Morphological and genetic correlates in the left-right asymmetric scale-eating cichlid fish of Lake Tanganyika

FRANCESCA RAFFINI1,2†, CARMELO FRUCIANO1,3,†‡ and AXEL MEYER1,2,4 *

1Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätsstrasse 10, 78464 Konstanz, Germany
2International Max Planck Research School (IMPRS) for Organismal Biology, University of Konstanz, Konstanz, Germany
3School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia
4Radcliffe Institute for Advance Study, Harvard University, 9 Garden Street, Cambridge, MA 02138, USA

Received 15 November 2017; revised 18 February 2018; accepted for publication 18 February 2018

The stable polymorphism in mouth asymmetry in the cichlid fish Perissodus microlepis is a textbook example of adaptive evolution accomplished by functionally relevant morphological changes, ecological specialization and negative frequency-dependent selection. Knowledge about the morphological and developmental basis of this stable polymorphism and the mechanisms driving intraspecific variation in this fish remains largely incomplete. Here, we focus on often-neglected but potentially important aspects of the biology of this fish. In particular, we explore patterns of body shape variation, neutral genome-wide genetic diversity across its geographical distribution, and the presence of asymmetry in eyes centroid size in relationship to mouth bending angle. Geographical space (or a factor associated with geography) has a significant effect on both morphological and genetic diversity, suggesting restricted gene flow across the range of this species. We discuss potential implications of these observed patterns, including the possibility that the genetic basis of asymmetry could vary among locations. A significant association between eye and mouth laterality suggests that the entire head might be involved in the asymmetry. These findings highlight that head asymmetry in P. microlepis is a complex polymorphism involving the interaction of the genetic basis of multiple potentially independent traits and the environment.


INTRODUCTION

Adaptive radiations of cichlid fishes in East Africa provide extraordinary model systems for the study of many fundamental questions in evolutionary biology. Cichlids are astonishingly rich in species, adaptations and ecological specializations, many of which evolved repeatedly in each of the three large species flocks of east Africa (reviewed by Meyer, 1993; Kocher, 2004; Salzburger & Meyer, 2004; Seehausen, 2006; Salzburger, 2009; Takahashi & Koblmüller, 2011; Henning & Meyer, 2014). Cichlids radiations are characterized by the rapid evolution of adaptive feeding modes through morphological changes that permitted ecological specializations (Fryer & Iles, 1972; Kerschbaumer & Sturmbauer, 2011; Takahashi & Koblmüller, 2011). This led to an expansion and successful exploitation of many trophic niches and to effective resource partitioning, facilitating coexistence (reviewed by Sturmbauer, 1998; Gross, Krutzler & Carlson, 2014).

One of the most notable examples of extreme ecological specialization coupled with functionally relevant morphological adaptation is Perissodus microlepis, a scale-eating cichlid fish endemic to Lake Tanganyika.
Tanganyika (Hori, 1993). Its frequency-balanced polymorphism in the morphology of the mouth, which is not symmetrically placed at the centre of the snout as in most fish but bends asymmetrically either to the left or to the right, is a particularly astonishing adaptation to its feeding behaviour (Hori, 1993). In this scale-eating species, individuals with a right-bending mouth (R morph) preferentially feed on scales from the left side of the prey, whereas fish possessing a mouth that bends to the left (L morph) mainly attack the prey’s right flank. Mouth asymmetry, when attacking the appropriate side, increases the area of contact between the predator’s oral cavity and the prey’s flank (Nshombo, Yanagisawa & Nagoshi, 1985; Hori, 1993; Takahashi, Moriwaki & Hori, 2007; Palmer, 2010; Van Dooren, van Goor & van Putten, 2010; Lee, Kusche & Meyer, 2012; Takeuchi, Hori & Oda, 2012; Takeuchi et al., 2016). This morphological and behavioural laterality contributes to increasing hunting success, which is generally very low in scale eaters (~20%; Hori, 1987; Takeuchi et al., 2012). The stable presence over time of both L and R mouth phenotypes within natural populations is thought to be maintained by negative frequency dependence, through the advantage of the less frequent morph, which better eludes prey’s surveillance since victims tend to guard more carefully the side that is more frequently attacked by the more abundant morph (Hori, 1993; Nakajima, Matsuda & Hori, 2004). Owing to the ecological and evolutionary effects of its lateralized morphology and behaviour, P. microlepis became one of the best-known textbook examples for ecological specialization and balanced polymorphism via negative frequency-dependent selection due to prey–predator interactions (Futuyma, 2009).

Recently, however, this famous model has been challenged by several lines of evidence. These include the presence of nearly symmetrical individuals and a likely polygenic nature of mouth asymmetry and, possibly, lateralized behaviour (Hori, 1991; Lee et al., 2010; Palmer, 2010; Stewart & Albertson, 2010; Van Dooren et al., 2010; Kusche, Lee & Meyer, 2012; Lee, Heim & Meyer, 2015; Raffini et al., 2017; Raffini F, Fruciano C, & Meyer A, unpublished), in contrast to the previously reported bimodal distribution and a simple Mendelian inheritance system that was initially hypothesized and which more easily explained how negative frequency-dependent selection acts on this polymorphism (Hori, 1993; Hori, Ochi & Kohda, 2007; Stewart & Albertson, 2010). To accommodate these findings, a more complex picture is emerging, suggesting that the genetic basis of asymmetry also interacts with other processes, potentially including intraspecific competition, physiological trade-offs, as well as random and non-random environmental effects (Palmer, 2010; Stewart & Albertson, 2010; Takeuchi et al., 2016; Van Dooren et al., 2010; Raffini F, Fruciano C, & Meyer A, unpublished).

