ABSTRACT

Hypothesis: Fluctuating asymmetry (FA) – random deviations from perfect symmetry that are used as a measure of developmental stability – is an effective indicator of stress and fitness in threespine stickleback.

Organisms: The threespine stickleback (Gasterosteus aculeatus) and two other species, the brook stickleback (Culaea inconstans) and ninespine stickleback (Pungitius pungitius), were the focus of a review of the literature. In addition, four populations of G. aculeatus – one anadromous population from the Kamchatka River, two marine populations from the White Sea, and one freshwater population from the White Sea basin – were studied in the field.

Methods: A review of the literature relating fluctuating asymmetry to different variables, and a comparison of fluctuating asymmetry in four populations of stickleback, which differed in geographical distribution and life history, using lateral plates and four cranial bones (operculum, lachrymal, third suborbital, quadrate).

Results: An appraisal of the literature on fluctuating asymmetry suggests that decreasing interest in FA studies has likely resulted from conflicting research results. To some extent, this problem is likely caused by the morphological structures used in FA analysis, which are generally limited to the lateral plates and pelvic fins. These structures can evolve quickly in response to various environmental changes, thus their fluctuating asymmetry reflects not only individual fitness and stress, but also multiple uncontrolled factors that may directly affect those same structures. Using four cranial bones in our analysis showed lower fluctuating asymmetry in anadromous stickleback from the Kamchatka Peninsula compared with marine and freshwater stickleback from the White Sea and its basin. This may be caused by more favourable feeding conditions in the North Pacific than in the White Sea. The different environmental conditions at these locations did not appear to have a significant effect on fluctuating asymmetry, although the comparison of freshwater, anadromous, and marine populations showed that the fluctuating asymmetry of the structures we used is responsive to these differences.
FA analysis of the selected bone structures reveals clear heterogeneity in stickleback with different life histories. We suggest that these structures can be considered reliable for studies of fluctuating asymmetry in stickleback fishes.

**Keywords**: cranial bones, fluctuating asymmetry, *Gasterosteus aculeatus*, literature review, threespine stickleback.

**INTRODUCTION**

The stickleback is a model species for studies in ecology and evolution (Woottton, 1984; Bell and Foster, 1994a; Östlund-Nilsson et al., 2007). Morphological analysis plays an important role in many such studies, as it can be used to address various research questions (Bell and Foster, 1994a). One promising morphological analysis is that of fluctuating asymmetry (FA), which measures deviations from perfect morphological symmetry. In the late 1980s and early 1990s, fluctuating asymmetry was presented as a simple, inexpensive, and universal measure of stress and fitness (Palmer and Strobeck, 1986; Zakharov, 1989; Parsons, 1990; Graham et al., 1993). Although there are many definitions of stress, here we follow Graham and colleagues’ understanding of the term as ‘... anything (physical, chemical, genetic, psychological, etc.) dissipating energy away from growth and production’ (2010, p. 501). Such stress can be caused by sub-optimal environmental conditions or by disruption of genetic co-adaptation (Graham et al., 1993).

Fluctuating asymmetry is the most widely used means of measuring developmental instability, which reflects an organism’s inability to develop the same phenotype under the same environmental conditions (Waddington, 1957), i.e. the inability to follow a developmental trajectory defined for a given genotype (Zakharov, 1989). Developmental instability represents a third, stochastic component of phenotypic variance, as important as genotypic variation and environmental heterogeneity (Lajus et al., 2003a). As stress reduces an organism’s available energy for growth (Parsons, 1990, 2005; Hoffmann and Parsons, 1991; Graham et al., 2010) and mechanisms directing growth also require energy (Koehn and Bayne, 1989; Sommer, 1996), stress may also reduce energy allocated to developmental control. Therefore, increased developmental instability, mediated by energy allocation in growing organisms, can be a direct consequence of stress (Lajus et al., 2014).

Research has shown that fluctuating asymmetry can increase due to sub-optimal temperatures, under both experimental (Leary et al., 1992; Campbell et al., 1998; Benderlioglu and Dow, 2017) and natural conditions (Alados et al., 1993; Yuritseva et al., 2014; Kozlov and Zverev, 2018), and high population densities (Wiener and Rago, 1987; Leary et al., 1991). The level of fluctuating asymmetry is inversely associated with growth rate (Zakharov, 1989; Lajus, 2001), resistance to parasites (Leberg and Vrijenhoek, 1994; Escos et al., 1997; Reimechen, 1997; Pojas and Tabugo, 2015), and heterozygosity (Blanco et al., 1990; Leary et al., 1992). Inbreeding can reduce developmental stability due to the expression of deleterious recessive alleles (Leary et al., 1983; Carter et al., 2009; Baker and Hoelzel, 2013). A number of studies have connected fluctuating asymmetry with the effects of pollutants (Valentine et al., 1973; Romanov and Kovalyev, 2004; Green and Lochman, 2006; Lajus et al., 2015a). Several other studies have associated higher fluctuating asymmetry with elevated radiation levels (Zakharov and Krysanov, 1996; Lajus et al., 2014). At the same time, some FA analyses yielded conflicting results that generated a great deal of scepticism as to the method’s usefulness as an indicator of stress and fitness (e.g. Bjorksten et al., 2000a, 2000b; but see Möller, 2000; van Dongen and Lens, 2000).
One reason for these disparate results may be the number and choice of characters used. In studies of fish, including stickleback, only a few characters have generally been utilized, such as the numbers of fin rays, gill rakers, and scales, as well as fin size and eye diameter (Valentine and Soule, 1973; Valentine et al., 1973; Ames et al., 1979; Graham and Felley, 1985; Zakharov, 1989; Leary et al., 1992; Kozhara, 1994; Prieto et al., 2005; Moodie et al., 2007; van Dongen et al., 2009) (see also evolutionary-ecology.com/data/3165Appendix.pdf). Character choice can lead to problems such as sampling error, measurement error, and departure from concordance in the fluctuating asymmetry of different characters across samples (Lajus et al., 2014). One reason for the high sampling error associated with FA studies is the low absolute value of fluctuating asymmetry of some meristic characters. For instance, asymmetry in a number of fin rays may occur in a small proportion of the population only; analysis would then require samples sometimes exceeding a thousand individuals, as in one study relating fluctuating asymmetry to pollution (Michaelsen et al., 2015).

