Children's Research Institute, Seattle, WA 98101, USA. [2]Friday Harbor Marine Laboratories, University of Washington, San Juan, WA 98250, USA. [3]Department of Orthopedics and Sports Medicine, University of Washington, Seattle, WA 98195, USA. [4]Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA 98109, USA. [5]Division of Craniofacial Medicine, Department of Pediatrics, University of Washington, Seattle, WA 98105, USA.

*Author for correspondence (KellyMGDiamond@gmail.com)

K.M.D., 0000-0001-8639-6795; S.M.R., 0000-0001-9514-7774; A.M.M., 0000-0002-7921-9018

(Kimmel et al., 2005), phenotypic plasticity (Navon et al., 2020), and adaptive radiations (Powder and Albertson, 2016) in fishes. Additionally, zebrafish are developing as a model system for quantifying phenotypic variability associated with human bone diseases, such as Osteogenesis Imperfecta (Busse et al., 2019; Gistelinck et al., 2018; Kwon et al., 2019). A longstanding challenge to analyzing the fish craniofacial skeleton is accurately capturing phenotypes that involve subtle alterations and complex 3D changes, including potential asymmetric alterations.

The traditional methods for quantifying cranial morphology use manually-placed homologous landmark points on 2D images of the lateral view of the head (i.e. Sidlauskas, 2008). However, manual placement limits potential for rapid-throughput applications. Further, the requirement for homologous structures limits landmark placement across the skull, and hence may miss the phenotypic variation in these areas. While MicroCT can help realize 3D structures, 3D landmark placement is complex as visualizations are dependent on both the scanner and rendering software settings used. Moreover, because of the close proximity of bones, segmentation-based approaches that are useful for axial skeleton are not amenable to those in the head. There is an urgent need to develop robust methods for phenotyping in the craniofacial skeleton that are sensitive to complex 3D changes while being amenable to rapid-throughput analyses.

Here, we propose using an atlas-based computational anatomy (CA) approach to build a reference template of the zebrafish skull and then using a pseudo-landmark pipeline to identify areas of the skull that vary among mutant and wild-type fish. Atlas-based approaches estimate an unbiased anatomical 'template' from a group of images (Guimond et al., 2006) and then use this template as the basis to assess shape differences among groups of interest (Ashburner and Friston, 2000). Atlas-based approaches have been used to characterize phenotypes in many neuroimaging studies in humans and fetal mice (Mandal et al., 2012; Mazziotta et al., 2001; KOMP2 project), as well as in the cranial skeleton of humans and mice (Darvann et al., 2011; Maga et al., 2017; Toussaint et al., 2021). We define pseudo-landmarks here as landmarks that are not morphologically homologous, but instead, are placed automatically on the surface of the template and transferred to each specimen, enforcing a degree of geometric homology.

We apply these methods to zebrafish with mutations in *bmp1a*, a gene implicated in later skeletal development. In humans, Bone Morphogenetic Protein 1 (BMP1) encodes for a secreted protein involved in procollagen processing. Individuals with mutations in BMP1 exhibit increased bone mineral density and recurrent fractures characteristic of Osteogenesis Imperfecta (OI; Asharani et al., 2012). Severe forms of OI are frequently associated with craniofacial abnormalities (Dagdeviren et al., 2019). Previous work in *bmp1a* and other zebrafish OI models have identified phenotypic abnormalities in the axial skeleton (Hur et al., 2017). However, due to the complicated structure of the fish cranial skeleton, craniofacial

Table 1. Welch two-sample *t*-test for difference between mutants and wild-type fish for each pair of manually segmented otolith volumes

Otolith	x mutant (mm³)	x wild type (mm³)	d.f.	t	P	UCL	LCL
Left asteriscus	0.044	0.039	15.975	3.383	0.004	0.002	0.008
Right asteriscus	0.044	0.040	14.62	3.232	0.006	0.002	0.007
Left lapilus	0.026	0.025	18.013	2.000	0.061	−0.0007	0.003
Right lapilus	0.026	0.025	17.693	1.554	0.138	−0.0004	0.003
Left sagitta	0.005	0.005	18.665	2.244	0.037	0.0003	0.009
Right sagitta	0.005	0.005	15.375	1.902	0.076	−0.0005	0.008

We provide the mean volumes (x) for mutants and wild-type groups, degrees of freedom (d.f.), test statistic (t), P-value (p), and confidence interval (UCL-LCL).

abnormalities in zebrafish OI models have mostly focused on qualitative phenotypes (Gistelinck et al., 2018), and little work has been done to quantify complex cranial phenotype. Here, we report complex craniofacial phenotype arising from disruptions in *bmp1a*. Our methods aim to identify areas of greatest variation among mutant and wild-type fish cranial phenotypes with minimal user intervention. We envision this method as a high-throughput first pass to identify areas for further exploration for phenotype-genotype associations in the skeletal system.

RESULTS AND DISCUSSION
When analyzing otoliths, we did not find significant differences between manually segmented volumes and atlas segmented volumes (t=-0.912, P=0.363; Fig. S1), though there were differences between some of the individual otoliths (Table S1, Fig. S1). The most apparent difference between *bmp1a* and wild-type fish is that *bmp1a* fish have larger otoliths than wild-type fish, especially for the asteriscus, the largest otoliths in the zebrafish. This difference was consistent in both manually and CA segmented otoliths (Table 1; Fig. S1). In contrast to

bone formation, in which the mineral phase is primarily hydroxyapatite, otoliths are formed via an accumulation of calcium carbonate in the acellular endolymph of the fish inner ear (Payan et al., 2004). Previous work found higher tissue mineral density in *bmp1a* fish across the axial skeleton (Hur et al., 2017) and this result suggests potential influence of *bmp1a* on other pathways associated with mineralized tissues.