Despite the substantial advances achieved in recent years, our understanding of the morphological and developmental basis of asymmetry and of the mechanisms that determine and drive intraspecific variation in this fish remain largely incomplete. The overall goal of this study is to contribute to a more complete knowledge of this study model, by providing a multi-level analysis of patterns and factors potentially influencing divergence in P. microlepis. To this end, we did not only consider variation in mouth asymmetry but also other processes and structures that might contribute to intraspecific polymorphism in this fish, a focus that has been overlooked in previous studies. To date, investigations of P. microlepis have focused mainly on mouth polymorphism and its quantitative genetic, environmental and behavioural covariates. This means that potentially important factors, such as restrictions to gene flow across geographical space and variation in other morphological structures, have been largely ignored. For instance, especially for complex phenotypes, restrictions in gene flow can be extremely important because variation in the same trait can be obtained through non-parallel genetic bases (e.g. Borowsky, 2008; Elmer & Meyer, 2011; Soria-Carrasco et al., 2014; Gross, 2016). Phylogeographical studies in P. microlepis conducted so far have been based on only a small number of markers (mitochondrial DNA and microsatellites; Koblmüller et al., 2009; Lee et al., 2010) and, importantly, on a limited portion of this species’ range, in the South of Lake Tanganyika.

Likewise, variation in other potentially relevant phenotypic traits has been largely ignored in this species. Detailed and inclusive analyses of variation in morphology can reveal cryptic differences in selection pressures or other underlying processes in natural environments. For example, divergence in external body shape is known to be a key component of intra- and interspecific variation and has important fitness consequences (e.g. in fish: Webb, 1978, 1984; Webb & Weihs, 1986; Pakkasmaa & Pihonen, 2000; Pettersson & Hedenström, 2000; Boily & Magnan, 2002; McGuigan et al., 2003; Blake et al., 2005; Fisher & Hogan, 2007; Langerhans, 2009; Rouleau, Glémet & Magnan, 2010; Tytell et al., 2010; Franchini et al., 2014; Fruciano et al., 2016a; Senay et al., 2017).

Another trait of potential extreme interest in P. microlepis is asymmetry in eye size. In fact, in this fish, the degree of lateralized feeding behaviour has been linked to differences in anatomical size and gene expression between the two hemispheres of several brain regions (Lee et al., 2017). Interestingly, the highest neuroanatomical and transcriptional divergence was
observed in the tectum opticum, the visual centre of the brain that processes optical information and interacts with motoneurons. More specifically, in this region the hemisphere processing data from the eye that is facing towards the prey during the attack (i.e. the right hemisphere for the left eye and vice versa) was bigger and had different levels of expression for many genes relative to the other hemisphere (Matsumoto & Hikosaka, 2007; Bianco & Wilson, 2009; Chen et al., 2009; Gutiérrez-Ibáñez et al., 2011; Ichijo et al., 2017; Lee et al., 2017; Mizumori & Baker, 2017). It is also known that eyes are an integrated part of the brain; asymmetry in eye use is linked to lateralized behaviour in fish, and preference for a specific eye was shown to be mirrored in an increase of the volume of the corresponding hemisphere in the tectum opticum (Bisazza, Pignatti & Vallortigara, 1997; Güntürkün, 1997; Bisazza et al., 1998a; Bisazza, Rogers & Vallortigara, 1998b; Facchin, Bisazza & Vallortigara, 1999; De Santi et al., 2001; Vallortigara & Rogers, 2005; Matsui, Takeuchi & Hori, 2013; Rogers, 1989, 2017). Scale-eating fish most probably use their eyes differentially during lateralized attacks, probably facing prey and coordinating the approach with the eye corresponding to the direction of assault, which might be the preferred one (Takeuchi et al., 2016; Ichijo et al., 2017). Eye function and size are tightly correlated in vertebrates; a bigger eye provides a larger visual field at higher resolution (Easter, Johns & Baumann, 1977; Hairston, Li & Easter, 1982; Wilkens & Strecker, 2003; Werner & Seifan, 2006; de Busserolles et al., 2013; Champ et al., 2014; Veilleux & Kirk, 2014). A larger eye on the side towards which the asymmetrical mouth is open probably contributes to improved *P. microlepis* hunting performance, and would thus be favoured by selection. All these lines of evidence suggest that asymmetry in the mouth of *P. microlepis* could be functionally associated with asymmetry in eye size, a correlation that has not been analysed previously in this fish or in other organisms.

Here, we analyse neutral genetic and morphological variation in eye size and body shape in an integrative framework. The aims are to identify genetic and shape variation in geographical space and to put mouth asymmetry in the context of variation of these potentially important morphological traits. In particular, our approach draws power from the analysis of genome-wide single nucleotide polymorphism (SNP) data and advanced morphometric techniques on largely overlapping sets of individuals spanning the distribution range of the species.

**MATERIAL AND METHODS**

Two hundred and sixty-eight wild-caught adult *P. microlepis* specimens were collected from five locations in Zambia (April 2010) and three sites in Congo (September 2013; Fig. 1; Supporting Information, Table S1), and preserved in ethanol at 4°C (Kusche et al., 2012; Raffini et al., 2017). To avoid

![Figure 1. Sampling sites along the northern (Congo) and southern (Zambia) coast of Lake Tanganyika (Africa).](https://academic.oup.com/biolinnean/advance-article-abstract/doi/10.1093/biolinnean/bly024/4951795)
potential artefacts, 36 fish were excluded from analysis (Supporting Information, Table S1) because pictures were not deemed appropriate to recover morphology (e.g. open mouth).

ACQUISITION OF MORPHOMETRIC DATA AND ANALYSIS OF MEASUREMENT ERROR

Photographs of the top of the head and both the left and the right side of each specimen were collected with a copy stand and using needles to reduce arching artefacts (Fruciano, Tigano & Ferrito, 2011b, 2012; Fruciano, 2016). The degree of mouth asymmetry of each individual was retrieved from two previous studies (Raffini et al., 2017; Raffini F, Fruciano C, & Meyer A, unpublished). Briefly, three landmarks on the most anterior part of the eye sockets and the tip of the snout were digitized as $x, y$ coordinates on the top view pictures using tpsDig v. 2.57 (Rohlf, 2015). These points were used to compute the angle at the left ($\alpha_L$) and right ($\beta_R$) eye. The mouth-bending angle was defined as the difference between these two angles ($\alpha_L - \beta_R$ in degrees) and used as a measure of asymmetry. Positive values identify those specimens whose mouth bends towards the left, whereas negative values characterize specimens with the mouth opening towards the right (Kusche et al., 2012; Raffini et al., 2017).