Measurement error is of great concern in FA studies. As the absolute value of fluctuating asymmetry of morphological structures usually approaches only a small fraction of its true size [1–5% according to Merilä and Björklund (1995)], measurement error associated with FA studies can be quite high, often exceeding 50% of variation in fluctuating asymmetry (Hubert and Alexander, 1995; Merilä and Björklund, 1995; Lajus and Alekseev, 2000; Lajus, 2001; Lajus et al., 2015b). This means that the real value of fluctuating asymmetry is usually significantly lower than the observed value. High measurement error also increases sampling error, since much observed variation in fluctuating asymmetry is associated with measurement noise and not biological signals.

It is also important to take account of other types of asymmetry, such as directional asymmetry and anti-symmetry, when assessing fluctuating asymmetry (Graham et al., 1998). Directional asymmetry occurs when a character on one side is consistently larger than on the other side (i.e. both the presence and direction of asymmetry are predetermined). Antisymmetry occurs when a structure is more pronounced on one side than on the other, but the side is not predetermined. These are adaptive types of symmetry. Structures with anti-symmetry and directional symmetry also possess fluctuating asymmetry.

The relationship between fluctuating asymmetry and trait size can assume various guises depending on the multiplicative and additive errors associated with structure growth (Graham et al., 1998). This needs to be accounted for in FA studies, and different techniques can be applied to address this type of error (Palmer and Strobeck, 2003).

Measuring developmental instability using fluctuating asymmetry is highly relevant to research on stickleback, which are widely used to explore questions in various fields of population biology and ecology (Wootton, 1984; Bell and Foster, 1994a; Östlund-Nilsson et al., 2007). Fluctuating asymmetry drew the interest of stickleback researchers soon after the technique appeared in other biological research (Moodie and Moodie, 1996; Reinchen, 1997). Yet today, fluctuating asymmetry is rarely used in stickleback research, despite its continued application in studies of other fish (Allenbach, 2011).

Here, we review the literature on stickleback FA studies, and develop a new FA analytical technique that employs the cranial bones as characters, using four populations of threespine stickleback that differ in their life history and geographical location.
METHODS AND MATERIALS

Literature review

For articles indexed in the Web of Science – All Databases, we searched for combinations of the keywords ‘stickleback’ and ‘fluctuating asymmetry’. In total, 41 articles were identified in the database, but on closer examination we narrowed this down to 28 original research studies describing the use of FA methods on stickleback (see Appendix). For each study we: (1) recorded the traits used in FA analysis; (2) characterized the populations studied by geographical location, life cycle (marine, anadromous, freshwater), and number of populations involved; (3) identified the variables correlated with fluctuating asymmetry; (4) described reported patterns of association between the studied variables and fluctuating asymmetry; and (5) analysed how measurement error was addressed, if at all. This method of data presentation allowed us to structure a variety of analytical procedures.

If more than one factor was analysed in a single paper, each factor was taken into account separately. Where patterns of correlation varied for multiple traits, only the mean value was used, i.e. if a positive correlation was found for one of three characters, with the other two showing no correlation, the overall correlation was considered to be positive.

Field sampling and bone preparation

For the methodological study, we used threespine stickleback populations representing different life cycles (resident freshwater, resident marine, and anadromous) and different geographical regions (the White Sea and the Kamchatka Peninsula). Altogether, four samples collected during the summer season of 2014 were used in the study (Table 1). Three of the samples were collected from the White Sea basin – one in freshwater Lake Goreloe, and two in geographically distant parts of the White Sea (Kandalaksha and Onega Bays). We also sampled an anadromous population in the Kamchatka River (Kamchatka

Table 1. Characteristics of the threespine stickleback populations sampled

<table>
<thead>
<tr>
<th>Life cycle</th>
<th>Population</th>
<th>Location</th>
<th>Date</th>
<th>No. males/females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshwater resident</td>
<td>Goreloe</td>
<td>Lake Goreloe, Malyy Gorelyy Island, Kandalaksha Bay, the White Sea Basin (66°17′49″N, 33°37′39″E)</td>
<td>28 June 2014</td>
<td>16/9</td>
</tr>
<tr>
<td>Anadromous</td>
<td>Kamchatka</td>
<td>Kamchatka River, Kamchatka Peninsula, the Bering Sea Basin (56°12′35″N, 161°59′1″E)</td>
<td>6–12 July 2014</td>
<td>16/8</td>
</tr>
<tr>
<td>Marine resident</td>
<td>Seldianaya</td>
<td>Seldianaya Inlet, Kandalaksha Bay, the White Sea (66°20′15″N, 33°37′26″E)</td>
<td>15 June 2014</td>
<td>15/15</td>
</tr>
<tr>
<td>Marine resident</td>
<td>Onega</td>
<td>Coastal waters of Onega Peninsula, the White Sea (64°56′47″N, 36°44′42″E)</td>
<td>9–12 July 2014</td>
<td>8/19</td>
</tr>
</tbody>
</table>
Peninsula, the Bering Sea). These samples reflect the variety of life cycles and geographical distribution of the threespine stickleback.