We found a similar pattern in normalized centroid size, with the pseudo-landmark (t=2.700, P=0.013) and the trends for the gold standard (t=1.956, P=0.063) methods finding *bmp1a* fish to be larger than wild-type fish (Fig. 1B). Previous work on the axial skeleton found that *bmp1a* fish had reduced standard length, but thicker bones (Watson et al., 2020). It is possible that the increase in bone thickness is contributing to the larger normalized centroid size observed in these analyses. Additionally, while these two methods to differ in normalized centroid size (F=18.66, P<0.001; Fig. 1B), we find similar patterns and across methods, which helps to validate our pseudo-landmark pipeline. The distances between the gold standard and the ALPACA transferred points of the gold standard,
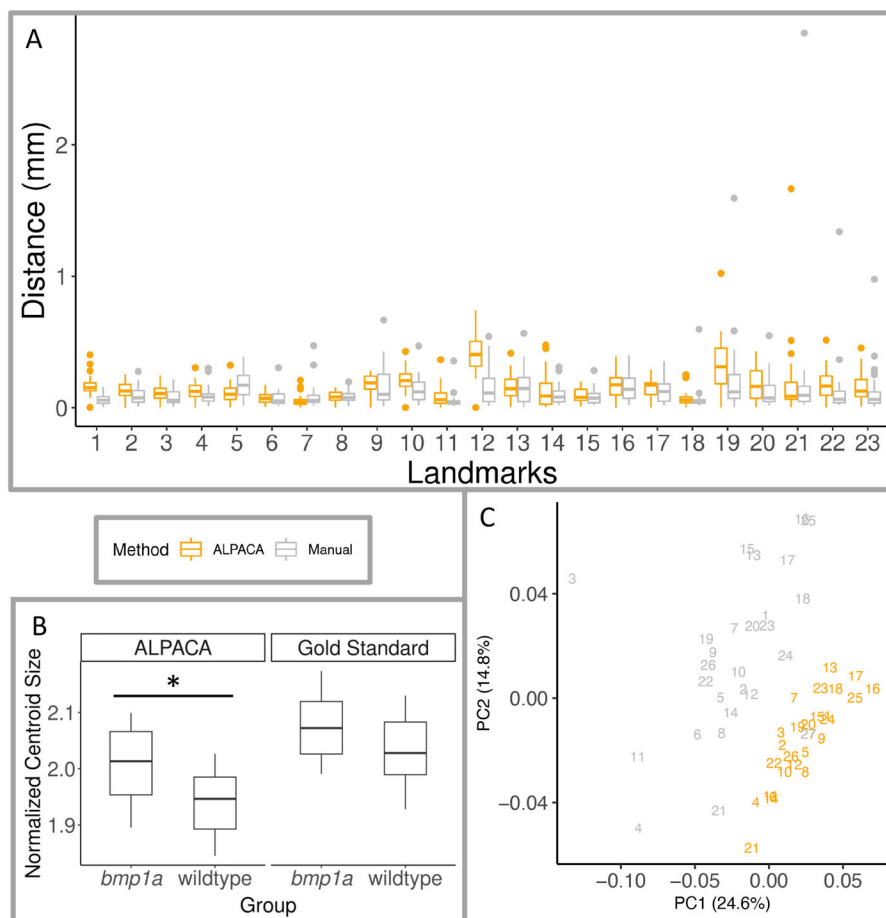
Fig. 1. Pipeline validation using 23 traditional landmarks. Boxplots are shown for (A) Euclidean distance between Gold Standard landmark locations and ALPACA transferred points (orange), and distance between the two manual landmark placements (grey), where color indicates method used for all panels, and (B) normalized centroid size of gold standard and ALPACA transferred pseudo-landmarks. Midline of boxplots show median value, with hinges corresponding to first and third quartiles, and whiskers extending to largest and smallest value no further than 1.5 times the interquartile range. Also shown (C) are the first two principal components of shape space from the combined GPA analysis of gold standard and ALPACA transferred 23 landmarks. The percent of variance for each PC is show in parenthesis of each axis. Individual fish are depicted as different numbers, fish 1–12 are crispant fish, fish 13–26 are wild-type fish, and fish 27 is the atlas.

Table 2. Digitization error associate with different methods and groups

| Landmark	Gold standard			ALPACA		
	bmp1a (N=12)	Wild type (N=14)	Pooled (N=27)	*bmp1a* (N=12)	Wild type (N=14)	Pooled (N=27)
1	0.072±0.053	0.061±0.038	0.067±0.044	**0.168±0.097**	**0.183±0.050**	**0.170±0.080**
2	0.108±0.073	0.078±0.060	0.091±0.066	0.129±0.055	0.139±0.066	0.130±0.065
3	0.094±0.055	0.071±0.059	0.079±0.058	0.128±0.045	0.106±0.064	0.112±0.059
4	0.136±0.084	0.063±0.032	0.094±0.070	0.129±0.057	**0.141±0.065**	**0.130±0.065**
5	0.186±0.113	0.183±0.105	0.182±0.105	0.123±0.065	0.113±0.071	**0.113±0.069**
6	0.073±0.041	0.078±0.079	0.075±0.062	0.077±0.040	0.080±0.048	0.076±0.045
7	0.136±0.132	0.050±0.018	0.087±0.098	0.078±0.061	0.041±0.017	0.056±0.047
8	0.076±0.030	0.091±0.055	0.085±0.045	0.092±0.043	0.085±0.031	0.085±0.039
9	0.269±0.202	0.094±0.051	0.171±0.163	0.190±0.087	**0.185±0.047**	0.180±0.075
10	0.137±0.125	0.142±0.088	0.137±0.103	**0.247±0.093**	0.204±0.070	**0.215±0.091**
11	0.050±0.027	0.062±0.086	0.055±0.064	0.100±0.060	0.077±0.090	0.084±0.077
12	0.151±0.080	0.187±0.185	0.165±0.144	**0.456±0.138**	**0.431±0.156**	**0.426±0.166**
13	0.133±0.160	0.186±0.113	0.161±0.134	0.179±0.093	0.149±0.082	0.157±0.090
14	0.130±0.091	0.071±0.045	0.098±0.073	0.225±0.142	0.067±0.051	0.135±0.129
15	0.070±0.030	0.106±0.079	0.088±0.063	**0.123±0.047**	0.078±0.054	0.095±0.056
16	0.210±0.132	0.126±0.072	0.161±0.109	0.187±0.093	0.172±0.088	0.172±0.094
17	0.136±0.097	0.132±0.092	0.130±0.093	0.184±0.070	0.135±0.050	0.152±0.069
18	0.091±0.161	0.056±0.021	0.071±0.107	0.069±0.062	0.077±0.050	0.071±0.056
19	0.281±0.446	0.186±0.128	0.226±0.308	0.299±0.275	**0.363±0.129**	**0.321±0.213**
20	0.128±0.144	0.114±0.094	0.119±0.115	0.216±0.141	0.197±0.128	0.199±0.135
21	0.348±0.793	0.096±0.057	0.209±0.533	0.345±0.441	0.083±0.047	0.196±0.319
22	0.196±0.364	0.081±0.090	0.131±0.252	0.171±0.118	**0.183±0.119**	**0.171±0.119**
23	0.164±0.267	0.106±0.120	0.132±0.195	0.171±0.124	0.149±0.090	0.153±0.108