To quantify the body shape of our sample of *Perissodus microlepis*, we digitized a set of 33 points, including landmarks, semi-landmarks and ‘helper points’ (Fig. 2). ‘Helper points’ are points that are treated as sliding semi-landmarks to aid the alignment of other points, then removed as they do not provide additional information (Zelditch, Swiderski & Sheets, 2004; Fruciano et al., 2016b). The configurations of points thus obtained were subjected to a generalized Procrustes analysis with sliding of semi-landmarks (Bookstein, 1997) in tpsRelW 1.54 (Rohlf, 2015).

As substantial non-biological variation can be introduced in geometric morphometric data owing to variation in presentation and digitization (Arnqvist & Martensson, 1998; Fruciano, 2016), and this might be aggravated in *P. microlepis* by artefactual variation attributable to mouth laterality, we carried out, on a subset of 20 specimens, a preliminary analysis of measurement error (Fruciano, 2016). This allowed us to choose as the experimental design two presentations (pictures) per side and two digitizations per presentation, for a total of eight configurations of points per specimen (two sides × two presentations × two digitizations = eight). We also determined that the procedure to remove body arching proposed by Valentín et al. (2008), otherwise useful and successfully used in other studies of fish body shape (Fruciano, Tigano & Ferrito, 2011a; Fruciano et al., 2012, 2014, 2016a; Franchini et al., 2014; Ingram, 2015), was not appropriate in our case because it was interacting non-linearly with the variation attributable to mouth bending (data not shown). After their alignment with sliding of semi-landmarks, the configurations of points (eight repetitions per specimen) were averaged across repetitions/sides by specimen, thus reducing measurement error (Arnqvist & Martensson, 1998; Fruciano, 2016). Body centroid size was also computed in tpsRelW for each repetition and averaged by specimen.

To ensure that the experimental design we had chosen was appropriate, we also estimated measurement error on our final dataset of eight configurations of points per specimen. To this end, we performed in MorphoJ (Klingenberg, 2011) a Procrustes ANOVA (Klingenberg & McIntyre, 1998; Klingenberg, Barluenga & Meyer, 2002) on the eight repetitions of both sides combined to quantify the relative contribution of variation among individuals, variation between sides of the same individual, and variation attributable to presentation and digitization error. We also performed a Procrustes ANOVA using individual as the only factor, so that we could compute an analogue of the intraclass correlation coefficient (repeatability), as suggested by Fruciano (2016).

To account for allometric variation in shape, we performed a multivariate regression of body shape on body centroid size in MorphoJ and used regression residuals in subsequent analyses.

MORPHOLOGICAL VARIATION IN GEOGRAPHICAL SPACE

To study body shape variation in geographical space, we used a set of complementary approaches. Using the R package Morpho v. 2.5.1 (Schlager, 2017), we performed

Figure 2. Landmarks (red filled circles), semi-landmarks (blue filled triangles) and helper points (blue open triangles) used to analyse the variation in body shape across *Perissodus microlepis* populations and asymmetry of the eyes. Here, helper points aid in the alignment of semi-landmarks for the eyes, thus preventing a variation in the digitization of these semi-landmarks relative to the antero-posterior axis of the body from being included in the analysis as a genuine variation in shape.
pairwise permutation tests (10 000 permutations) for the difference in mean shape between sampling sites. Given that we have a nested sampling scheme, we also performed a permutational MANOVA (Anderson, 2001), as implemented in the R package geomorph (Adams & Otárola-Castillo, 2013) with a residual randomization procedure for hypothesis testing (Collyer, Sekora & Adams, 2015), using country as the main factor and location nested within country. As an exploratory tool, we used between-group principal component analysis (PCA; Boulesteix, 2005), as implemented in Morpho (Schlager, 2017). Scatterplots of between-group principal component scores are increasingly used in geometric morphometrics (Firmat et al., 2012; Seetah, Cardini & Miracle, 2012; Fruciano et al., 2014, 2016b, 2017; Schmieder et al., 2015; Franchini et al., 2016) as an alternative to scatterplots of canonical variate scores, because scatterplots of between-group principal component scores do not exaggerate the extent of separation between groups (Mitteroecker & Bookstein, 2011). Here, as our samples came from distinct countries, Congo and Zambia, we performed between-group PCA using both sampling site and country as groups and we repeated the analysis within countries (Congo or Zambia; Anderson, 2001; Adams & Otárola-Castillo, 2013; Collyer et al., 2015). In addition to these ‘spatially naïve’ analyses, we also performed spatially explicit analyses (i.e. statistical analyses that explicitly incorporate spatial information). Spatially explicit approaches are rare in geometric morphometrics (but see Cardini, Jansson & Elton, 2007; Fruciano et al., 2011a) but have the advantage of describing spatial patterns better and more directly. Here, we use bearing analysis (Falsetti & Sokal, 1993), as implemented in PASSaGE 2 (Rosenberg & Anderson, 2011) to test for the presence of a phenotypic cline in geographical space using morphometric data (Fruciano et al., 2011a). Briefly, this analysis tests the correlation between geographical and phenotypic distances using a Mantel test (Mantel, 1967). However, pairwise geographical distances in bearing analysis are weighted relative to an a priori defined direction in geographical space before the Mantel test. A high and significant correlation, then, suggests a cline in geographical space. Here, we tested pairwise distances among individuals for all directions in 5° increments and using Euclidean distances obtained from the morphometric data.

**Analysis of asymmetry in eye size**

To investigate variation in size between the two eyes of each individual, we retrieved the size of each eye of each specimen as the centroid size of the eye’s semi-landmarks and helper points (Fig. 2). Centroid size was computed for both the right and the left eye of each fish and for each of the four repetitions (two pictures and two digitizations per picture).