The White Sea is a semi-enclosed sea connected to the Barents Sea to the north, with reduced salinity (about 25 ppm at the surface) and a relatively continental climate. Kandalaksha Bay is situated in the northwest of the White Sea. It features a steep, embayed coastline with abundant inlets carpeted with macrophytes (brown algae and seagrass). In contrast, the Onega Peninsula is situated in the southern part of the White Sea between two bays: Onega and Dvina. Its coastline is smoother and the adjacent waters are shallower. Stickleback density on the Kandalaksha Bay spawning grounds is much higher than in the rest of the White Sea, including Onega Bay. The distribution of adult and juvenile stickleback in Kandalaksha Bay is primarily associated with the seagrass *Zostera marina* (Ivanova et al., 2016; Rybkina et al., 2017). Goreloe, a small humified forest lake with a surface area of about 2 hectares and average depth of 2.8 m, is situated on a small island in Kandalaksha Bay (Kuznetsova et al., 2007). At ~5 m above sea level, Goreloe Lake is not connected to the sea. The Kamchatka River is the largest river on the Kamchatka Peninsula. Stickleback spawn where it flows into the Bering Sea, and juvenile fish spend their first several weeks there before moving offshore (Bugaev et al., 2007).

Freshwater and marine stickleback were caught using a 7.5-m beach seine within 30 m of the shore, photographed and preserved in 70% ethanol. Anadromous stickleback were collected with a hoop-net in the Kamchatka River and preserved in 10% formaldehyde. The sex of each fish was determined via dissection. Total length was measured from photos of the fresh fish (White Sea samples) using ImageJ software (http://imagej.nih.gov/ij/index.html), or measured with a ruler after capture (Kamchatka).

To obtain cranial bones, fish heads were kept in 2% NaOH solution at a temperature of 50 °C for a few hours until the bones could be separated from the soft tissues. Then, the bones were cleaned in water with a brush, dried at room temperature, and stored in Eppendorf tubes.

**Morphological analyses**

We analysed one meristic and 41 morphometric characters. The meristic character – the number of lateral plates – was analysed under a stereo-microscope (MBS-10). For morphometric analyses, we chose four relatively large, flat, paired cranial bones: the operculum, lachrymal, third suborbital, and quadrate. They were scanned using an Epson Perfection 4490 Photo scanner at a resolution of 1200 dpi. We used the coordinates of 20 distinguishing features, numbered as landmarks, on the digital images (Fig. 1) using ImageJ software. For the morphometric characters, we used the distances between the landmarks on each bone, calculated in Microsoft Excel (e.g. the distance between landmarks 3 and 5 was given the label ‘character 3-5’). To estimate measurement error, each bone was scanned twice after a slight change of position. Thus, each character was measured four times, since both left and right structures were measured twice. Lateral plates were also measured twice on the left and right sides of the body. Measurements were not performed according to sample order, but instead randomly and blindly, so that the operators were unclear which images or samples they assessed. This minimized possible bias due to prior operator experience when reading sample results, which has been observed in some studies (Lajus et al., 2003b; Kozlov and Zvereva, 2015). Statistical analysis was performed using the Statistica v.10.0 software package.
The fluctuating asymmetry of each of the 41 individual morphometric characters was calculated using the index: $\text{FA} = \frac{|R - L|}{0.5(R + L)}$, where $R$ and $L$ are the values of the characters on the right and left side of the body, respectively (Palmer and Strobeck, 1986). To normalize $|R - L|$ distributions, we used the Box-Cox transformation (Box and Cox, 1964): $\left(|R - L| + \lambda_1\right)^{\lambda_2}$, with $\lambda_1 = 0.016$ and $\lambda_2$ varying from 0.2 to 0.6 for different characters ($\lambda_1$ and $\lambda_2$ values minimize departure from the normal distribution). For lateral plates, $\lambda_1 = 0.0159$ and $\lambda_2 = 0.4$.

The generalized FA index for an individual was based on the FA values of the 41 morphometric characters, according to Clarke and McKenzie (1992). We standardized all individual FA values by variance to equalize the contribution of different characters in the generalized index, then determined the mean FA value for all the characters of an individual fish.

---

Fig. 1. Bones and landmarks used for analysis of fluctuating asymmetry in threespine stickleback.
Before FA analysis, we performed a character selection procedure to avoid replicating FA information from structurally related characters, such as characters 11-14 and 11-15. After the first visual selection, we calculated the Pearson correlation coefficients between the absolute FA values of all possible character pairs in a pooled sample, and removed pairs with significant correlations ($df = 105; P < 0.05$). Therefore, only characters with non-correlated FA values were used in the generalized index.

The significance of directional asymmetry (systematic differences between values on the left and right sides) was estimated using mixed-model analysis of variance (ANOVA), with ‘body side’ (left or right) as a fixed factor and ‘individual’ as a random factor, with subsequent Bonferroni correction (Palmer and Strobeck, 2003). Only characters without significant directional asymmetry were used in the calculations.

The characters used in the final FA analysis were as follows: 1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5, 6-7, 6-8, 6-10, 6-11, 7-8, 7-10, 7-11, 8-10, 8-11, 10-11, 12-14, 12-15, 12-16, 12-17, 12-18, 14-15, 14-16, 14-17, 14-18, 15-16, 15-17, 15-18, 16-17, 16-18, 17-18, and 19-20 (see Fig. 1 for numbered landmarks).

Measurement error was estimated using one-way ANOVA of the first and second replicate FA measurements, with ‘individual’ as a factor (Palmer and Strobeck, 2003). This allowed us to estimate the contribution of measurement error to the observed fluctuating asymmetry.

RESULTS

Publications on fluctuating asymmetry in sticklebacks

G.E.E. Moodie published the first paper examining fluctuating asymmetry in sticklebacks in 1977, and the latest paper in our survey appeared in 2014 (Fig. 2; Appendix). During the 1990s and early 2000s, the number of publications on fluctuating asymmetry in sticklebacks rose, as did publications featuring the keywords ‘fluctuating asymmetry’ and ‘stickleback’. Subsequently, publications on fluctuating asymmetry in stickleback declined, whereas publications on fluctuating asymmetry in general remained stable, and stickleback publications rapidly increased (Fig. 2).