Mean±s.d. given. For each method we compared *bmp1a* fish, wild-type fish, and pooled groups which included the atlas. Units are in mm. A paired Mann–Whitney *U*-test was used to test for differences in digitization errors with respect to gold standard at P=0.01, indicated in bold.

as well as the distance between the two independent manual landmark positions. We found that while the manual landmark placement showed shorter distances than the ALPACA method (F=14.690, P<0.001; Table 2, Fig. 1A), both methods show similar patterns in variation across the landmark points. For further validation of our pipeline, we tested if landmarking method varied across the combined shape space. While we do find differences among manual and ALPACA methods (F=9.860, Z=4.883, p=0.001), both methods show similar patterns of variation within the different treatment groups (Fig. 1C).

To assess if there were areas of the skull that had local shape variability more subtle than what could be determined from our sparse landmark analysis or manual LM-based Procrustes analysis, we deployed a pseudo-landmark approach, placing 372 geometrically placed pseudo-landmarks across the outer surface of the cranial skeleton. In our symmetry analysis of pseudo-landmark points, we found significant differences in symmetry between groups for both the symmetric (F=3.573, Z=2.708, P=0.011) and asymmetric (F=3.830, Z=3.124, P=0.002) components of shape variation. The symmetric differences in shape variation between groups were concentrated in the anterior frontal bone and the dorsal portion of the operculum (Fig. 2). While the asymmetric differences between groups were concentrated in the posterior portion of the parietal bone and ventral portion of the operculum (Fig. 2). Tables of PC scores for both analyses can be found in the supplemental information (Tables S2, S3).

The results of separate PCA of each shape component suggest the asymmetric component of shape may be contributing more to the variation between groups in our dataset. For the symmetric component of variation, we found significant differences between
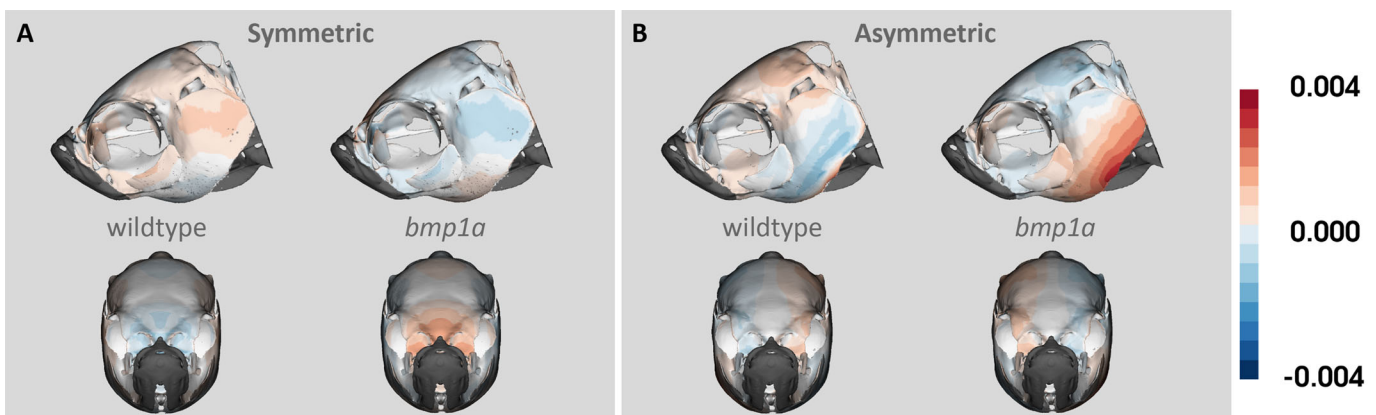

Fig. 2. Heat map of (A) symmetric and (B) asymmetric components of shape variation. Lateral and anterior views are shown for each group (wild type and *bmp1a*) within both components of shape variation. Colors show variation in shape from the symmetric atlas, with deeper colors representing greater variation from the atlas.

bmp1a and wild-type fish along PC2, which explained 12.3% of the variation in the data (*F*=7.018; *Z*=2.006; *P*=0.002), but not along PC1, which explained 42.9% of the variation in the data (*F*=2.583; *Z*=1.124; *P*=0.092; Fig. 3) or any other PCs. Whereas in the asymmetric shape space, we found differences between groups along PC1, which explained 35.0% of the variation, (*F*=6.305, *Z*=1.753, *P*= 0.009), but not along PC2, which explained 17.1% of the variation (*F*= 0.318, *Z*=-0.374, *P*=0.677; Fig. 3), or any other PCs. We find that the greatest variation across the skull is observed in the posterior operculum, as observed along the first PC in both the symmetrical and asymmetrical analyses (Fig. 3). In the symmetric shape analysis, we also observe variation in the anterior portion of the frontal bone across PC1 (Fig. 3). The asymmetric component of PC1 shows variation in the lateral parietal and supraocular regions (Fig. 3). Relative to the asymmetric component, smaller changes (lighter coloration in Fig. 3) are observed in the symmetric component of PC2 in the frontal, parietal, and ventral opercular bones. The asymmetric component of the PC2 axis is again concentrated around the opercular and ocular regions (Fig. 3). As we removed pseudo-landmark points associated with areas of the skull that varied due to preservation or scanning methods, this variation

represents areas of interest for exploring how phenotype differs between mutant and wild-type fishes.

