Aging three-spined sticklebacks (*Gasterosteus aculeatus*): comparison of estimates from three structures

Anastasia Yurtseva* [1,2], Kristina Noreikiene[3,4], Dmitry Lajus[2], Zitong Li [3], Tarja Alapassi[5], Tatiana Ivanova[2], Mikhail Ivanov[2], Pavel Golovin[2], Sami Vesala[5], Juha Merilä[3]

[1] Zoological Institute of Russian Academy of Sciences, St. Petersburg, Russia

[2] Saint-Petersburg State University, St. Petersburg, Russia

[3] Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Program, Faculty Biological and Environmental Sciences, University of Helsinki, Finland

[4] Institute of Veterinary Medicine, Estonian University of Life Sciences, Tartu, Estonia

[5] Natural Resources Institute, Finland

* **Correspondence**. Anastasia Yurtseva, Laboratory of Ichthyology, Zoological Institute of Russian Academy of Sciences. Universitetskaya nab., 1, 199034 St. Petersburg, Russia. Tel. +7 812 3280612, fax +7 812 3282941. Email: ayurtseva@gmail.com.

Abstract

Studies of fish life-history evolution and population demography require knowledge of populations' age structure. However, reliable aging of wild-collected individuals poses practical and logistic challenges associated with aging protocols. In order to assess the accuracy and reliability of age estimates from calcified structures in the three-spined stickleback (*Gasterosteus aculeatus*), we evaluated intra- and inter-reader repeatability from three structures (*viz*. otoliths, gill covers and pelvic spines). Average age estimates were also compared between the structures. The overall intra-reader repeatabilities of age estimates were highest for otoliths (69%), lowest for gill covers (53%), and intermediate for spine cross-sections (63%). Although four of the seven readers had the highest intra-reader repeatability score for spine cross-sections, the inter-reader variance in this structure was much higher than in others. Otoliths were the easiest in terms of their pre-analysis treatment and exchange of materials (as digital images) between readers. In addition, otoliths are more well-studied compared to the other structures with respect to their development through ontogenesis; hence, age estimates based on otoliths should be the most reliable. Therefore, our recommendation is that whenever possible, analysis of otoliths should be the preferred approach for aging three-spined sticklebacks.

Keywords: age, *Gasterosteus aculeatus*, gill cover, otolith, pelvic spine, repeatabilities

Significance Statement

The three-spined stickleback is a popular model in ecology and evolutionary biology. The age of wild-caught sticklebacks is often crucial for understanding their biology, but since different aging techniques are widely used, substantial heterogeneity in age estimates is possible. We compared intra- and inter-reader repeatabilities of age estimates obtained from three calcified structures (otoliths, gill covers, and pelvic spines). Otoliths provided the most consistent results and should therefore be the preferred structure for aging three-spined sticklebacks.

1. INTRODUCTION

Age analysis is a highly important component of fish population studies (e.g. Chugunova, 1959; Pravdin, 1966; Schneider *et al.,* 2000; Campana & Thorrold, 2001). Aging of fish, as well as other ectothermic vertebrates, is commonly done by analyzing annual or daily growth marks on calcified structures such as bones and otoliths. These analyses are based on the fact that distinct zones, called annuli, are formed during periods of slow growth (Craig, 1985; Wright & Huntingford, 1993; Secor *et al.,* 1995; Campana & Thorrold, 2001). While not without caveats and difficulties, including processing time and reliability, analysis of annual or daily growth increments on calcified structures still remains a powerful tool for providing age estimates (Campana, 2001).

The three-spined stickleback (*Gasterosteus aculeatus*) is a widely distributed small teleost that has become a model species in ecological, evolutionary and genetic studies (Schluter, 2000; McKinnon & Rundle, 2002; Bell *et al.,* 2004; Gibson, 2005; Cresko *et al.,* 2007; Barber & Nettleship, 2010). For this reason, the assessment of age structure in wild populations is especially important, and could provide advances for many fields (e.g. Baker *et al.,* 2008; DeFaveri & Merilä, 2013). However, age estimation in three-spined sticklebacks is challenging, as a universal method for aging has yet to be established. The absence of scales, which are among the most widely used structures for age estimation in fish (Campana & Thorrold, 2001), has motivated researchers to look for other structures that would be suitable for aging sticklebacks.

The most popular approach for stickleback age estimation is based on the analysis of their sagittal otoliths: the method was developed in the early 1950's (Jones & Hynes, 1950) and is still used today (Pichugin *et al.,* 2008; Herczeg *et al.,* 2009; Moser *et al.,* 2012; von Hippel *et al.,* 2013; Bergström *et al.,* 2015; Golovin *et al.,* 2015). However, otoliths can easily dissolve in non-buffered alcohol- or formaldehyde-preserved material, and hence, cannot be used for aging purposes in many collections.

To a lesser extent, other structures such as pelvic or dorsal spines (Reimchen, 1992; Gambling & Reimchen, 2012; DeFaveri *et al.*, 2014), and gill covers (Mukhomediarov, 1966; Patimar *et al.*, 2010; Yershov & Sukhotin, 2015) have also been used for aging sticklebacks. Worryingly, studies conducted in other fish species have shown that there may be large discrepancies in age estimates depending on which structure was used for the analysis (Zymonas & McMahon, 2009; Hüssy *et al.*, 2012; van der Meulen *et al.*, 2013; Sotola *et al.*, 2014; Baudouin *et al.*, 2015; Elzey *et al.*, 2015; Khan *et al.*, 2015; Watkins *et al.*, 2015; Zhu *et al.*, 2015). Two recent stickleback studies have analyzed several structures (DeFaveri & Merilä, 2013; DeFaveri *et al.*, 2014), but no study has formally compared the age estimates from different structures and their reliability in sticklebacks. Thus, the relative accuracy of stickleback age estimation using the aforementioned structures remains uncertain. Such comparisons would be of great practical relevance, especially since otoliths can be damaged during preservation. Furthermore, otolith and spine cross-section-based age estimation is time consuming, whereas analyses of gill covers are much faster to conduct. Hence, if gill cover analysis is deemed to be a reliable aging technique in sticklebacks, considerable amounts of time can be saved.