Twenty-four of 28 publications focused on threespine stickleback, three on brook stickleback (*Culaaea inconstans*) and one on ninespine stickleback (*Pungitius pungitius*). Three studies (Reimchen, 1997; Reimchen and Nosil, 2001a, 2001b) did not separate fluctuating and directional asymmetry, but did focus in the main on directional asymmetry (DA). Fluctuating asymmetry creates additional noise in DA studies, which increases minor sampling error. However, if fluctuating asymmetry is the main interest, the presence of directional asymmetry will lead to magnitude-level changes, thus directional asymmetry must be removed for accurate FA analysis (Graham et al., 1998; Palmer and Strobeck, 2003). Therefore, we removed the three DA-focused studies from further analyses.

Traits

Two types of traits were used in stickleback FA studies – meristic or counted traits, and morphometric or measured traits. Meristic traits included the number of fin rays, lateral plates, gill rakers on the first arch, forks on the ascending branch of the pelvic girdle, and neuromasts/neuromast pores in lateral line structures. The most commonly used traits were the number of lateral plates, and of rays on the pectoral fin. Morphometric characters included the length of pelvic spines and pectoral fins, the width of pectoral fins, the position
and length of lateral plates, the diameter of the eye, the height and width of the ascending process, the width at the base of the pectoral fins, the pelvic–dorsal distance, the length of the anterior process, ascending branch, and pelvic spines, and the width and mass of gill covers.

The number of traits used in analyses was quite limited in most cases: 11 publications used just one trait, four studies used two traits, and another ten used 3–15 traits. Often, variation in structures pertaining to external defence was analysed: 18 studies analysed lateral plates (their number and position), while a further ten examined the length of pelvic spines. Pectoral fins (their length and number of fin rays) were used in nine studies. In total, about 35 different characters were used in the various studies (see Appendix).

**Measurement error**

Measurement error (ME) was addressed in 21 studies. In two studies (Moodie, 1977; Aguirre et al., 2004), no ME analysis was performed, while in two further studies (Mazzi et al., 2003, 2004), the authors looked at the association between female preference and male fluctuating asymmetry using computer images of males, and hence ME analysis was not appropriate. In six of the 21 studies that did address measurement error (Moodie and Moodie, 1996; Bergstrom and Reimchen, 2005; Moodie et al., 2007; Reimchen et al., 2008; Reimchen and Bergstrom, 2009; Trokovic et al., 2012), the authors assumed that measurement error was negligible based on previous studies. Lescak et al. (2013) did not estimate the measurement error, but implicitly accounted for it by scoring samples in random order to maximize measurement consistency and to minimize the potential for shifting observer bias, thereby avoiding temporal trends in measurement error. In all other studies, replicate measurements were undertaken, although one study did not present the results of this procedure (Hermida et al., 2002). Two studies (Mazzi and Bakker, 2001;
Mazzi et al., 2002) undertook replicate measurements and used the mean value to halve the variance caused by measurement error. Three studies examined whether measurement error confounded fluctuating asymmetry (Hechter et al., 2000; Prieto et al., 2005; Kenney and von Hippel, 2014) but found no such effect. Five studies provided data on the overall ratio of measurement error in observed fluctuating asymmetry (Bergstrom and Reimchen, 2000; Bakker et al., 2006; van Dongen et al., 2009; Loehr et al., 2012, 2013).

### Directional asymmetry

Directional asymmetry was often reported in the stickleback studies. Although we assess only papers focusing on fluctuating asymmetry, we note that directional asymmetry does inflate fluctuating asymmetry. Directional asymmetry was observed most often in pelvic spines and lateral plates (left-biased directional asymmetry in both cases). In their detailed study of the asymmetry patterns of lateral plates, Reimchen and Bergstrom (2009) found that either type of asymmetry (directional or fluctuating asymmetry) may vary from plate to plate and depend upon their position. At the same time, some studies reported no directional asymmetry in plates or pelvic spines (Bergstrom and Reimchen, 2003; Bakker et al., 2006), though directional asymmetry was observed in other structures (van Dongen et al., 2009; Lescak et al., 2013; Kenney and von Hippel, 2014).

### Concordance in fluctuating asymmetry among characters

Fourteen studies used more than one trait, so that the concordance of fluctuating asymmetry across characters can be a potential concern. Of these studies, Trokovic et al. (2012) used a composite FA index and did not address concordance. Eight other studies explicitly analysed concordance, and three (Bergstrom and Reimchen, 2002; Robinson and Wardrop, 2002; Prieto et al., 2005) reported significant concordance in the variation in fluctuating asymmetry across characters. One of these three studies (Robinson and Wardrop, 2002) reported significant concordance in fluctuating asymmetry among individuals within a population, and three others found concordance among populations. Three studies reported a lack of concordance in the fluctuating asymmetry of traits across samples, and four others reported similar results across individuals within a population.

### Relationship between fluctuating asymmetry and various environmental and fitness characteristics

The relationship between fluctuating asymmetry and abiotic variables was the focus of several studies. Three reported geographical heterogeneity in fluctuating asymmetry, but did not link fluctuating asymmetry to a particular environmental factor (Prieto et al., 2005; van Dongen et al., 2009; Loehr et al., 2013). Fluctuating asymmetry was found to decrease with increased water clarity (the authors argued that this association was mediated by higher pressure from predators in clearer waters), increased altitude and pH (Bergstrom and Reimchen, 2002, 2003). Depending on the trait, the fluctuating asymmetry was either higher or lower in stream than in lake dwellers (Moodie, 1977). Two studies found no association between fluctuating asymmetry and acidification or chemical water pollution (Mazzi and Bakker, 2001; Kenney and von Hippel, 2014). Fluctuating asymmetry was found to be lower in marine than in freshwater habitats, a difference probably due to local adaptation rather than to direct responses to different environments (Trokovic et al., 2012).