As fish skulls are particularly kinetic, it is possible that the differences observed, especially those observed with the opercula, could be the result of preservation and or imaging techniques. We also note that because we used a symmetric template to generate the pseudo-landmark points, fine scale aspects of anatomical variation, such as differences in lateral line pits, may not be captured by our current pipeline. However, we note that this pseudo-landmark method is intended to identify areas of phenotypic variability across complex morphology for which traditional landmarking methods may be too sparse to capture, and not to directly correlate genotype to phenotype. Together these results provide support for phenotypic effects of the *bmp1a* mutation on the cranial phenotype of zebrafish. Future work should expand the number of families to ensure this is not unique to this particular family and include other potential factors, such as differences in sex. We have shown how our pipeline can identify areas of greatest variation among groups of animals. In combination with additional morphological analyses, we hope this pipeline will enable researchers to better define the links between genotype and phenotype.
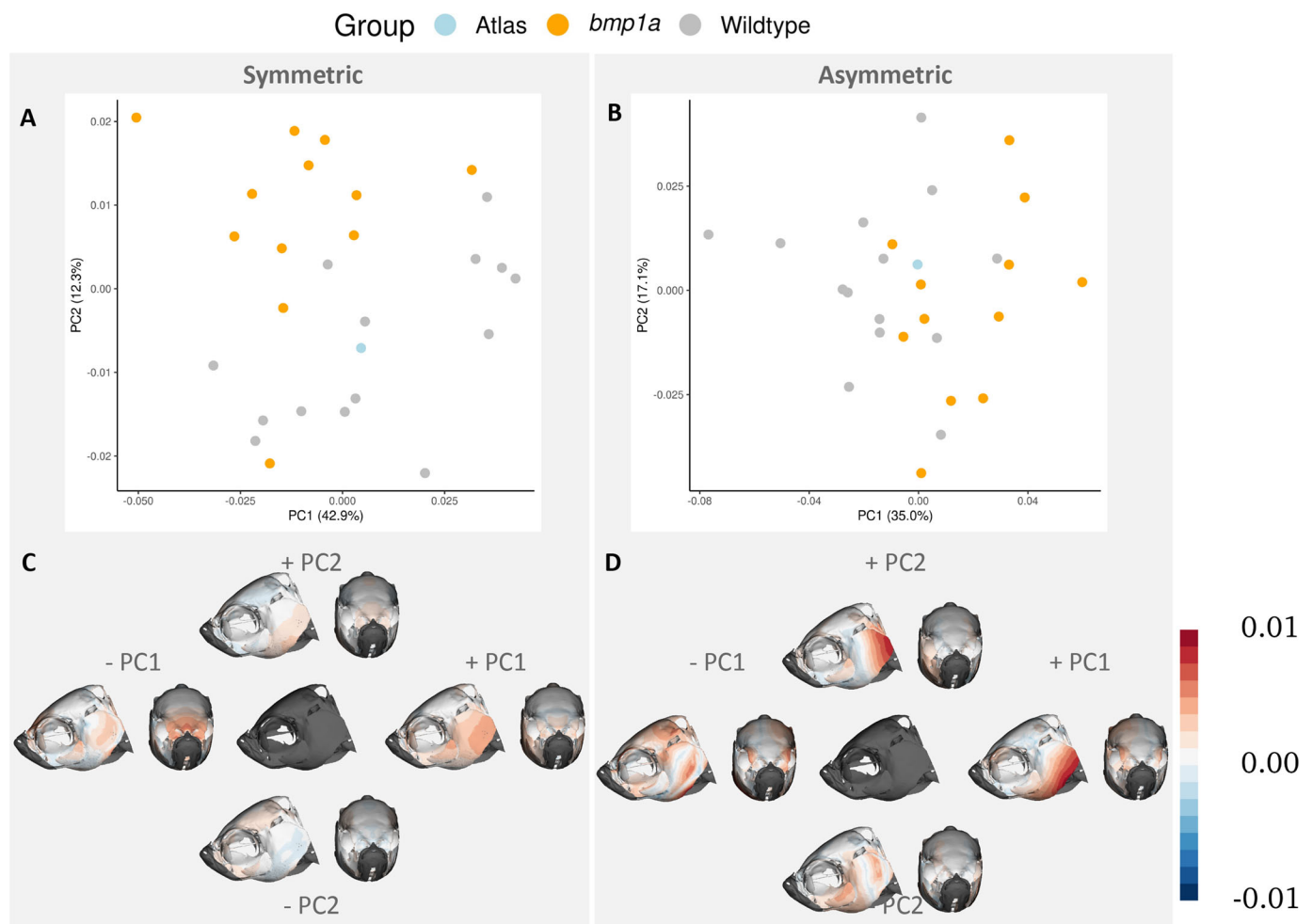
Fig. 3. First two principal components of symmetry analysis. PC plots show separation of groups (represented by color) along the first and second PCs (A,B). Heat maps of the same PCs represent where shape variation occurs across each axis (C,D). Columns represent symmetric (A,C) and asymmetric (B,D) components of shape variation. The central image in C and D represents mean shape of each component. Color in C and D represents the Procrustes distance between the average shape and the shape occupying the ends of each PC axis. Deeper colors represent larger differences, and the specific colors refer to differences in direction relative to the average image.