An exploratory analysis indicated that the most appropriate index to measure the degree of difference in size between the two eyes is the scaled index FA8 (Palmer, 1994), as suggested by Palmer & Strobeck (2003). This index was obtained by calculating the logarithm of the ratio between the left and right eye centroid size for each individual, without the use of variance since our focus was at the individual level and not at the population level. The repeatability of FA8 and centroid size was computed as the intraclass correlation coefficient (Fisher, 1958a). Additionally, we estimated the repeatability of FA8 obtained by averaging repeated measurements (Fleiss & Shrout, 1977; Arnvist & Martensson, 1998). The Shapiro–Wilk normality test (Shapiro & Wilk, 1965), computed in R, did not show a departure from normality (P = 0.208). The relationship between FA8 and the mouth bending angle was assessed using the Pearson correlation coefficient and tested by performing multiple linear regression in R. In particular, our linear model had FA8 as the dependent variable and mouth bending angle, body centroid size (as proxy for fish age, another factor that could contribute to asymmetry; Raffini F, Fruciano C, & Meyer A, unpublished) and their interaction as predictors. The interaction was visualized by plotting the effect of one predictor variable on the coefficients of the other using the R package interplot (Solt & Hu, 2015).

**Genetic differentiation**

A genome-wide array of SNPs developed in a previous study (Raffini et al., 2017) was used to explore geographical genetic variation. In contrast to our former study, we analysed the three sampling locations in the northern part of Lake Tanganyika (Democratic Republic of Congo) as distinct sampling sites, notwithstanding their small sample size. Forty-four individuals (Supporting Information, Table S1) were individually sequenced using the double-digested restriction-associated DNA tags (ddRAD) sequencing approach (Miller et al., 2007; Baird et al., 2008; Peterson et al., 2012), obtaining 94 717 SNPs after removal of loci significantly deviating from the Hardy–Weinberg equilibrium and of the locus containing the SNP 56537-113 that we previously identified as differentiated between the left and right morph (molecular and bioinformatics pipelines described by Raffini et al., 2017). The removal of these loci allowed our patterns to be dominated by neutral genetic variation.
The Stacks v. 1.35 populations module (Catchen et al., 2013) was used to analyse genetic differentiation between locations. The minimum percentage of individuals in a population required to process a locus for that population (-r) was set at 0.4, together with five individual minimum stack depth required for individuals at a locus (-m). Pairwise $F_{ST}$ (Weir & Cockerham, 1984; Nielsen & Beaumont, 2009) and corresponding $P$-values according to the Fisher’s exact test (Fisher, 1958b) were computed for each locus. These $P$-values were corrected for multiple tests in SGOF+ v. 3.8 (Carvajal-Rodriguez & de Uña-Alvarez, 2011) using the Benjamini & Hochberg (1995) correction method. The genome-wide bootstrapped $F_{ST}$ (Wright, 1949; Weir & Cockerham, 1984) for each pairwise comparison was computed using 1000 bootstrap samples with the R library STAMPP (Pembleton, Cogan & Forster, 2013). ARLEQUIN v. 3.5 (Excoffier & Lischer, 2010) was used to perform hierarchical analyses of molecular variance (AMOVA, 10 000 permutations; Excoffier, Smouse & Quattro, 1992) with locations nested in country. Accordingly, genetic variation is partitioned into four components: between countries; among locations within country; among individuals within locations; and within individuals. For this analysis, three random subsets of 10 000 SNPs were generated through the procedure reported in the Stacks documentation.

The correlation between individual genetic (Prevosti distance calculated in R v. 3.3.3 using the package poppr v. 2.2.1; Prevosti, Ocaña & Alonso, 1975; Kamvar, Tabima & Grünwald, 2014; R Core Team, 2016) and spatial (coastal distances in kilometres) distances was tested by performing a Mantel test (R package vegan v. 2.4-1; Oksanen et al., 2016) using 999 permutations and the Spearman correlation. Maximum likelihood estimation of individual ancestries was run in Admixture v. 1.23 (Alexander, Novembre & Lange, 2009). This analysis was performed for $K$ (number of clusters) ranging from one to ten and implementing the cross-validation (CV) procedure to identify the most likely subdivision in clusters. The outputs of these tests were plotted using R. A PCA was computed using the R library adegenet v. 1.4-2 (Jombart & Ahmed, 2011). As a further tool to visualize the genetic relationship between the geographical sites, we computed a neighbor-joining (NJ) tree (Saitou & Nei, 1987) based on the Prevosti distance using the R library ape (Paradis, Claude & Strimmer, 2004). The same matrix of distances was used to perform a bearing analysis. Demographic parameters, such as the number of segregating sites and the neutrality indices Tajima’s $D$, $F$ and $D$ and $F$ and $D$ (Tajima, 1989; Fu & Li, 1993), were computed both individually for each population and pooling individuals by country (Congo or Zambia) using the R package PopGenome (Pfeifer et al., 2014). The populations Bemba and Bangue were excluded from neutrality tests due to their small sample sizes.

**COMBINED ANALYSIS OF GENETIC AND MORPHOMETRIC DATA**

For the fish with both genetic and morphometric data ($N = 44$; Supporting Information, Table S1), we computed the correlation between Euclidean distances of the morphometric dataset and Prevosti distances of the genetic dataset. We then tested this correlation with a Mantel test. Furthermore, we tested for the correlation of genetic and morphometric distances while controlling for geographical linear distances using a partial Mantel test (Smouse, Long & Sokal, 1986). As Mantel and partial Mantel tests have encountered much criticism (e.g. Legendre & Fortin, 2010; Guillot & Rousset, 2013), we used a complementary and conceptually similar approach to ensure the robustness of our results. Here, we performed a principal coordinate analysis on the matrix of Prevosti distances, retaining the first 27 axes (i.e. all the dimensions accounting for $\geq 1\%$ of variance). Then we fitted two separate linear models using as a predictor the matrix of latitude and longitude coordinates and as response the shape data and the principal coordinates of the genetic data, respectively. We then used a two-block partial least squares analysis (Rohlf & Corti, 2000), as implemented in geomorph, to assess and test for the association of the two sets of residuals (shape and genetic) obtained from the linear models.