Of the biotic factors, studies focused on parasites and predators. A positive association between fluctuating asymmetry and parasite infection would confirm expectations that
individuals with lower fitness have lower developmental stability and less resistance to parasites. Such an association was found in two cases (Bergstrom and Reimchen, 2002; Reimchen and Bergstrom, 2005), but only for population-level comparisons (Bergstrom and Reimchen, 2002), and for only one of three studied parasites, *Eustrongylides* sp. (Reimchen and Bergstrom, 2005). Prieto and colleagues (2005) reported a negative association between fluctuating asymmetry and parasite loading by *Eustrongylides* sp. in one of four studied populations and for one of six studied characters (no data on correction for multiple comparisons was provided).

Some research looked at whether stickleback with lower fitness, and therefore higher fluctuating asymmetry, are more vulnerable to predators. Although it is difficult to test this hypothesis directly, three studies used it to interpret empirical results (Moodie, 1977; Bergstrom and Reimchen, 2003, 2005), two of which concluded that fluctuating asymmetry is lower in the presence of predators.

The correlation of fluctuating asymmetry with various fitness proxies is expected to be negative. Contrary to expectations, however, Moodie and Moodie (1996) found a positive correlation between fluctuating asymmetry and characteristics of male reproductive success (number of fry and eggs in nests, and proportion of empty nests). In another study with a similar design, Bakker et al. (2006) found this correlation to be negative. Turning to females, Hechter et al. (2000) reported that fluctuating asymmetry is negatively correlated with characteristics of female reproduction success (fecundity and weight of ovaries). Mazzi and colleagues (2002) found a negative correlation between fluctuating asymmetry (one of three characters) and levels of heterozygosity. Robinson and Wardrop (2002) observed fluctuating asymmetry to increase in fish at higher water temperatures and with richer diets, resulting in faster growth. Although this contradicts theory, the authors explained the incongruity via confounding environmental effects. Also contrary to theory, Moodie et al. (2007) noted a positive correlation between fluctuating asymmetry and growth rate in large males, but not in small males or females. Mazzi et al. (2003, 2004) showed that inbred females preferred more symmetrical males, which confirmed theory, but outbred females did not. Van Dongen et al. (2009) tested hypotheses of whether directional selection and evolutionary changes increase average levels of fluctuating asymmetry, and whether they also increase the strength of association between fluctuating asymmetry and population-level genetic variation. Their results supported the second hypothesis, but not the first. Lescak et al. (2013) reported an increase in fluctuating asymmetry with increased pelvic expression but did not interpret these results in terms of fluctuating asymmetry–fitness relationships. Analysis of FA heritability showed that it was significant in one case (Loehr et al., 2012) but non-significant in two others (Hermida et al., 2002; Aguirre et al., 2004).

In summary, 11 studies confirmed expected correlations (i.e. positive with stress indices and negative with fitness proxies), three contradicted expectations, and in nine cases no correlation was observed.

**Field analysis of fluctuating asymmetry in lateral plates and cranial bones**

Stickleback mean length was $44.5 \pm 0.89$, $64.4 \pm 0.61$, $64.6 \pm 1.06$, and $83 \pm 0.79$ mm for the freshwater, two marine (Onega Bay and Seldianaya Inlet in Kandalaksha Bay), and anadromous populations, respectively. Paired comparisons showed significant differences between all pairs of samples ($t$-test, $P < 0.05$), except between the two marine populations.

Mean measurement error in fluctuating asymmetry for all characters in all samples was 0.327, with the lowest measurement error for character 3-5 (0.150) and the highest for
character 19-20 (0.668). We tested for directional asymmetry using the paired Student’s \( t \)-test. One character (2-3) showed statistically significant directional asymmetry (\( df = 104, P < 0.05 \)), which became non-significant after sequential Bonferroni correction (Rice, 1989).

For cranial bones, fluctuating asymmetry was at its the highest in the resident freshwater and two marine populations; for lateral plates, it was highest in the freshwater population. The anadromous population displayed relatively low fluctuating asymmetry for both cranial bones and lateral plates (Fig. 3). Significant differences (\( t \)-test, \( P < 0.05 \)) in the fluctuating asymmetry of cranial characters were found in all comparisons of anadromous fish with the other populations, as well in all comparisons of freshwater fish with the other populations, whereas the two populations of marine stickleback did not differ from one another. No differences were observed in fluctuating asymmetry for cranial bones or lateral plates between males and females in all samples.
Number of lateral plates divided the populations into two groups: (1) anadromous (32.94 ± 0.166 plates) and marine (Seldianaya) (32.88 ± 0.124) stickleback with a relatively high number of plates, and (2) freshwater (Goreloe) (30.76 ± 0.499) and marine (Onega) (31.05 ± 0.203) stickleback with significantly fewer plates. Pairwise comparisons in all cases revealed significant (t-test, \( P < 0.05 \)) differences between groups (1) and (2), but no differences between populations from the same group. In the freshwater population, almost half of the fish we examined (11 out of 25) had missing plates on the middle parts of their bodies, something that was not observed in the other populations.

**DISCUSSION**

**Analysis of publications**

Although the first article on fluctuating asymmetry in stickleback appeared in 1977, real interest only began to be shown by stickleback researchers some 20 years later, after FA studies of other organisms had become common (Fig. 2). However, interest waned in the mid-2000s, despite continued research on fluctuating asymmetry in other species and an exponential growth in the number of publications on stickleback (Fig. 2). Evidently, FA research results did not meet stickleback researchers’ expectations, so we analysed their publications to determine why.