Table 3. Manual landmark definitions are provided for each landmark number (#)

#	Definition	#	Definition
1	Anterior frontal	13	Left posterior opercula
2	Dorsal epiphyseal bar	14	Left meeting of dentary/quadrate
3	Dorsal meeting of frontal/parietal	15	Right epiphyseal bar
4	Left meeting of parietal/supraoccipital	16	Right meeting of frontal/parietal
5	Dorsal meeting of parietal/supraoccipital	17	Right posterior pterotic
6	Right meeting of parietal/supraoccipital	18	Right meeting of opercle/hyomandibular
7	Dorsal cranium – 1st vertebrae	19	Right meeting of opercle/interopercle
8	Left epiphyseal bar	20	Right posterior opercula
9	Left meeting of frontal/parietal	21	Right meeting of dentary/quadrate
10	Left posterior pterotic	22	Anterior hyoid
11	Left meeting of opercle/hyomandibular	23	Posterior hyoid
12	Left meeting of opercle/interopercle		

For an example of landmark placement see Fig. 4.

MATERIALS AND METHODS

Generation of mutant zebrafish (*Danio rerio*) and collection of MicroCT scans were performed as part of a previous study (Watson et al., 2020). Fish were scanned using a vivaCT40 MicroCT scanner (Scanco Medical, Switzerland), with 21 mm isotropic voxel resolution, 55kVp, 145 mA, 1024 samples, 500proj/180°, 200 ms integration time (Watson et al., 2020). We used a total of 23 wild-type fish from two clutches ('wildtype fish') to build our atlas and used 12 *bmp1a* somatic mutants ('*bmp1a* fish'), from a single clutch. To eliminate clutch-effect, we included only wild-type fish (N=14) from the same clutch as the *bmp1a* mutants used in this study to analyze shape differences between groups, as such the final number of samples in the analysis was 26, not including the atlas. Watson et al. (2020) performed a comparison of *bmp1a* somatic and germline mutants and showed that somatic *bmp1a* mutants recapitulate germline *bmp1a* mutant phenotypes but possess additional phenotypic variability due to mosaicism. We focused our analyses on *bmp1a* somatic mutants as they provide a real-world sample of phenotypic variability likely to be encountered in CRISPR-based reverse genetic screens (Shah et al., 2015; Watson et al., 2020).
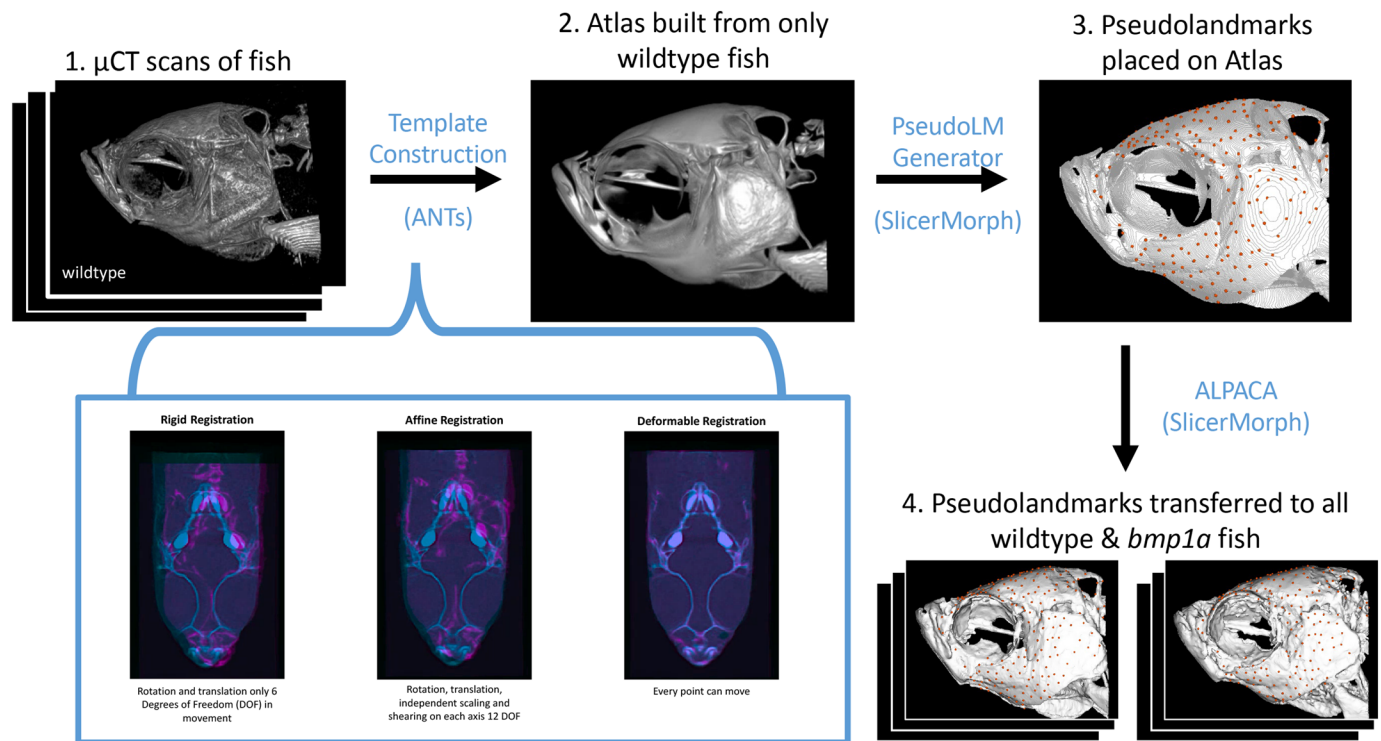
Atlas building