**RESULTS**

The Procrustes ANOVA we performed to quantify measurement error and variation between sides revealed a substantial and significant proportion of variation attributable to side and presentation (picture), with $\sim 10\%$ of variance explained by each of the terms side, individual × side and presentation (Supporting Information, Table S2). Repeatability across the eight configurations per individual (two sides × two pictures × two digitizations) was moderate (0.75). Overall, these results on the full dataset reinforce our choice, based on a preliminary analysis, of taking repeated measurements and averaging them as a sensible method to reduce measurement error.

**VARIATION IN GEOGRAPHICAL SPACE**

Performing exploratory between-group PCA of morphological differentiation using all sampling sites with country as a group reveals a very high degree
of overlap among sampling sites (Fig. 3A), but some level of separation of specimens from Zambia and specimens from Congo (Fig. 3B). This pattern becomes clearer when performing a between-group PCA using country as a group, with scores along the first between-group principal component showing distinct (but overlapping) distributions for Zambia and Congo (Fig. 3B, C). The direction of variation between country means shows that, after controlling for allometry, *P. microlepis* from Congo are slightly more elongated compared with fish from Zambia (Fig. 3D). Performing a between-group PCA within countries reveals a good separation only between Congo sampling sites (Fig. 3E, F). The bimodal distribution of the scores on the country between-group PCA for the Congolese samples (Fig. 3C) is probably attributable to the good separation of the Congolese sampling sites from each other (Fig. 3E) and the fact that the two between-group principal component axes computed for Congo samples only (Fig. 3E) are relatively similar (i.e. account for similar shape change; data not shown) to the between-group principal component axis obtained using country as a group.

Tests of difference in mean shape between sampling sites (Table 1) reveal overall significant differences among sampling sites, with non-significant differences restricted to sites with small sample sizes, where the absence of a significant difference in means can be attributed to lack of statistical power. Significant differences in mean shape between countries \((P < 0.0001)\) mirror the exploratory results obtained with between-group PCA. The permutational MANOVA with residual randomization procedure with location as a nested factor within country confirms that, although there are significant differences between countries and between sampling sites within countries, most of the variation among individuals is not accounted for by these two terms, with the residuals explaining \(-84.5\%\) of total variance \(\text{country: mean squares (MS)} = 0.0129, F_{1,266} = 15.400, P = 0.001;\) location nested in country: \(\text{MS} = 0.0045, F_{5,266} = 5.352, P = 0.001;\) residuals: \(\text{d.f.} = 259, 266, MS = 0.0008\). Bearing analysis failed to detect a clear cline in geographical space (Supporting Information, Fig. S1A). In fact, except for a small range of directions \((15–25^\circ)\) where correlation is even lower, for all the other angles tested the correlation was significant but always low and with a restricted range of values \((0.16–0.19)\). This suggests that there is a relationship between geographical and morphological distances, but not along a precise cline.

The analyses of genome-wide differentiation between sampling sites were all highly significant (Table 2). The \(F_{\text{ST}}\) values increased with growing coastal distance, except that the Congo locations show lower differentiation from Mbita than the other Zambian populations. This overall trend was in agreement with the results of the Mantel test, exhibiting a significant correlation between genetic and spatial distances (Mantel’s \(r = 0.625; P = 0.001\)).

The AMOVA analyses indicated that differentiation between countries and locations is significant and accounts for 30 and 5\% of genetic variation respectively; 48 and 17\% of genetic variance is explained by the between- and within-individual component, respectively (Table 3). Principal component analysis and NJ tree (Supporting Information, Figs S2, S3) were concordant in suggesting that most of the genetic variation is found between Congo and Zambia, followed by the four Zambian populations, with a considerable overlap between Kasakalawe and Mbita. This was confirmed by the estimation of individual ancestries (Fig. 4), which identified two clusters of individuals corresponding to Congo and Zambia, and three (best supported by the CV procedure; Supporting Information, Fig. S4) or four clusters corresponding to the four Zambia populations, with weak distinction between Kasakalawe and Mbita. The three populations in Congo were not genetically differentiated in any analysis. Negative values of the neutrality tests suggested that Congo and Zambia (population pooled by country) had an excess of rare alleles associated with recent population expansion or the occurrence of positive selection (Supporting Information, Table S3).

At the individual population level, positive neutrality statistics in Lahanga, Katoto and Toby might indicate potential population bottleneck or balancing selection events. Bearing analysis on Prevosti distances revealed a pattern qualitatively similar to the one shown by morphometric distances, with most directions having very similar and significant levels of correlation between genetic distances and geographical distances (Supporting Information, Fig. S1B). The main difference between the results of the bearing analysis based on morphometric and genetic data is that genetic data show higher correlations \((0.59–0.65)\).

**Combined analysis of genetic and morphometric data**

Individuals having both genetic and morphometric data were tested for correlations between morphometric and genetic distances, which were low but significant (Mantel test, \(r = 0.17, P = 0.003\)). However, when controlling for geographical distances using a partial Mantel test, the correlation was nullified and not significant \((r = -0.01, P = 0.56)\). Our complementary analysis based on a PCA of genetic distances, linear models accounting for geographical location and partial least squares analysis to test for association of the residuals mirrored our results based on the partial Mantel test. Indeed, the correlation...
Figure 3. Morphological variation in the geographical space. A, B, plot of the scores along the first two between-group principal components (sampling site used as group), with colour-coding for sampling sites (A) or countries (B). These two axes cumulatively account for 58.54% of variance among individuals. C, density plot of the scores along the between-group principal component using country as the group (this component accounts for 35.29% of the variance among individuals). D, variation in body shape between the two countries (difference in mean shape, exaggerated twofold). E, F, plot of the scores along the first two between-group principal components (sampling sites used as group), analysing each country separately: Congo (E; accounting for 38.67 and 8.85% of variance) and Zambia (F; accounting for 18.67 and 5.05% of variance). Abbreviation: bwgPC, between-group principal component.
The correlation between the singular scores (partial least squares scores) of the genetic and morphometric blocks was low and not significant ($r = 0.38$, $P = 0.39$). These results together suggest that, although both genetic and morphometric data show significant and 'diffused' variation in geographical space (i.e. variation that does not follow a precise cline) and although genetic and morphometric variation are correlated, these patterns are attributable to geographical distances.