The goal of this study was to assess and compare the reliability of three-spined stickleback age estimates based on otoliths, gill covers and pelvic spine cross-sections. This was done by comparing repeatabilities of age estimates within and between readers for all three types of aging structures. Given the large number of readers (n = 8) and several (two to three) repeated measures per reader per structure, the results should provide information about the magnitude and sources of variation in stickleback age estimates. We also discuss and reflect upon practical matters related to obtaining reliable age estimates in sticklebacks.

2. MATERIALS AND METHODS

2.1. SAMPLE COLLECTION AND PREPARATION

Three-spined sticklebacks were collected from two locations in the White Sea using a beach seine (7.5 ×1.5 m, 5 mm mesh size in wings and 1 mm in the cod end). Forty-eight adults (6 males, 42 females; total length [± SD]: 70.8 ± 9.4 mm) of unknown age were caught on June 30th 2015 from Seldianaya Inlet (66.3378°N, 33.6234°E), and used as the focal individuals for the age estimation protocols. Fifteen one-year-old juvenile fish (40.9 ± 3.6 mm) were caught in the Koliushkovaya Lagoon (66.3137°N, 33.6440°E) on June 23rd 2015, and were used as references when aging the focal individuals. Their age was inferred to be one year based on the absence of breeding coloration and their small size. Other reference individuals were three sticklebacks that were obtained from the Sukhaya Salma Strait (66.3116°N, 33.6470°E) in August 2014 when they were roughly one month old, and subsequently released into a freshwater pond (66.2637°N, 33.4295°E) uninhabited by stickleback. They were re-captured after 11 months, in July 2015. Their body lengths were 54, 54 and 56 mm. All fish were stored at -20°C until further analyses. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national guides, and have been approved by the Institutional Animal Ethics Committee (Zoological Institute RAS, St. Petersburg, Russia).

Three types of structures were analyzed: otoliths, gill covers, and pelvic spines. Fish were thawed, and all structures were extracted using forceps and scissors. To obtain otoliths, the head was cut on the top midline between the frontals, and the lateral parts were opened. The brain was removed, and otoliths located in the lateral sides of the skull in the otic area were extracted. Otoliths and gill covers were rinsed with water to clear residual soft tissues. All structures were dried at room temperature and stored in Eppendorf tubes.

2.2. AGE ASSESSMENT

In order to assess inter- and intra-reader variability in age estimates, structures were analyzed by eight different readers, who performed aging independently from each other. Each reader scored the age of each fish either twice (gill covers) or three times (otoliths and spines) with two to seven days between scoring rounds. All readers were blind to the identity of the fish. The total and per reader sample sizes for each structure type are given in Table S1. Six readers (TA, SV, DL, TI, MI, PG) had previous experience with otolith aging, whereas two (AYu, KN) did not. None of the readers had previous experience in aging from spine cross-sections, and only two readers (DL, TI) had experience in aging from gill covers.

2.2.1. OTOLITHS

Otolith analyses were conducted using digital images taken with a Leica DMLB microscope under 10× magnification with an integrated digital Leica DC 300 camera. Only the largest otoliths, known as sagitta, were used for age estimates (Jones & Hynes, 1950). Otoliths were placed on a glass slide and immersed in 1,2-propanediol to obtain more transparent and clear images (DeFaveri & Merilä, 2013; DeFaveri *et al.*, 2014). In an earlier study, otoliths were polished in order to obtain more clearly demarcated annuli (Jones & Hynes, 1950). However, in the current study all annuli in otoliths were deemed to be easily visible in a drop of 1,2-propanediol (Fig. 1). Hence, as in many previous studies (Greenbank & Nelson, 1959; Tiller, 1972; Allen & Wootton, 1982; Pichugin *et al.*, 2008, Golovin *et al.*, 2015), otoliths were not polished in this study.

Images were organized in PDF-format files for further analyses, where each page included two otoliths from each focal individual (Fig. 1 A-B), and two reference otoliths of marine yearlings (the largest and smallest ones individuals in the sample; Fig. 1 C-D). Images of all four otoliths on the page were in the same scale, allowing comparison between focal and reference otoliths, particularly to estimate the size of the first year annuli. Hence, a final PDF file included 48 pages – one for each focal individual. Three copies of this final file, each having randomly numbered pages, were distributed to each reader in order to estimate intra-reader variance in replicate age estimates. Otoliths of the 11-month-old pond-transplanted juveniles (Fig. 1 E-F) were not included in the PDF files. However, they were used to verify the age of the reference juveniles caught in the lagoon. Specifically, the patterns in their otoliths were similar to those in the marine juveniles (cf. Fig. 1 C-D *vs* E-F), whose growth patterns should be similar to those of the experimental fish. Hence, the pond-transplanted juveniles were used to confirm that the lagoon-caught individuals were yearlings, and these were further used as reference fish.

2.2.2. GILL COVERS

Gill covers were analyzed according to the method applied earlier to sticklebacks (Mukhomediarov, 1966; Patimar *et al.,* 2010; Yershov & Sukhotin, 2015). Gill covers were removed with scissors and mechanically cleared of soft tissues. The age was determined directly from the physical samples by counting the annuli on the operculum bone. This was carried out under normal transmitted light conditions with a stereomicroscope without the use of optics. Gill covers of the yearlings from the sea were used as a reference.