To investigate potential asymmetric patterns, we built a symmetrical atlas of wildtype fish (N=23) by first reflecting all MicroCT volumes along the sagittal plane using the reflectImage function of the AntsR package in R (Avants, 2020). For each of the 23 fish scanned, two volumes (original and reflected) were used to build the atlas to minimize asymmetries between right and left sides of the atlas. A fully symmetrized atlas was generated using the antsMultivariateTemplateConstruction2.sh script as provided by
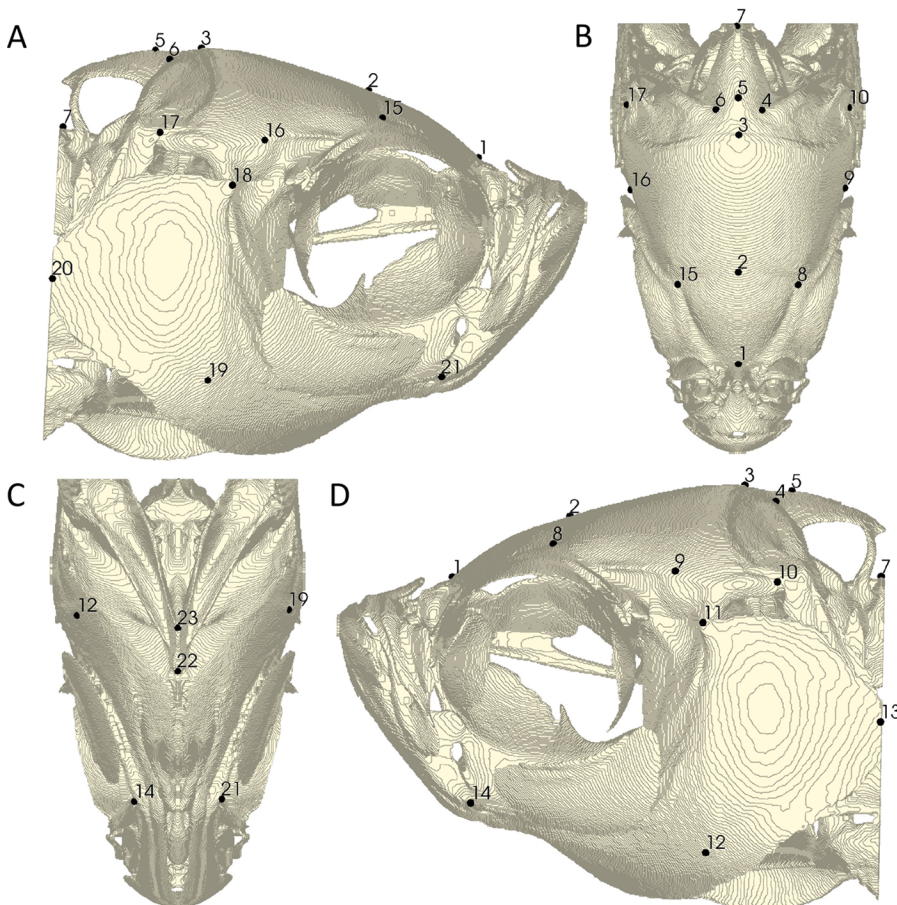
Fig. 4. Gold standard of 23 manual landmarks. Landmark placement represents the average location of two independent landmark placements by the same author. Right lateral (A), dorsal (B), ventral (C), and left lateral (D) views of the atlas mesh are shown. Landmark definitions can be found in Table 3.

Fig. 5. Pipeline for atlas building, pseudo-landmark generation, and transferring pseudo-landmarks to individual fish. Blue text notes the software used between each step. (1) Starting with µCT scans of wild-type fish, ANTs uses a series of rigid, affine, and deformable registrations to create an average image, or (2) Atlas. The PseudoLMGenerator tool in SlicerMorph was used to (3) place 372 pseudo-landmarks on the atlas. The ALPACA tool in SlicerMorph was used to (4) transfer points from the atlas to wild-type and bmp1a fish for comparisons between groups.

the Advanced Normalization Tools (ANTs). For a full explanation of adjustable parameters see Avants et al. (2009). The following settings were used to generate the atlas: type of transformation: greedy symmetric normalization (SyN); similarity metric: cross correlation (CC); iterations: 4; N4 bias field correction: off; smoothing factor: 3×2×1; shrinkage factor: 6×4×2. This atlas building script estimates the best average shape by iteratively estimating an average, or template, and computing deformation fields that map each image to the template. The resultant deformation fields are applied to samples, and a new average is estimated and then used as a new reference for the next step of registrations. Four iterations were sufficient to obtain a symmetrical and anatomically detailed template from which individual bones and landmark locations could be identified.

Atlas validation

To validate our atlas, we compared automated and manual segmentations and landmarks. Before continuing we would like to note the underlying assumption of these comparisons, that the manual dataset represents a ground truth, is likely incorrect (Robinson and Terhune, 2017), and all analyses should be considered in this context.

To quantitatively validate the atlas and our computational anatomy (CA) framework, we first created manual segmentations of individual otoliths from every sample using the open-source 3D Slicer application (Fedorov et al., 2012). We chose otoliths because they are dense, spread out along the dorsoventral axis of the crania, and do not touch any bones, which minimized the potential for human error (or interpretation) in our manual segmentations that serve as the ground truth data. The otoliths from the atlas were segmented in the same manner. Using the antsApplyTransforms function of the ANTsR package (Avants, 2020), for each image in our sample we applied the deformable transformation field, generated during the atlas building step, to the manual atlas otolith segmentation. This function essentially maps the atlas otolith segmentation into the subject space of each image by inversing the calculated transformation fields, effectively creating an automated segmentation of the otoliths for each fish. From this mapping, we calculated the volumes of CA derived segmentation and statistically

compared them to ground truth manual segmentations using Welch two-sample t-tests (Table 1). We also visually inspected all automatically segmented volumes in 3D Slicer to ensure they were segmenting the full and correct structures. All statistical analysis and image registrations were done using the R extensions of the ANTs ecosystem (Avants, 2020).