Variation in Eye Asymmetry

Next, we analysed the relationship between eye size and mouth asymmetry. The repeatability of the centroid size measurement was 0.94 and 0.92 for the left and right eye, respectively, and it was 0.54 for FA8. Therefore, we used the mean of repeated centroid size measurements to compute FA8, which improved the repeatability of this index to 0.82. Asymmetry in the eye size was significantly associated with mouth laterality (Fig. 5A; Pearson's product–moment correlation: $r = 0.22$, $t = 3.596$, d.f. = 264, $P = 0.0004$; linear regression: $R^2 = 0.047$, $F_{1.264} = 12.93$, $P = 0.0004$). The difference in size between the two eyes increased with growing mouth bending angle, and the direction of asymmetry in the eye mirrored the one in the mouth; that is, individuals with a mouth bending towards the right (negative mouth angle values) showed a bigger right eye (negative FA8 values), and vice versa. Body size had a significant effect on eye asymmetry and on the coefficients of mouth bending angle in the interaction term (Table 4; Fig. 5B, C).

Discussion

Perissodus microlepis is famously known for its intraspecific polymorphism in mouth shape, which is thought to represent an adaptation to its specialized feeding mode (Hori, 1993). Nonetheless, patterns of external morphological variation and neutral genomic diversity in this cichlid fish remain largely unstudied, although they could contribute to a more comprehensive understanding of this striking study system. We combined genomic and geometric morphometric approaches to explore body shape variation and genetic neutral diversity in geographical space. Additionally, we investigated the presence of divergence in eye size and its association with mouth asymmetry. Our results suggest that body morphology...
Table 3. Genetic variation among and within countries, locations and individuals

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Dataset</th>
<th>Random subset 1</th>
<th>Random subset 2</th>
<th>Random subset 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between countries</td>
<td>Sum of squares</td>
<td>42972.288</td>
<td>42403.042</td>
<td>43011.006</td>
</tr>
<tr>
<td></td>
<td>Variance components</td>
<td>1273.667</td>
<td>1251.681</td>
<td>1272.800</td>
</tr>
<tr>
<td></td>
<td>Percentage of variation</td>
<td>30.473</td>
<td>29.913</td>
<td>30.222</td>
</tr>
<tr>
<td>Among locations within country</td>
<td>Sum of squares</td>
<td>36519.379</td>
<td>36814.466</td>
<td>36815.493</td>
</tr>
<tr>
<td></td>
<td>Variance components</td>
<td>203.780</td>
<td>203.440</td>
<td>202.688</td>
</tr>
<tr>
<td></td>
<td>Percentage of variation</td>
<td>4.875</td>
<td>4.862</td>
<td>4.813</td>
</tr>
<tr>
<td>Among individuals within locations</td>
<td>Sum of squares</td>
<td>174231.129</td>
<td>176575.083</td>
<td>176936.989</td>
</tr>
<tr>
<td></td>
<td>Variance components</td>
<td>2006.759</td>
<td>2042.968</td>
<td>2046.148</td>
</tr>
<tr>
<td></td>
<td>Percentage of variation</td>
<td>48.013</td>
<td>48.823</td>
<td>48.586</td>
</tr>
<tr>
<td>Within individuals</td>
<td>Sum of squares</td>
<td>30.599</td>
<td>30.200</td>
<td>30.350.500</td>
</tr>
<tr>
<td></td>
<td>Variance components</td>
<td>695.432</td>
<td>686.364</td>
<td>689.784</td>
</tr>
<tr>
<td></td>
<td>Percentage of variation</td>
<td>16.639</td>
<td>16.403</td>
<td>16.379</td>
</tr>
<tr>
<td>Total</td>
<td>Sum of squares</td>
<td>284321.795</td>
<td>285992.591</td>
<td>287113.989</td>
</tr>
<tr>
<td></td>
<td>Variance components</td>
<td>4179.637</td>
<td>4184.452</td>
<td>4211.420</td>
</tr>
</tbody>
</table>

Hierarchical analyses of molecular variance (AMOVA). All the tests were statistically highly significant ($P < 0.01$).

Figure 4. Genetic relationship between geographical sites: Admixture barplots for the genetic dataset using values of $K$ from two (upper plot) to four (lower plot). Each vertical line represents an individual, and the colours within each line correspond to the inferred proportion of ancestry. Populations are separated by black dotted vertical lines. Abbreviations: Ba, Bangue; Be, Bemba; C, Congo; La, Lahanga; Kas, Kasakalawe; Kat, Katoto; Mb, Mbita; To, Toby; Z, Zambia.
and neutral genetic variation are explained by geography (or a factor associated with geography), and that the mouth bending angle is correlated with the amount and direction of asymmetry in eye size. These findings add to the growing evidence that multiple factors and structures underlie intraspecific polymorphism in *P. microlepis*.