2.2.3. PELVIC SPINES

Pelvic spines were extracted from the fish by cutting them as close to the spine base as possible. Spine slices were prepared for age analysis using a previously applied method (DeFaveri & Merilä, 2013, DeFaveri *et al.,* 2014). Briefly, pelvic spines were set in epoxy and sectioned using a Struers Accutom 100 (slice width 200 µm, motor rotation speed 3000 rpm; Struers ApS, Ballerup, Denmark). Later, spine sections were mounted on an object glass and photographed with Leica DMLB microscope with an integrated digital Leica DC 300 camera.

Similar to the otoliths, digital images of spine cross-sections were organized into a PDF file. Each file included images of 30 individuals randomly taken from the sample of 48 fish (Fig. 2). Three copies, each with images in a randomly shuffled order, were distributed to each reader for intra-reader variance estimates. No reference images of juveniles were used for this structure.

2.3. STATISTICAL ANALYSES

The ages of the 48 focal fish were independently estimated by eight readers, two to three times for each of the three structures. This generated a total of 2337 age estimates; the distribution across readers and structure types are summarized in Table S1.

We adopted two different linear mixed effect models (LMM) to evaluate which of the three structure types would provide the most accurate basis for age estimation. The first model was defined as:

$$y_{ijkl} = \beta_0 + x_{ijkl}\beta_i + \alpha_j + \alpha_{ik} + e_{ijkl}, \qquad (1)$$

where *i*, *j*, *k*, *l* represent structure, sample ID, reader and replication (within structure and reader), respectively. The parameter β_0 is the fixed intercept, β_i is the fixed structure effect, $\alpha_j \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0,\sigma^2)$ is the random effect of sample ID, $\alpha_{ik} \overset{\text{i.i.d}}{\sim} N(0,\sigma_i^2)$ is the random effect of the interaction between structure and reader (which have structure-specific variances), and $e_{ijkl} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0,\sigma_0^2)$ is the residual error.

Equation (1) was used to evaluate the variances of age estimates, $\hat{\sigma}_i^2$, among the eight readers for each structure. The smaller the variance, the more consistent the age estimates among different readers are. Hence, this model offers a means to identify which aging structure is likely to provide the most reliable age estimates for three-spined sticklebacks.

The second model was defined as:

$$y_{ijkl} = \beta_0 + x_{ijkl}\beta_i + \alpha_j + \alpha_{jk} + e_{ijkl}, \qquad (2)$$

where $\alpha_j \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0,\sigma_1^2)$ is the random effect of sample ID, $\alpha_{jk} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0,\sigma_2^2)$ is the random effect of the interaction between ID and reader, and the residual error $e_{ijkl} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0,\sigma_{0i}^2)$ is assumed to follow a normal distribution with structure-specific variances. All other variables and parameters were defined in the same way as in equation (1). Hence, the difference between this and the previous model is that the random interaction effect between fish ID and reader was used instead of the interaction effect between structure and reader, and the residual variances were assumed to be structure-specific. This model provides a means to estimate the technical repeatability of the whole data from the equation:

$$\hat{\sigma}_i^2 = \hat{\sigma}_1^2[\hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \mathrm{mean}(\hat{\sigma}_{0i}^2)]^{-1} \qquad (3)$$

where $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_{0i}^2$ are the estimated variances of the random effects on the sample ID, the interaction between ID and reader, and the structure, respectively. In other words, the equation gives the proportion of the variance explained by different fish (but not by structures and readers) of the total variance equalling the overall repeatability of age estimates. Conversely, the reciprocal of *X* (i.e. 1-*X*) equals the proportion of variance explained by measurement/estimation error. Repatabilities are thus expressed as proportions ranging from zero to one; the higher the value, the higher the repeatability.

Furthermore, LMMs similar to those described in equation (2) were used to calculate the technical repeatability separately for each structure type:

$$y_{jkl} = b_0 + a_j + a_{jk} + e_{jkl}, \qquad (4)$$

as well as to estimate the technical repeatability separately for each reader:

$$y_{ijl} = \beta_0 + x_{ijl}\beta_i + \alpha_j + e_{ijl}, \tag{5}$$

and to estimate the technical repeatability separately for each structure and reader:

$$y_{jl} = \beta_0 + \alpha_j + e_{jl}, \tag{6}$$

A summary of the equations used in data analysis is provided in Table 1. In these analyses, a potential concern is that the dataset was incomplete, in the sense that there was unbalanced representation of repeated measurements in different groups (c.f. structures and readers). In principle, LMM should be able to provide reasonable estimates of the between- and within-group variances, even when the data are unbalanced (Cnaan *et al.*, 1997). However, to ensure that our conclusions were not biased because of the imbalance, we performed extra analyses using the same models (1)–(2), and (4)–(6) on a subset of the data, by randomly selecting roughly equivalent numbers of repeated measurements within each reader and/or structure.

To provide intuitive illustrations of how age estimates obtained from different structures are related, we calculated their degree of correlation. For these analyses, we averaged the estimates for a given fish over the repeated measures within and across the readers. We estimated both Pearson product moment correlation and linear least squares regression statistics for these comparisons.

All the LMM analyses were performed using the statistical packages “nlme” (https://cran.r-project.org/web/packages/nlme/index.html) and “lme4” (https://cran.r-project.org/web/packages/lme4/index.html) with R3.3 software (R Core Team, 2013).

3. RESULTS

3.1. REPEATABILITIES OF AGE ESTIMATES

In the LMM analysis of the unbalanced data using model (1), the estimated inter-reader variances were 0.030 (SE = 0.003, 95% CI: [0.024, 0.036]) for otoliths, 0.067 (SE = 0.017, 95% CI: [0.034, 0.100]) for gill covers, and 0.191 (SE = 0.041, 95% CI: [0.111, 0.271]) for spine cross-sections. This suggests that the ages estimated from otoliths are generally more consistent across readers than from the two other structure types. LMM analysis of models (2) and (4) revealed that technical repeatabilities ranged from 0.53 to 0.69, with otoliths having the highest technical repeatability (Table 2).