To further validate our atlas and overall pipeline, we placed 23 traditional landmark points on the meshes of each specimen in our sample (Table 3, Fig. 4). Each specimen was landmarked twice by the same author (K.M.D.) and the average of the two were used as the gold standard for comparison. To validate the ALPACA transfer of points, which uses linear and deformable point cloud registration, we used ALPACA to transfer the 23 gold standard points from the atlas to all meshes in the study (Porto et al., 2021) and calculated the Euclidean distance between ALPACA transferred points to the original gold standard landmark points. We also calculated the Euclidean distances between the gold standard landmark points and each of the two manual landmark placements to establish a comparison between the automated and manual landmarking methods and used an ANOVA to test for differences in Euclidean distances from the gold standard among methods overall. Paired Mann–Whitney U test was used to test for differences in digitization errors for each landmark (Table 2).

To compare object sizes and shapes among our methods and groups, we performed a joint generalized Procrustes superimposition on the combined set of gold standard and ALPACA transferred landmark points (Rohlf and Slice, 1990) and ran a Procrustes ANOVA with landmarking method as a factor using the procD.lm function in the geomorph R package (Adams and Otárola-Castillo, 2013). Using the joint superimposition, we also tested whether Procrustes variances of each method are significantly different from each other using a permutation procedure where the vectors of residuals are randomized among methods using the morphol.disparity function in the geomorph R package (Adams and Otárola-Castillo, 2013). Finally, we compared the normalized centroid size, or centroid size divided by the number of landmarks for each method (Toussaint et al., 2021). Centroid size was calculated for each method using the Geomorph package in R (Adams and Otárola-Castillo, 2013).

Analysis of ZF cranial shape difference in wild types and mutants

To identify regions that might differ outside of the manual landmarks, we opted to use a pseudo-landmark-based analysis between *bmp1a* and wild-type fish. To place pseudo-landmark points (pLMs) on each of our specimens, we first created 3D models of from our CT volumes using the Segment Editor module of 3D Slicer (Fedorov et al., 2012). To generate a set of pLMs on our atlas model, we used the PseudoLMGenerator module in the SlicerMorph extension of 3D Slicer which uses the original mesh geometry and a sagittal plane as the axis of symmetry, to generate a dense symmetric set of surface points (Rolfe et al., 2021). One author (K.M.D.) then went through the pLMs and removed points that were on both jaws and the pectoral girdle using the MarkupEditor tool in 3D Slicer (Rolfe et al., 2021; Fig. 5). Both of these structures are highly prone to post-mortem deformation due to handling and preservation, as such they represent confounding non-biological variation and should be excluded from the analysis. The final number of pLMs on the template was 372. To transfer the pLMs from the atlas to all other models in the study, we used the ALPACA module in the SlicerMorph extension of 3D Slicer (Porto et al., 2021). We skipped the optional scaling step of the ALPACA pipeline as all of our samples were of similar size and used default settings (see Porto et al., 2021 for full details on parameter options) to transfer pLMs from the atlas to all meshes in our sample (Fig. 5; Fig. S2).

To examine differences between *bmp1a* and wild-type fish, we ran a Generalized Procrustes Analyses (GPA) on the all pLMs, allowing all pLMs to slide along the surface using the gpagen function of the geomorph (Adams and Otárola-Castillo, 2013). We ran a symmetry analysis on the GPA coordinates using the bilat.symmetry function in geomorph (Adams and Otárola-Castillo, 2013). From this output, we first ran Procrustes ANOVAs using the procD.lm function in geomorph to determine if the symmetric and fluctuating asymmetric components of shape variation differ between groups. We also ran separate principal components analyses on both the symmetric and asymmetric components of variation from the symmetry analysis using the geomorph (Adams and Otárola-Castillo, 2013). Visualizations were created in the SlicerMorph extension of 3D Slicer (Rolfe et al., 2021) and using ggplot in R (Wickham, 2016).

References
Adams, D. C. and Otárola-Castillo, E. (2013). Geomorph: An r package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393-399. doi:10.1111/2041-210X.12035
Asharani, P. V., Keupp, K., Semler, O., Wang, W., Li, Y., Thiele, H., Yigit, G., Pohl, E., Becker, J., Frommolt, P. et al. (2012). Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and Zebrafish. *Am. J. Hum. Genet.* **90**, 661-674. doi:10.1016/j.ajhg.2012.02.026