**Geography underlies body morphological and genome-wide variation**

Geography was a major predictor of genetic and morphological variation between sampling sites in *P. microlepis*, although a precise cline was not detected. As expected, the two countries Congo and Zambia, which represent the northern and southern end of the Lake Tanganyika, respectively, and are separated by ~660 km, account for most of the variation. This geographical effect could be linked to multi-directional patterns of isolation-by-distance, as seen at a much smaller spatial scale along Zambian rocky shorelines using mitochondrial DNA (sampling localities separated by a few dozen kilometres; Koblmüller et al., 2009). It could also reflect changes in the network dynamics of populations, for example due to fluctuations in the lake level originating and reshuffling barriers to gene flow. Intermittent connectivity, oscillations in population size and colonization events are known to affect diversity distribution (examples in Lake Tanganyika cichlids: Koblmüller et al., 2011; Nevado et al., 2013; Sefc et al., 2017). Our genetic data showed signatures of population expansion in Congo and Zambia (locations pooled by country), whereas it revealed bottlenecks in three local populations (Lahanga, Katoto and Toby). This partial discrepancy between large and small spatial scale could emerge from reduced sample sizes of non-pooled populations or might reflect differences in microevolutionary histories (e.g. between eastern and western Zambia populations; Koblmüller et al., 2009). Larger sample sizes and estimates of mutation rates might allow discrimination between these two scenarios and provide divergence time estimates to be compared with geological events in the area (e.g. fluctuations in the lake’s level; see, for example, Sefc et al., 2017).

Another factor potentially affecting geographical diversity is the presence of spatial variation in ecological cues (e.g. environmental abiotic and biotic factors) and/or selective regimes, as repeatedly shown in the Tanganyikan cichlid fish radiations (e.g. Clabaut et al., 2007; Day, Cotton & Barraclough, 2008; Clabaut et al., 2007).

**Figure 5.** The relationship between eye asymmetry, mouth asymmetry and body centroid size. A, B, plot of the association between asymmetry in eye size (FA8 index) and degree of mouth bending angle (A) or body centroid size (B). C, interaction between body size and mouth bending angle: interplot showing how a change in body size (as a proxy for fish age) affects the coefficient of mouth asymmetry in the model. Shaded area represents the 96% confidence interval.
Muschick, Indermaur & Salzburger, 2012; Janzen et al., 2017). These between-species comparative studies revealed that environmental and ecological differences play a key role in generating morphological diversity and barriers to gene flow potentially attributable to divergent selection, and it might be also the case for P. microlepis within-species populations. Northern and southern lake shores present different abiotic features, such as water depth or type of bottom (steeper, deeper and considerably more sandy in Congo; Coenen, Hanek & Kotilainen, 1993; International Lake Environment Commitee, 1999–2017). In line with this habitat variation, we found that fish from Congo are on average more streamlined compared with fish from Zambia, similar to the more elongated bodies observed in cichlids inhabiting deeper waters (Clabaut et al., 2007). Additionally, the scale-hunting technique of this scale eater includes quick body flexions and aggressive mimicry (hiding itself between schools of prey or other fish thanks to its close resemblance to these species in overall body shape and colour; Supporting Information, Fig. S5; Hori & Watanabe, 2000; Koblmüller et al., 2007, 2009; Takeuchi et al., 2012). The morphological variation in the body could be influenced by local community structure and mirror the natural phenotypic diversity of prey across Lake Tanganyika. The slender body presented by P. microlepis in northern Congo might be beneficial for improving swimming ability (as shown in other fishes; e.g. Webb, 1978, 1982, 1984; Webb, Kostecki & Stevens, 1984; Webb & Weihs, 1986; Blake et al., 2005; Langerhans, 2009; Rouleau et al., 2010; Senay et al., 2017) and removing scales from prey, which might naturally include faster swimmers or more limetic species. Alternatively, the northern schooling fishes could be more arrow shaped than the Zambian populations and, consequently, P. microlepis evolved a more streamlined body for better camouflage. Differences between fish populations related to aggressive mimicry have been shown for colour polymorphism (e.g. in fish: Munday, 2009; Rouleau, 1978, 1982, 1984; Takeuchi & Oda, 2017). Here, we showed that the eyes, an integrated part of the brain, are asymmetrical in size, and this polymorphism is associated with the asymmetry in the mouth. In the light of the significant relationship between transcriptional and neuroanatomical asymmetry in the tectum opticum and lateralized attacking predilection (Lee et al., 2017) and between eye preference and the volume of the corresponding tectum opticum’s hemisphere shown in other organisms (Rogers, 1989, 2017; Bisazza et al., 1997, 1999a, 1999b; Güntürkün, 1997; Fachin et al., 1999; De Santi et al., 2001; Vallortigara & Rogers, 2005; Matsui et al., 2013), eye asymmetry in P. microlepis could be correlated with both mouth

Asymmetry in neuronal circuits in sensory and/or motor systems (tectum opticum, habenula and M-cells) has recently been linked to lateralized behaviour or mouth asymmetry of P. microlepis (Takeuchi et al., 2012, 2016; Ichijo et al., 2017; Lee et al., 2017; Takeuchi & Oda, 2017). Here, we showed that the eyes, an integrated part of the brain, are asymmetrical in size, and this polymorphism is associated with the asymmetry in the mouth. In the light of the significant relationship between transcriptional and neuroanatomical asymmetry in the tectum opticum and lateralized attacking predilection (Lee et al., 2017) and between eye preference and the volume of the corresponding tectum opticum’s hemisphere shown in other organisms (Rogers, 1989, 2017; Bisazza et al., 1997, 1999a, 1999b; Güntürkün, 1997; Fachin et al., 1999; De Santi et al., 2001; Vallortigara & Rogers, 2005; Matsui et al., 2013), eye asymmetry in P. microlepis could be correlated with both mouth

Table 4. The relationship between eye asymmetry, mouth bending angle and body centroid size as modelled in a general linear regression model

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>d.f., num, den</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouth angle</td>
<td>1, 262</td>
<td>-0.0069</td>
<td>0.0019</td>
<td>-3.602</td>
<td>0.0004</td>
</tr>
<tr>
<td>Body size</td>
<td>1, 262</td>
<td>0.0025</td>
<td>0.0013</td>
<td>2.010</td>
<td>0.0454</td>
</tr>
<tr>
<td>Mouth angle:body size</td>
<td>1, 262</td>
<td>0.0066</td>
<td>0.0001</td>
<td>4.339</td>
<td>2.05e-05</td>
</tr>
</tbody>
</table>

Abbreviations: den, denominator; d.f., degree of freedoms; num, numerator. Statistically significant (P < 0.05) results are highlighted in boldface.