Intra-reader repeatability varied greatly depending on the structure used (Table 2). For otoliths, which were scored by eight readers, repeatabilities ranged from 0.52 to 0.88. Gill covers were scored by five readers, and repeatabilities ranged from 0.37 to 0.73. Repeatabilities for spine-based estimates (scored by seven readers) ranged from 0.67 to 0.91. In fact, for four out of the seven readers, spine cross-sections gave the highest repeatabilities (0.75–0.91).

To further evaluate the significance of inter-reader variance, model (1) was compared with a LMM without the structure-by-reader random effect using ANOVA. Model (1), which included the structure-by-reader effect, was indeed favored over a model without this effect ($P = 2\times10^{-16}$), which indicates that there is a significant inter-reader effect in the data.

To further validate the results, analyses conducted on a randomly selected subset of the data revealed qualitatively similar results to those obtained with the entire dataset. The inter-reader variances were 0.034 for otoliths, 0.041 for gill covers, and 0.192 for spine cross-sections. Likewise, although the overall technical repeatabilities were lower in the subset than in the entire dataset (Table S2), the repeatabilities across the entire (Table 2) and subsampled (Table S2) data were highly correlated $r^2 = 0.85$, $P = 4\times10^{-9}$). Hence, the unbalanced data structure did not have a marked effect on within- and between-group variances.

3.2. DIFFERENCES IN AGE ESTIMATES AMONG STRUCTURES

Age estimates were as follows (Mean ± SE): otoliths: 2.0 ± 0.7, gill covers: 2.2 ± 0.7, spine cross-sections 2.6 ± 1.0 years. Differences in age estimates were significant in all pairwise comparisons between structures (ANOVA, $P < 0.01$). These differences resulted mainly from a higher proportion of younger age groups estimated with otoliths, and older age groups estimated with spines (Fig. 3). Despite these discrepancies, two-year-old fishes were the modal age group, regardless of the structure analyzed (Fig. 3). Age estimates were correlated between gill covers and otoliths ($r^2 = 0.66$, $P = 2.8\times10^{-7}$), and between gill covers and spines ($r^2 = 0.49$; $P = 0.005$), but not between otoliths and spines ($r^2 = 0.32$, $P = 0.09$; Fig. S2).

4. DISCUSSION

We assessed the reliability of three different structures for age estimation in three-spined sticklebacks and found that otoliths provided the most repeatable source of data for age assessment in this species. For gill covers and spine cross-sections, intra- and inter-reader repeatabilities were lower, thus providing less reliable age estimates. Indeed, the differences in age estimates between the three assessed structures were significant. This result is in agreement with other studies that have also observed tendencies to under- or overestimate fish age depending on the structure used (e.g. Elzey *et al.*, 2015). Significant correlations were found between age estimates based on gill covers *vs*. otoliths, and between gill covers *vs*. pelvic spines, but not between otoliths *vs*. pelvic spines. Moreover, we also demonstrated a significant reader effect, which may have a considerable influence on the reliability of age scorings from different structures. In the sections below we will discuss these findings and their implications, as well as reflect upon practical considerations when choosing particular structures for aging sticklebacks.

4.1. RELIABILITY OF AGE ESTIMATES

Similar to the current study, earlier studies of other fish species have also used repeatabilities as a measure of the reproducibility of measurements, i.e. a measure of precision (Campana, 2001) or reliability of age estimates. Although there are doubts about the use of repeatability as a reliable measure of accuracy of age estimates (Beamish & McFarlane, 1983), this approach is often the only one that is applicable when fish age is unknown, particularly in wild-caught individuals. Other related measures of aging precision have been widely used (e.g. average percent error, coefficient of variation, percent agreement between readers; Beamish & Fournier, 1981; Campana, 2001; Sotola *et al.*, 2014; Elsey *et al.*, 2015; Khan *et al.*, 2015). However, use of these different approaches has made comparison among different studies difficult (Campana, 2001).

Among the studies that have compared aging structures for their reliability, otoliths are reported to be superior to other structures such as gill covers, spine or vertebrae sections (Zymonas & McMahon, 2009; Ma *et al.*, 2011; van der Meulen *et al.*, 2013; Elzey *et al.*, 2015). Our results complement these earlier findings in other fish species, and suggest that otoliths are also the most reliable structure for aging three-spined sticklebacks. This notion is supported by the finding that among the eight readers, variances for age estimates were lowest and technical repeatabilities highest (average = 0.69) for otoliths. Pelvic spine cross-sections had intermediate variances and repeatabilities (0.63), whereas gill covers proved to be the least reliable, yielding the highest variances and lowest inter-reader repeatabilities (0.53). This contrasts with studies on *Micropterus sp.* (Sotola *et al.*, 2014) and *Coregonus lavaretus* (Raitaniemi *et al.*, 1998), which found opercules to be the best structure for aging the analyzed species. A combination of factors such as availability of reference material of known age (otoliths, gill covers), inherent difficulty of reading annuli, and existing knowledge about their formation could have led to differences among readers in the current study. Indeed, otoliths were considered by readers as convenient structures for age estimation due to the presence of reference yearling otolith photos (Fig. 1), relatively clear annuli, and available information on their formation (Jones & Hynes, 1950; Tiller, 1972; Allen & Wootton, 1982; Singkam & MacColl, 2018). The repeatabilities for otoliths could have been even higher if not for extra markings observed in the central part of larger otoliths, which were absent in yearlings (Fig. 1 A&B *vs.* C&D). Similar marks have been seen in earlier studies (Tiller, 1972; Pichugin *et al.*, 2008; Moser *et al.*, 2012), and a method to remove these "false" rings has been developed (Odelström *et al.*, 2014). However, these rings can be seen in relatively large otoliths of fishes older than yearlings, and it is unclear if they should be removed from age estimates. Hence, having both known-age yearlings and older fish as references during analyses would most likely further improve repeatabilities and accuracy of age estimates.