In many cases, the results of FA analyses did not show the expected associations with stress or fitness. And many more negative results likely remained unpublished because of the difficulties associated with publishing negative results in high ranking journals (Kicinski, 2013; Duyx et al., 2017). This probably reduced the incentive for continuing research on fluctuating asymmetry in stickleback. Scepticism about fluctuating asymmetry as an indicator of stress and fitness became widespread in the early 2000s – not only in stickleback fishes, but in other organisms as well (see references in Introduction). However, publications since that time have remained steady overall (Fig. 2), which suggests that interest in the approach has persisted, but that standards for publishing FA results have become more stringent, requiring tests for measurement error, other types of asymmetry, size dependence, and so on. Conducting these tests reduces noise and yields FA data that can support detailed interpretation (Palmer and Strobeck, 1986, 2003), but this is rarely done in stickleback FA studies. Below we consider different issues affecting FA analyses.

**Measurement error**

Although the importance of measurement error has been recognized for a long time in FA research (Palmer and Strobeck, 1986), it still receives relatively little attention. According to Allenbach (2011), measurement error was assessed in only 30 out of 81 studies of naturally occurring fluctuating asymmetry in fish caused by exogenous stress. This is important because high measurement error can influence the results of sample comparisons. As measurement error can differ in different samples, when measured by different operators, or even by the same operators at different times (see, for instance, Lajus et al., 2003b), it may confound real differences between samples or suggest spurious differences. There are two ways to address this issue. The first is to analyse specimens from different samples in random order, which equalizes high measurement error across samples and does not affect the results of sample comparison. This is effective when the magnitude of measurement error is affected not only by heterogeneous analytical conditions, but also by the specific patterns of morphological
structures. It is always important to estimate the contribution of measurement error to the observed fluctuating asymmetry. The second way to address this issue is to assess measurement error in each sample separately, expressed as a proportion of fluctuating asymmetry, and to use true fluctuating asymmetry (i.e. fluctuating asymmetry minus measurement error) when comparing samples. This is the only rigorous way to compare samples that differ intrinsically in morphological structures, such as the muscle prints on the internal surface of mollusc shells (Lajus et al., 2015b). Not accounting for measurement error in stickleback FA studies likely caused noise in the results, and introduced ambiguity when interpreting them.

Concordance in fluctuating asymmetry among characters

Often when comparing samples in FA studies, concordance in fluctuating asymmetry of different characters is implicitly assumed as a precondition for considering developmental instability as a genome-wide effect (Dufour and Weatherhead, 1996; Clarke, 1998; Zakharov, 2003; but see Leung and Forbes, 1996 and Bjorksten et al., 2000a, 2000b for a contrary opinion). If this is not the case, developmental instability is merely a property of individual traits. There are conflicting empirical results for concordance in fluctuating asymmetry across populations, however. Möller and Swaddle (1997) reported finding concordance in 12 of 17 papers. Clarke (1998) found concordance in four out of his 11 data sets. Quantitative estimates of concordance, as measured by the Kendall coefficient of concordance, are usually low (Leamy, 1992; Leary et al., 1992; Kozhara, 1994; Auffray et al., 1999). This is because different characters may respond differently to stress (Hoffmann and Woods, 2003; Freeman et al., 2005). In particular, functionally important characters display lower fluctuating asymmetry and may be less suitable for FA studies (Palmer and Strobeck, 1986; Leung and Forbes, 1996). Different patterns of fluctuating asymmetry may be explained by the different development of structures. For instance, meristic characters such as number of fin rays are determined early in ontogenesis and do not change later, while morphometric characters change throughout the life of a fish. Departure from concordance also results from different patterns of association between the size of characters and the magnitude of fluctuating asymmetry in different populations (Lajus, 2001). Thus, it is intuitively clear that departure from concordance can confound correlations between fluctuating asymmetry and both stress and fitness; however, we lack a simple means of accounting for it in practical research.

In eight of the 24 studies that addressed the concordance in fluctuating asymmetry of different traits when comparing samples, four reported concordance while the other four reported departure from concordance. Such ambiguity can exacerbate the heterogeneity of results, thus making their interpretation more difficult. Concordance can also be a problem in studies where only one trait, or a composite FA index based on several traits, is used (almost half of the total number). Although addressed explicitly, concordance can still contribute to widely disparate findings.

Trait selection

Trait selection predetermines the success of FA studies because it materially affects the chances of distinguishing differences among several samples. Thus, trait selection should be optimized before testing the association of fluctuating asymmetry with stress or fitness (Lajus et al., 2015a). Unfortunately, such optimization rarely occurs. An ideal trait for FA analysis must be easily measured or counted without unnecessary preparation and manifest ideal fluctuating asymmetry, i.e. where normally distributed $R - L$ values approach zero, are not correlated with size, display negligible measurement error, and are not subject to
unaccounted environmental or genetic effects that complicate interpretation of the results. The majority of stickleback studies have used lateral plates and pelvic fins specific to the species, thus not making comparisons with other fish possible. These features can evolve quickly in populations responding to various selection regimes (van Dongen et al., 2009). For this reason, fluctuating asymmetry of these characters can change not only due to one of the traits being considered, such as environmental stress or fitness characteristics, but also as a result of selection pressure on that particular trait.

This can confound the interpretation of stickleback FA studies and introduce inconsistency in the analysis of different traits. For instance, Bergstrom and Reimchen (2005) reported that the relationship between the fluctuating asymmetry of the number of lateral plates of three-spine stickleback and their parasite load is stronger in habitats where the plates’ functional importance and selection against their asymmetry was reduced. The authors suggested that the inconsistent connection between fluctuating asymmetry and parasitism may be due to ‘undetected biomechanical selection against asymmetry’. In general, stickleback researchers have usually studied the fluctuating asymmetry of traits that are also widely used to study phenotypic plasticity – and may not produce the best results. Since criticism of fluctuating asymmetry in the early 2000s, research on other species became more discriminating in selecting structures for analysis; however, stickleback offer a limited choice of structures and this may have stifled research. In addition, lateral plates and pelvic spines may manifest directional asymmetry, which is of interest in a number of stickleback studies, but may inflate fluctuating asymmetry if not addressed properly. While some studies have considered directional asymmetry as an indicator of environmental stress, most presented directional asymmetry as part of normal development and inappropriate for indicating environmental stress (Graham et al., 1993, 1998; Marques et al., 2005).