© 2018 The Linnean Society of London, Biological Journal of the Linnean Society, 2018, XX, 1–18
asymmetry (as shown in the present study) and lateralized behaviour. The polymorphism in the eyes might be adaptive, because improved vision (bigger eye) and visual processing (larger tectum opticum) on the side at which scale eaters approach prey is likely to be beneficial for hunting success.

We observed a significant association between body size (a proxy for age; Raffini F, Fruciano C, & Meyer A, unpublished) and asymmetry in both the eyes and the mouth, and a significant interaction between body size and mouth asymmetry; the positive relationship between eye asymmetry and mouth asymmetry becomes stronger in larger fish. Direct and indirect evidence indicates that the amount of mouth asymmetry is influenced by feeding experience (Stewart & Albertson, 2010; Van Dooren et al., 2010; Kusche et al., 2012; Lee et al., 2015; Takeuchi et al., 2016; Raffini F, Fruciano C, & Meyer A, unpublished) and that lateralized behaviour is established during development through practice; fish initially attack prey on both flanks and learn on which side they are more efficient in removing scales, depending on their asymmetrical mouth and lateralized kinetics (Takeuchi et al., 2016; Takeuchi & Oda, 2017). Our findings might also suggest that other morphological structures, such as the eyes, contribute to this dynamic, since asymmetry of the eyes size mirrored that of the mouth and, possibly, behaviour laterality. Vision could be adjusted plastically during growth to match asymmetry in the mouth and thus improve hunting success, or a larger eye on the side towards which the mouth opens might increase the amount of fruitful attempts during correctly directed lateralized attacks, leading to a more asymmetrical mouth due to the plastic effects of the higher number of ingested scales (through the mechanism described by Raffini F, Fruciano C, & Meyer A, unpublished). Owing to the presence of left–right differences in eyes size, the two tectum opticum hemispheres most probably receive and/or process asymmetrical visual information that is then output to downstream cerebral structures. This asymmetrical flow might possibly contribute to the establishment of lateralized neuronal circuits, for example in the habenula or hindbrain (M-cells), which are known to be involved in motor responses to visual stimuli and have been proposed to play a key role in P. microlepis asymmetry (Takeuchi et al., 2016; Ichijo et al., 2017; Lee et al., 2017; Takeuchi & Oda, 2017).

Taken together, these findings on eyes provide the first indication of lateralization in eye size (and, indirectly, eye function) related to asymmetry in the mouth and, possibly, in the brain and behaviour; an association that has been overlooked in this and other organisms. We advocate future studies to explore this multi-level interrelationship further, particularly its genetic or non-genetic basis (at present relatively clear only for mouth asymmetry) and its relative role and timing in the developmental and evolutionary establishment and maintenance of laterality in this cichlid fish.

**CONCLUSION**

In summary, in this study on P. microlepis we have focused on a suite of traits and features largely overshadowed by the striking mouth polymorphism of this species. In particular, we document a reduction in gene flow across Lake Tanganyika, which opens the possibility that the genetic basis of mouth asymmetry could vary among locations. Perhaps most importantly, we show that functionally relevant traits (body shape) vary over large-scale geographical distances and that there is an association between eye and mouth asymmetry, thus suggesting the possibility that variation in these traits interacts with the genetic basis of mouth asymmetry and the environment to produce (and, possibly, to maintain over time) the striking variation in mouth asymmetry of P. microlepis.

**ACKNOWLEDGEMENTS**

We thank Henrik Kusche and Gabriele Legant for help in collecting fishes in 2010, and Stephan Koblmüller and Christian Sturmbauer in 2013. We are grateful to Niclas Kolm and Ralf Schneider for useful suggestions about eye asymmetry. We thank two anonymous reviewers for their helpful comments. F.R. is supported by the International Max Planck Research School (IMPRS). C.F. was partially funded by a Marie Curie IEF Fellowship (Grant Agreement 327875 – PlasticitySpeciation). Funding for this project came from the Deutsche Forschungsgemeinschaft grant ME2848-1-1 (to Hyuk Je Lee and A.M.) and the IMPRS Emergency Grant to F.R.

**AUTHOR CONTRIBUTIONS**

F.R., C.F. and A.M. designed the study. Morphological data were collected by F.R. and analysed by C.F. The analysis of eye asymmetry was performed by F.R. under supervision by C.F. Both F.R. and C.F. analysed the genetic data. F.R. drafted the manuscript. All authors edited and agreed to the manuscript.

**REFERENCES**


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Table S1.** Detailed breakdown of the individual specimens used for the morphometric, genetic and combined analysis (indicated with a X in the dedicated column), respectively.

**Table S2.** Measurement error and variation between sides as quantified by Procrustes ANOVA and by an analogous of the intraclass correlation coefficient (‘repeatability’).

**Table S3.** Demographic analysis: number of segregating sites and neutrality indices for each population and country.

**Table S4.** Regression residuals of shape (Procrustes aligned coordinates) on centroid size, as used in the analyses of shape variation among sites and in the combined analyses of genetic and morphometric data.

**Figure S1.** Bearing plot of the correlation between morphometric and geographical distances (A) and genetic (Prevosti) and geographical distances (B).

**Figure S2.** Genetic relationships between the geographical sites: plot of the scores along the first two principal components of the genetic dataset.

**Figure S3.** Genetic relationships between the geographical sites: neighbor-joining tree.

**Figure S4.** Admixture cross-validation plot. K is the number of clusters.

**Figure S5.** Aggressive mimicry in *Perissodus microlepis*. This cichlid fish (second, fourth, fifth and sixth specimens from top) resembles the poeciliid *Lamprichthys tanganicanus* (first and second from top).

**SHARED DATA**

Data on individual specimens (unique identity, Illumina barcode, sampling location, mouth bending angle and Procrustes aligned coordinates on centroid size) used in this study are provided in the Supporting Information (Tables S1 and S4). Raw Illumina sequences have been archived to the National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA) database with Accession no. PRJNA324541.

© 2018 The Linnean Society of London, Biological Journal of the Linnean Society, 2018, XX, 1–18