In this study, age was scored by eight independent readers, allowing analysis of a reader effect on repeatability estimates. Indeed, this effect was significant. Readers differed not only in average repeatability (0.48-0.86), but also in the structure that proved to give the most repeatable age estimates. For instance, as discussed above, otoliths seem to be the most reliable structure when compared at the inter-reader level. However, half of the readers reached their highest repeatabilities when scoring pelvic spine cross-sections. This was quite unexpected, since none of the readers had previous experience with age scoring from this structure. In general, marks on spine cross-sections were quite clear, although there were some difficulties in finding the inner marks and counting the outer marks. Apparently, most readers counted these marks in their own way with quite high repeatability, but each reader did this differently, resulting in the relatively low inter-reader repeatability (0.63). This could be explained by the lack of comparative material (no reference yearlings as for otoliths), making inter-reader standardization impossible. Similar to the case of otolith-based scoring, having reference samples of known age – ideally both young and old – could help increase repeatabilities for spine cross-sections.

Gill covers provided the fastest material to prepare for aging sticklebacks, and building on evidence from other species (Raitaniemi *et al.*, 1998; Sotola *et al.*, 2014), proved to be an efficient method for age assessment when large sample sizes had to be processed. However, gill covers had the lowest repeatability of all three structures, and only one out of five readers who scored age from gill covers more than once reached their highest repeatability with this structure. We attribute such results to the difficulty of reading annuli in three-spined stickleback gill covers. In contrast to the other two structures, age from gill covers was scored from physical material rather than photos. This was done primarily because we hoped to see if gill covers might be a quick and reliable way to score age in sticklebacks. We also tried to score annuli from digital images, but this proved to be even more difficult because annuli were more poorly visible in photos than in the physical structures. Consequently, we allowed readers to analyze materials under the conditions they found best (stereoscopes, light regime, etc.), and this lack of standardized protocol most likely increased the variability in age estimates. While the overall variance for age estimates was highest for gill covers, some readers managed to have very high technical repeatabilities (0.70–0.73). This would suggest that gill covers could indeed be a quick method for repeatable age estimation, as reported in studies of other species (Raitaniemi *et al.*, 1998; Sotola *et al.*, 2014). However, individual readers can differ greatly in their consistency when using this structure, even when having access to reference material. Hence, along with additional considerations (see below), this suggests that gill covers may lead to overestimation of true fish age.

4.2. DATA ON OTOLITH FORMATION TO INCREASE RELIABILITY OF AGE SCORING

Growth increments are formed on otoliths on a daily basis (von Hippel *et al.*, 2013), which generates a pattern of zones (rings) that differ in their optical characteristics depending on the season in which they are formed (Jones & Hynes, 1950; Tiller, 1972; Allen & Wootton, 1982). At the beginning of otolith development, a centre (core/nucleus) develops first, which can be observed during and just after the breeding season (Jones & Hynes, 1950). According to Jones & Hynes (1950), a transparent ring begins to lay down thereafter, followed by an opaque zone laid down later on. However, later studies demonstrated different patterns, where opaque zones were thought to be formed during the summer and transparent zones during other seasons (Tiller, 1972; Allen & Wootton, 1982; Singkam & MacColl, 2018).

Our material on yearlings reared in the pond provided valuable data on otolith development in northern stickleback populations. Otoliths of yearlings sampled in early July had a clearly visible wide central opaque zone surrounded by a narrow transparent zone (Fig. 1, E&F). This allowed us to conclude that a wide opaque zone appears around the nucleus first in the summer (not clearly distinguishable off nucleus in figures). Therefore, this opaque zone corresponds to a period of intense growth, as proposed earlier (Tiller, 1972). Following this growth phase, a narrow transparent zone develops and remains to be the outermost until the next summer (growing) season. This pattern of otolith formation in northern stickleback is different from that described earlier (Jones & Hynes, 1950), but agrees with later observations (Tiller, 1972; Allen & Wootton, 1982; Singkam & MacColl, 2018).

4.3. DIFFERENCES BETWEEN STRUCTURES IN AVERAGE AGE ESTIMATES

We found that the estimated age structure of our sample was skewed towards older individuals when scored from gill covers and spine cross-sections as compared to otoliths. A possible explanation for these differences can be attributed to ontogenetic properties. Specifically, otoliths are mainly inorganic structures in which minerals are deposited consistently throughout ontogenesis (Brothers *et al.*, 1976; Campana & Thorrold, 2001; von Hippel *et al.*, 2013), whereas osseous tissues undergo a dynamic process of remodeling (Witten & Hall, 2015). Therefore, physiological processes such as stress and compensatory growth can leave extra markings on osseous skeletal structures, which can be incorrectly counted as annuli. Since otoliths are not skeletal, similar processes do not leave such marks on these structures, rendering them less susceptible to age overestimation. This possibility is supported by studies that have found higher age estimates when using bones instead of otoliths (Elzey *et al.*, 2015; Blackwell *et al.*, 2016; Yates *et al.*, 2016).

An important consideration in age scoring is that due to inflation of growth with age, there is a progressive reduction in the distances between annuli on aging structures as age increases, leading to less clear annuli on older individuals. This tendency affects both otoliths and skeletal structures. However, in some cases annuli may be easier to identify from cross-sections of skeletal structures than from those of otoliths.