**Ameliorations**

Although the above issues should be taken into account in FA research – and failure to do so has likely obscured the expected correlations between fluctuating asymmetry and stress and fitness (for reviews papers, see Polak, 2003) – there is no simple way to address them. Nevertheless, we suggest that using bony structures, such as cranial bones, may increase considerably the power and reliability of FA analysis. Few analyses of the fluctuating asymmetry of internal fish bones have been conducted (Romanov, 1983; Kozhara, 1989; Masson et al., 2011). Bone shape depends on developmental conditions (Ornsrud et al., 2004; Georgakopoulou et al., 2007; Yurtseva et al., 2010; Berg et al., 2012) or on genetics (Sutterlin et al., 1987; Benfey, 2001; Sadler et al., 2001). While they require some preparation before analysis (dissolving of soft tissues, cleaning), bones provide a large number of characters that are suitable for digital analysis. In our studies of several fish species, we have used 24–48 characters per species (Lajus, 1991, 2001; Lajus et al., 2003c, 2014, 2015a; Yurtseva et al., 2010, 2014). Measurement error was variable, but was usually below 20% in morphometric characters (Lajus, 2001; Lajus et al., 2015a). An important advantage of using cranial bones is the large number of convenient characters, which reduces sampling error as well as the number of fish required to obtain statistically significant results. Nevertheless, all of our FA studies using multiple bone traits displayed statistically significant differences among samples (Lajus, 2001; Lajus et al., 2003c, 2014, 2015a; Yurtseva et al., 2010, 2014), some of which could not be associated with genetic or environmental variables and, therefore, severely limited their interpretation.

Recently, tomographic techniques have become more widely used in morphological studies. These techniques do not require labour-intensive preparation of individual bones and allow for collocation of different bones in relation to one another (Lauridsen et al., 2011;
Hilton et al., 2015). Tomography, therefore, may have great potential in studies of fluctuating asymmetry.

**Field studies of fluctuating asymmetry in stickleback**

We obtained conflicting results when comparing different characters in our four populations of stickleback. For cranial bones, fluctuating asymmetry was most marked in the resident freshwater and the two marine populations, whereas for lateral plates it was highest in the freshwater population. The anadromous population showed relatively low fluctuating asymmetry in both cranial bones and lateral plates.

How should we interpret the observed differences? The number of lateral plates in three-spine stickleback depends on their life history. Marine and anadromous populations usually have more than 30 lateral plates, whereas freshwater populations usually have up to 14 (Hagen, 1967; Zugerov, 1983; Wootton, 1984; Bell and Foster, 1994b; Reimchen, 1994). In our case, plates in the freshwater population were similar in number to marine forms, despite the fact that this population has probably inhabited Goreloe Lake for a thousand years, given that it lies 5 m above sea level and the rate of post-glacial rebound in this area is about 5 mm per year (Kolka et al., 2005; Romanenko and Shilova, 2012). This suggests that the Goreloe stickleback population has maintained its full number of lateral plates, as has been observed in other freshwater populations (Mori, 1990). However, we observed fewer lateral plates on some of the lake fish, whereas the marine and anadromous populations presented full sets of plates. This may mark the beginning of the process of lateral plate reduction, which may increase fluctuating asymmetry in the number of lateral plates, and thus requires careful interpretation when relating differences in fluctuating asymmetry to population fitness.

Interpreting fluctuating asymmetry in cranial bones does not present such problems. As a large number of characters are involved in FA analysis, the somewhat high sampling error that can occur in FA studies is avoided (Fuller and Houle, 2003; van Dongen, 2007). Another advantage of using these structures is that they lend themselves to analysis via computerized techniques, which are more precise than manual methods and less influenced by factors that increase measurement error (Muñoz-Muñoz and Perpiñán, 2010; Lajus et al., 2015b). Moreover, bones growing throughout an animal’s life should reflect the totality of its environmental conditions, whereas the number of lateral plates is fixed early in ontogenesis. For this reason, we advocate the use of cranial structures when interpreting the results of FA analyses, which likely indicate a great deal more about the overall developmental stability of a population.

Why did an anadromous stickleback population from the Pacific Ocean have higher fitness than a freshwater population from a lake near the White Sea, or marine stickleback in the White Sea itself? First, fluctuating asymmetry is inversely correlated with the size of adult fish – the smaller the fish, the more asymmetrical they are. Considering growth as another fitness surrogate (Arnold, 1983; Palmer, 1983) confirms the differences we observed among the four populations.

Anadromous stickleback on the Kamchatka Peninsula spend the first weeks of their lives in a river, before migrating to the sea (Bugaev, 1992). Their growth – and hence their bone formation – takes place mostly at sea (Bugaev, 1992). The habitats and migration patterns of North Pacific stickleback are not well known, but the White Sea is covered with ice for 8–10 months in the north and is completely ice-free in the south, and likely offers a wide variety of environmental conditions for migrating fish. The North Pacific Ocean supports the largest fisheries worldwide (Garcia and Newton, 1994) due to its high biological productivity compared...
with other regions of the world’s oceans (Grebmeier et al., 2006). This suggests that there are favourable conditions for fish there in different seasons. Another factor in organism fitness is quality of food items. According to data obtained in spring at Avacha Bay on the Kamchatka Peninsula, planktonic crustaceans form the basis of the marine zooplankton community, particularly copepods (Batishcheva, 2008), which have high nutrition value compared with freshwater invertebrates (Watanabe et al., 1981, 1983; cited by Novoselova, 2012). They likely provide a rich food source for stickleback.