Ashburner, J. and Friston, K. J. (2000). Voxel-based morphometry - The methods. *Neuroimage* **11**, 805-821. doi:10.1006/nimg.2000.0582
Avants, B. B. (2020). ANTsR: ANTs in R: Quantification tools for biomedical images version 0.5.7.5.
Avants, B. B., Tustison, N. J. and Song, G. (2009). Advanced normalization tools (ANTS). *Insight J.* **2**, 1-35.
Busse, B., Galloway, J. L., Gray, R. S., Harris, M. P. and Kwon, R. Y. (2019). Zebrafish: an emerging model for orthopedic research. *J. Orthop. Res.* **38**, 925-936. doi:10.1002/jor.24539
Dagdeviren, D., Tamimi, F., Lee, B., Sutton, R., Rauch, F. and Retrouvey, J.-M. (2019). Dental and craniofacial characteristics caused by the p.Ser40Leu mutation in IFITM5. *Am. J. Med. Genet.* **176**, 65-70. doi:10.1002/ajmg.a.40383
Darvann, T. A., Hermann, N. V., Demant, S., Larsen, P., Ólafsdóttir, H., Thorup, S. S., Zak, M., Lipira, A. B., Kane, A. A., Govier, D. et al. (2011). Automated quantification and analysis of facial asymmetry in children with arthritis in the temporomandibular joint. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1193-1196.
Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., JC, F.-R., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M. et al. (2012). 3D Slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging* **30**, 1323-1341. doi:10.1016/j.mri.2012.05.001
Gistelinck, C., Kwon, R. Y., Malfait, F., Symoens, S., Harris, M. P., Henke, K., Hawkins, M. B., Fisher, S., Sips, P., Guillemyn, B. et al. (2018). Zebrafish type I collagen mutants faithfully recapitulate human type I collagenopathies. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E8037-E8046. doi:10.1073/pnas.1722200115
Guimond, A., Meunier, J. and Thirion, J. (2006). Average Brain Models : A Convergence Study To cite this version : Average Brain Models: A Convergence Study.
Hur, M., Gistelinck, C. A., Huber, P., Lee, J., Thompson, M. H., Monstad-Rios, A. T., Watson, C. J., McMenamin, S. K., Willaert, A., Parichy, D. M. et al. (2017). MicroCT-Based Phenomics in the Zebrafish Skeleton Reveals Virtues of Deep Phenotyping in a Distributed Organ System. *Elife* **6**, e26014. doi:10.7554/eLife.26014
Kimmel, C. B., Ullmann, B., Walker, C., Wilson, C., Currey, M., Phillips, P. C., Bell, M. A., Postlethwait, J. H. and Cresko, W. A. (2005). Evolution and development of facial bone morphology in threespine sticklebacks. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5791-5796. doi:10.1073/pnas.0408533102
Kimmel, C. B., Wind, A. L., Oliva, W., Ahlquist, S. D., Walker, C., Dowd, J., Blanco-Sánchez, B., Titus, T. A., Batzel, P., Talbot, J. C. et al. (2020). Transgene-mediated skeletal phenotypic variation in zebrafish. *J. Fish Biol.* **98**, 956-970. doi:10.1111/jfb.14300
Kwon, R. Y., Watson, C. J. and Karasik, D. (2019). Using zebrafish to study skeletal genomics. *Bone* **126**, 37-50. doi:10.1016/j.bone.2019.02.009
Maga, A. M., Tustison, N. J. and Avants, B. B. (2017). A population level atlas of Mus musculus craniofacial skeleton and automated image-based shape analysis. *J. Anat.* **231**, 433-443. doi:10.1111/joa.12645
Mandal, P. K., Mahajan, R. and Dinov, I. D. (2012). Structural brain atlases: Design, rationale, and applications in normal and pathological cohorts. *J. Alzheimer's Dis.* **31** Suppl. 3, S169-S188. doi:10.3233/JAD-2012-120412
Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B. et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 1293-1322. doi:10.1098/rstb.2001.0915
Miller, C. T., Swartz, M. E., Khuu, P. A., Walker, M. B., Eberhart, J. K. and Kimmel, C. B. (2007). mef2ca is required in cranial neural crest to effect Endothelin1 signaling in zebrafish. *Dev. Biol.* **308**, 144-157. doi:10.1016/j.ydbio.2007.05.018
Navon, D., Male, I., Tetrault, E. R., Aaronson, B., Karlstrom, R. O. and Craig Albertson, R. (2020). Hedgehog signaling is necessary and sufficient to mediate craniofacial plasticity in teleosts. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 19321-19327. doi:10.1073/pnas.1921856117
Payan, P., De Pontual, H., Bœuf, G. and Mayer-Gostan, N. (2004). Endolymph chemistry and otolith growth in fish. *Comptes Rendus - Palevol* **3**, 535-547. doi:10.1016/j.crpv.2004.07.013
Porto, A., Rolfe, S. M. and Maga, A. M. (2021). ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. *Methods Ecol. Evol.* **12**, 2129-2144. doi:10.1111/2041-210X.13689
Powder, K. E. and Albertson, R. C. (2016). Cichlid fishes as a model to understand normal and clinical craniofacial variation. *Dev. Biol.* **415**, 338-346. doi:10.1016/j.ydbio.2015.12.018
Robinson, C. and Terhune, C. E. (2017). Error in geometric morphometric data collection: Combining data from multiple sources. *Am. J. Phys. Anthropol.* **164**, 62-75. doi:10.1002/ajpa.23257
Rohlf, F. J. and Slice, D. (1990). Extensions of the procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* **39**, 40-59. doi:10.2307/2992207
Rolfe, S., Pieper, S., Porto, A., Diamond, K., Winchester, J., Shan, S., Kirveslahti, H., Boyer, D., Summers, A. and Maga, A. M. (2021).

Biology Open

SlicerMorph: an open and extensible platform to retrieve, visualize and analyze 3D morphology. *Methods Ecol. Evol.* **12**, 1816-1825. doi:10.1111/2041-210X.13669

Shah, A. N., Davey, C. F., Whitebirch, A. C., Miller, A. C. and Moens, C. B. (2015). Rapid reverse genetic screening using CRISPR in zebrafish. *Nat. Methods* **12**, 535-540. doi:10.1038/nmeth.3360

Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution (N. Y).* **62**, 3135-3156. doi:10.1111/j.1558-5646.2008.00519.x

Toussaint, N., Redhead, Y., Vidal-García, M., Lo Vercio, L., Liu, W., Fisher, E. M. C. C., Hallgrímsson, B., Tybulewicz, V. L. J. J., Schnabel, J. A., Green, J. B. A. A. et al. (2021). A landmark-free morphometrics pipeline for high-resolution phenotyping: application to a mouse model of Down syndrome. *Development* **148**, 1-12. doi:10.1242/dev.188631

Watson, C. J., Monstad-Rios, A. T., Bhimani, R. M., Gistelinck, C., Willaert, A., Coucke, P., Hsu, Y. H. and Kwon, R. Y. (2020). Phenomics-based quantification of CRISPR-induced mosaicism in Zebrafish. *Cell Syst.* **10**, 275-286.e5. doi:10.1016/j.cels.2020.02.007

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.