The majority of stickleback studies that have used otoliths for aging have found that individuals die after their first breeding season, typically at the age of two to three years (Table S3, Jones & Hynes, 1950; Tiller, 1972; Allen & Wootton, 1982; Wootton, 1984; Reimchen, 1992; Pichugin *et al.*, 2008; Bergström *et al.*, 2015). Studies that have used pelvic spine cross-sections (or a combination of spines and otoliths) found that sticklebacks can reach six to eight years of age (Table S3; Reimchen, 1992; Gambling & Reimchen, 2012; DeFaveri & Merilä, 2013). Similarly, relatively high age estimates (up to seven years old) were obtained for the closely related nine-spined stickleback (Herczeg *et al.*, 2009; DeFaveri *et al.*, 2014). This is almost twice the known maximum age observed in populations within the five genera of Gasterosteidae (Reimchen, 1992). However, since the study by Reimchen (1992) was based on a mark-recapture approach, it is unlikely that use of pelvic spine cross-sections led to overestimated lifespans. Likewise, the Finnish studies (Herczeg *et al.*, 2009; DeFaveri & Merilä, 2013; DeFaveri *et al.*, 2014) utilized both otoliths and spine cross-sections, which should reduce the likelihood of upwardly-biased age estimates. Similarly, there is a clear tendency for the average age in stickleback populations to increase with increasing latitude (DeFaveri

& Merilä, 2013), and the abovementioned aging studies (conducted with spine cross-sections) tend to come from more northern populations than those conducted with otoliths.

High age estimates from bone structures have also been suggested in other studies of White Sea sticklebacks. Age estimated with gill covers demonstrated that three-year-old fish prevailed in samples from the Kandalaksha Bay, while one-year-old fish were not found (Ivanova *et al.*, 2007; Yershov & Sukhotin, 2015). This is puzzling since other studies using otoliths (this study, Golovin *et al.*, 2015), found mature one-year-old stickleback from the same area, despite having smaller sample sizes. Furthermore, the average size of four-year-old fish was smaller in studies where age was estimated with gill covers (Yershov & Sukhotin, 2015) than with otoliths (Mukhomediarov, 1966). This could indicate that studies using gill covers for aging, especially in the absence of reference individuals, may overestimate stickleback age.

Despite some discrepancies in age estimates between structures and readers (Fig. S1), overall stickleback age estimates were quite similar, and ranged from one to five years, with a modal age group of two years old (Fig. 3). Therefore, estimates of modal population age were quite similar, and the differences between structures and readers were found in the 'marginal' age groups. Regular discussion between readers (researchers), and knowledge of how growth affects different structures can reduce discrepancies in age estimates (Golovin *et al.*, 2015).

Finally, as to the sampling of material from different populations, our study used material only from one particular area, the White Sea. Repeating this study with material from additional populations from different geographic locations and environmental conditions – such as from lakes, especially those where sticklebacks are known to reach older ages (e.g. Herczeg *et al.*, 2009; DeFaveri & Merilä, 2013; DeFaveri *et al.*, 2014) – could provide additional insights into the reliability of different structures in age estimation.

4.4. PRACTICAL CONSIDERATIONS FOR THE CHOICE OF STRUCTURE

When selecting a method for aging, practical considerations (summarized in Table 3) should be weighed. For instance, otolith extraction may require special training, and spine sectioning requires specific equipment. Age scoring was rather uncomplicated for otoliths due to the obvious annuli; this structure had the lowest intra-reader variability in age estimates. Pelvic spine slices were moderately difficult to read because annuli located closer to the outer edge were dense. Annuli on gill covers were the least pronounced, thus leading to uncertainty within and between readers. Although all structures can be safely stored dry at room temperature, it is important to caution that alcohol or formaldehyde can eventually dissolve otoliths (Jones & Hynes, 1950; Pichugin *et al.*, 2008), if not properly buffered (Kristoffersen & Salvanes, 1998; Gagliano *et al.*, 2006; Schnell *et al.,* 2016). Pelvic spines and opercular bones can be more easily prepared in the field without compromising the specimens' integrity. However, use of these structures for aging should be interpreted with caution, as illustrated by the results of the current study.

In this study we used digital images for scoring otoliths and spine sections. It is likely that this approach increased the repeatability of the age estimates. Using digital images offers several benefits, including their easy exchange and simple arrangement for repeatability scoring. Images of gill covers were not used because annuli could not be seen in pictures as clearly as in physical samples.

4.5 CONCLUSIONS

In conclusion, we found that stickleback ages estimated from otoliths were more precise (repeatable) as compared to those from pelvic spines and gill covers. Otoliths were also relatively easy structures to prepare and exchange as digital images among readers. Furthermore, since the development of otoliths through ontogenesis has been thoroughly studied, this increases their reliability in age estimation. Therefore, it is our recommendation that whenever possible, analysis of otoliths should be the preferred approach for aging three-spined sticklebacks. In studies where otolith extraction is incompatible with the main purpose of the study, for instance where head anatomy is being studied, or when preservation of specimens in either ethanol or formaldehyde had led to otolith breakdown, other structures could be used. The results of this study provide important insights into the differences in age scoring that can result from using different structures within the same fish.

Acknowledgements

The authors thank J. DeFaveri for her comments and language editing on the earlier versions of the manuscript. The administration of the Educational and Research station of the Saint Petersburg State University “Belomorskaia” is thanked for providing opportunity to conduct fieldwork in the White Sea. Our work was supported by the Exchange Programme between the University of Helsinki and Saint-Petersburg State University (travel grants #1.57.1150.2015 and 1.57.189.2016, to AYu), grants from Academy Finland (#129662, 134728 and 218343 to JM), Societas pro Fauna et Flora Fennica and Nordenskiöld foundations (to KN), Russian Foundation for Basic Research (grant № 18 04 01052 AFBR for MI, TI and PG), and the Ministry of Science and Higher Education of the Russian Federation, project AAAA-A19-119020790033-9).