In inshore areas of the White Sea, the diet of juvenile threespine stickleback consists to a large extent of copepods, mostly Temora longicornis and Microsetella norvegica (Demchuk et al., 2015; Rybkina et al., 2016). In the deeper marine waters, the copepods Calanus glacialis, Metridia longa, and Pseudocalanus minutus predominate (Kosobokova and Pertsova, 2005), and these high-quality food items are likely consumed by stickleback. At the same time, subarctic winter conditions on the inland White Sea – the Kandalaksha Gulf is covered with ice for six months – are probably more severe than in the Bering Sea. Because the White Sea is located near the northern border of the threespine stickleback distribution, habitat conditions may be suboptimal and stickleback populations will be sensitive to environmental changes. Significant fluctuations in stickleback abundance in the White Sea bolster this supposition (Lajus et al., 2013). We know little about any differences between the two White Sea marine populations in this study. It has been found, however, that the gonads of Onega Bay stickleback have higher levels of essential polyunsaturated (docosahexaenoic and eicosapentaenoic) and monounsaturated (palmitoleic) fatty acids, which may reflect their more favourable trophic status (Murzina et al., 2018). These differences, however, are not reflected in fluctuating asymmetry.

Environmental conditions are likely even more severe in Goreloe Lake. Food organisms available for stickleback in the lake are quite different from those in the White Sea. The rotifers Bipalpus hudsoni and Asplanchna priodonta predominate in number, and the cladoceran Polyphemus pediculus in biomass (Kuznetsova et al., 2007). Stickleback can also feed on benthic organisms. Comparing food organisms available to fish in different environments suggests more favourable feeding conditions for marine and anadromous stickleback than for the landlocked freshwater population. Additional stress in the lake can be caused by the tapeworm Schistocephalus solidus, which uses stickleback as an intermediate host and is common in lakes and reservoirs with low water flow (Barber and Scharsack, 2010). The larvae of this parasite can be very efficient energy consumers, requiring four to five times more energy to grow compared with their host fish (Walkey and Meakins, 1970). In previous studies of the threespine stickleback (Reimchen and Nosil, 2001a, 2001b; Bergstrom and Reimchen, 2002), as well as other species such as Atlantic salmon, Salmo salar (Yurtseva et al., 2010), higher fluctuating asymmetry and lower developmental stability were found in populations infected with parasites. Studies of parasites in marine stickleback (Shulman and Shulman-Albova, 1953; Rybkina et al., 2016) did not reveal species as debilitating as S. solidus. Like the White Sea, Goreloe Lake is completely covered with ice in winter, and conditions there can be even more severe than in the White Sea.

Although mortality among stickleback males is higher than among females in the White Sea, which is likely associated with a female-biased sex ratio (Golovin et al., 2019), no differences in fluctuating asymmetry between the sexes was observed. This suggests that the higher male mortality is due to differences in the life history of the sexes, not their fitness.

In summary, the results of our analysis of fluctuating asymmetry in four populations of stickleback that differ in life cycle and habitat, can be meaningfully interpreted in terms of species biology.
CONCLUSION
Interest in fluctuating asymmetry in stickleback, expressed as the number of publications with the keywords ‘fluctuating asymmetry’ and ‘stickleback’, showed growth in the late 1990s, a plateau in the early 2000s, and a falling away of interest thereafter. At the same time, the number of publications with the keywords ‘fluctuating asymmetry’ also grew quickly but remained constant as the 2000s progressed, while publications with the keyword ‘stickleback’ grew rapidly throughout. Critical analysis of these publications suggests that the decline in interest in stickleback fluctuating asymmetry was due to highly disparate research results and difficulties interpreting the relationship between fluctuating asymmetry and stress and fitness. To some extent, these difficulties may have been caused by the limited number of morphological structures used in FA analysis in stickleback, particularly lateral plates and pelvic fins which can quickly evolve due to various environmental changes. Thus, changes in fluctuating asymmetry in these structures can change not only due to stress and fitness, but also due to multiple uncontrolled factors that may specifically affect these very same structures.

In this study, we used cranial bones, which had not been used for FA analysis in stickleback before, but have shown promise in other fish species. Our analyses showed less fluctuating asymmetry in cranial bones in anadromous stickleback (Kamchatka Peninsula) than marine and freshwater stickleback from the White Sea. This may be because of more favourable feeding conditions in the North Pacific than in the White Sea locations. At the same time, we found similarities in fluctuating asymmetry among the White Sea populations. The different environmental conditions in the lake and at sea were not reflected in marked differences in fluctuating asymmetry, although comparison with the anadromous fish from the North Pacific shows that the fluctuating asymmetry of the structures we used is sensitive to such conditions. Therefore, our results show that FA analysis using cranial structures reveals patterns in stickleback heterogeneity that can be interpreted in terms of different life histories. Thus, cranial bones can be considered suitable structures for FA studies of stickleback.

ACKNOWLEDGEMENTS
We are grateful to A. Kucheryavyy for providing material from the Kamchatka River, V. Spiridonov for organizing an expedition to the ‘Onezhskoe Pomor’e’ National Park, Arkhangelsk region, and the administration of the ‘Belomorskaya’ Marine Biological Station of Saint-Petersburg State University (SPbSU). Karen Alexander helped improve our English. Financial support was provided by the Ministry of Science and Higher Education of the Russian Federation, project AAAA-A19-119020790033-9, to A. Yurtseva, and SPbSU grant 1.40.529.2017 to the authors from SPbSU.

REFERENCES


Fluctuating asymmetry as an indicator of stress and fitness in stickleback


Michaelsen, S., Schaefer, J. and Peterson, M.S. 2015. Fluctuating asymmetry in Menidia beryllina before and after the 2010 Deepwater Horizon oil spill. PLoS One, 10: e0118742. Available at: https://doi.org/10.1371/journal.pone.0118742.