SUPPORTING INFORMATION

Figure S1 Box plots of the age estimates for the three structures types (Gill covers: red, Otoliths: green, and Spines: blue for all the readers (on top left) and for each reader separately (the rest).

Figure S2 Pairwise scatter plots of age estimates (averaged over readers and replicate measurements) based on three different aging structures. The red lines represent linear regression lines ($P = 2.8\times10^{-7}$ for gill covers *vs*. otoliths; $P = 0.005$ for gill covers *vs*. pelvic spines; $P = 0.09$ for otoliths *vs*. pelvic spines).

Table S1. Sample sizes (number of individuals) in the study together with number of repeated age measurements (in parentheses) per reader (columns) and structure type (rows).

Table S2. Technical repeatability of age estimates for different structure types and readers as estimated on the basis of a randomly selected subset of data, where each reader and structure had the same number of measurements**.**

Table S3. Age estimates of the three-spined stickleback from different parts of their distribution range with use of three types of reference structures: otoliths, gill covers and pelvic spines.

ORCID

Anastasia Yurtseva http://orcid.org/0000-0003-1404-7152

Kristina Noreikiene https://orcid.org/0000-0001-7529-4902

Dmitry Lajus http://orcid.org/0000-0002-2264-5825

Zitong Li http://orcid.org/0000-0001-8469-7295

Tatiana Ivanova http://orcid.org/0000-0003-4186-7084

Mikhail Ivanov http://orcid.org/0000-0002-8277-7387

Juha Merilä http://orcid.org/0000-0001-9614-0072

Contributions

A.Yu.: ideas, data generation and analysis, manuscript preparation; K.N.: ideas, data generation and analysis, manuscript preparation; D.L.: ideas, data generation and analysis, manuscript preparation; Z.L.: data analysis, manuscript preparation; T.A.: data generation. T.I.: ideas, data generation; M.I.: ideas, data generation; P.G.: ideas, data generation; S.V.: data generation; J.M.: ideas, manuscript preparation.

References

Allen, J. R. M. & Wootton, R. J. (1982) Age, growth and rate of food consumption in an upland population of the three-spined stickleback, *Gasterosteus aculeatus* L. *Journal of Fish Biology* **21**, 95–105.

Baker, J. A., Heins, D. C., Foster, S. A. & King, R. W. (2008) An overview of life-history variation in female threespine stickleback. *Behaviour* **145**, 579–602.

Barber, I. & Nettleship, S. (2010) From 'trash fish' to supermodel: the rise and rise of the threespined stickleback in evolution and ecology. *Biologist* **57**, 15–21.

Baudouin, M., Marengo, M., Pere, A., Culioli, J. M., Santoni, M. C., Marchand, B. & Durieux, E. D. H. (2015) Comparison of otolith and scale readings for age and growth estimation of common dentex *Dentex dentex. Journal of Fish Biology* **88**, 760–766. doi.org/10.1111/jfb.12816

Beamish, R. J. & McFarlane, G. A. (1983) The forgotten requirements for age validation in fisheries biology. *Transactions of the American Fisheries Society* **112**, 735–743.

Bell, M. A., Aguirre, W. E. & Buck, N. J. (2004) Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**, 814–824.

Bergström, U., Olsson, J., Casini, M., Eriksson, B. K., Fredriksson, R., Wennhage, H. & Appelberg, M. (2015) Stickleback increase in the Baltic Sea: a thorny issue for coastal predatory fish. *Estuarine, Coastal and Shelf Science* **163**, 134–142.

Blackwell, B.G., Kaufman, T.M. & Moos, T.S. (2016) An assessment of calcified structures for estimating northern pike ages. *North American Journal of Fisheries Management* **36**, 964–974.

Brothers, E., Mathews, C. P. & Lasker, R. (1976) Daily growth increments in otoliths from larval and adult fishes. *Fishery Bulletin* **74**, 1–8.

Campana, S. E. (2001) Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of Fish Biology* **59**, 197–242.

Campana, S. E. & Thorrold, S. R. (2001) Otoliths, increments and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 30–38.

Chugunova, N. I. (1959) *Rukovodstvo po izucheniyu vozrasta i rosta ryb [Manual for studying fish age and growth]*. Moscow. AN SSSR. 164 p. Translated from Russian: Israel Program for Scientific Translations, 1963, Jerusalem, Israel.

Cnaan, A., Laird, N. M. & Slasor, P. (1997) Tutorial in statistics: using general linear mixed model to analyze unbalanced repeated measures and longitudinal data. *Statistics in Medicine* **16**, 2349–2380.

Craig, C. F. (1985) Aging in fish. *Canadian Journal of Zoology* **63**, 1–8.

Cresko, W. A., McGuigan, K. L., Phillips, P. C. & Postlethwait, J. H. (2007) Studies of threespine stickleback developmental evolution: progress and promise. *Genetica* **129**, 105–126. doi.org/10.1007/s10709-006-0036-z

DeFaveri, J. & Merilä, J. (2013) Variation in age and size in Fennoscandian three-spined sticklebacks (*Gasterosteus aculeatus*). *PLoS One* **8**, e80866. doi.org/10.1371/journal.pone.0080866

DeFaveri, J., Shikano, T. & Merilä, J. (2014) Geographic variation in age structure and longevity in the nine-spined stickleback (*Pungitius pungitius*). *PLoS One* **9**, e102660. doi.org/10.1371/journal.pone.0102660

Elzey, S. P., Rogers, K. A. & Trull, K. J. (2015) Comparison of 4 aging structures in the American shad (*Alosa sapidissima*). *Fishery Bulletin* **113**, 47–54.

Gagliano, M., Kowalewsky, S. & McCormick, M. I. (2006). An alternative method for the preservation of tropical fish larvae. *Journal of Fish Biology* **68**, 634–639. doi.org/10.1111/j.0022-1112.2006.00911.x

Gambling, S. J. & Reimchen, T. E. (2012) Prolonged life span among endemic *Gasterosteus* populations. *Canadian Journal of Zoology* **90**, 284–290.

Gibson, G. (2005) The synthesis and evolution of a supermodel. *Science* **307**, 1890–1891.

Golovin, P. V., Ivanov, M. V., Ivanova, T. S. & Lajus, D. L. (2015) Age estimation in the three-spined stickleback *Gasterosteus aculeatus* L. *Vestnik SPbGU* **3**, 47–60. In Russian.

Greenbank, J. & Nelson, P. R. (1959) Life history of the threespine stickleback *Gasterosteus aculeatus* Linnaeus in Karluk Lake and Bare Lake, Kodiak Island, Alaska. *Fishery Bulletin* **59**, 537–559.

Herczeg, G., Gonda, A. & Merilä, J. (2009) Evolution of gigantism in nine-spined sticklebacks. *Evolution* **63**, 3190–3200. doi.org/10.1111/j.1558-5646.2009.00781.x

Hüssy, K., Coad, J. O., Farrell, E. D., Clausen, L. A. W. & Clarke, M. W. (2012) Age verification of boarfish (*Capros aper*) in the Northeast Atlantic. *ICES Journal of Marine Science* **69**, 34–40.

Ivanova, T. S., Korobulin, V. V, Polyakova, N. V. & Lajus, D. L. (2007) Population characteristics of threespine stickleback from Keret Archipelago of the Kandalaksha Bay, White Sea. 8 nauchnaya sessia MBS SPBGU. 8 February 2007. Saint Petersburg, Saint Petersburg State University, p. 23–25 (in Russian). Available at: http://mbs.spbu.ru/wp-content/uploads/2016/07/2007.pdf

Jones, J. W. & Hynes, H. B. N. (1950) The age and growth of *Gasterosteus aculeatus*, *Pygosteus pungitius* and *Spinachia vulgaris* as shown by their otoliths. *Journal of Animal Ecology* **19**, 59–73.

Khan, S., Khan, M.A., Miyan, K. & Lone, F.A. (2015) Precision of age estimates from different ageing structures in selected freshwater teleosts. *Journal of Environmental Biology* **36**, 507–512.

Kristoffersen, J. B. & Gro Vea Salvanes, A. (1998). Effect of formaldehyde and ethanol preservation on body and otoliths of *Maurolicus muelleri* and *Benthosema glaciale*. Sarsia **83**, 95–102. doi.org/10.1080/00364827.1998.10413675

Ma, B., Xie, C., Huo, B., Yang, X. & Li, P. (2011) Age validation, and comparison of otolith, vertebra and opercular bone for estimating age of *Schizothorax o'connori* in the Yarlung Tsangpo River, Tibet. *Environmental Biology of Fishes* **90**, 159–169.

McKinnon, J. S. & Rundle, H. D. (2002) Speciation in nature: the threespine stickleback model systems. *Trends in Ecology and Evolution* **17**, 480–488. doi.org/10.1016/S0169-5347(02)02579-X

Moser, D., Roesti, M. & Berner, D. (2012) Repeated lake-stream divergence in stickleback life history within a Central European Lake Basin. *PLoS One* **7**, e50620. doi.org/10.1371/journal.pone.0050620

Mukhomediarov, F. B. (1966) Three-spined stickleback (*Gasterosteus aculeatus* L.) from Kandalaksha Bay of the White Sea. *Voprosy Ikhtiologii* **6**, 454–467. In Russian.

Odelström, A., Ångström, C., Olsson, J. & Bergström, U. (2014) Staining removes false rings in the otolith of three-spined stickleback. In: Life history/Age structure of three-spined stickleback in the Baltic Sea (book of abstracts). 5th International Otolith Symposium. October 20-24, 2014. Mallorca, Spain. P. 231–232.

Patimar, R., Najafabadi, M. H. & Souraki, M. G. (2010) Life history features of the nonindigenous three-spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758) in the Gomishan wetland (southeast Caspian Sea, Iran). *Turkish Journal of Zoology* **34**, 461–470.

Pichugin, M. Yu., Pavlov, D. S. & Savvaitova, K. A. (2008) Life cycle and structure of populations of three-spined stickleback *Gasterosteus aculeatus* (Fam. Gasterosteidae) in rivers of

Northwestern Kamchatka (with reference to the Utkholok River). *Journal of Ichthyology* **48**, 151–161.

Pravdin, I. F. (1966) *Rukovodstvo po izucheniyu ryb [A Manual on Fish Studies*]. Pishchevaya Promyshlennost', Moscow (in Russian).

R Core Team (2013) A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Raitaniemi, J., Bergstrand, E., Floystad, L., Hokki, R., Kleiven, E., Rask, M., Reizenstein, M., Saksgård, R. & Ångström, C. (1998) The reliability of whitefish (*Coregonus lavaretus* (L.)) age determination - differences between methods and between readers. *Ecology of Freshwater Fish* **7**, 25–35.

Reimchen, T. E. (1992) Extended longevity in a large-bodied stickleback, *Gasterosteus*, population. *Canadian Field-Naturalist* **106**, 122–125.

Schluter, D. (2000) Ecological character displacement in adaptive radiation. *American Naturalist* **156**, S4–S16. doi.org/10.1086/303412

Schneider, J. C., Laarman, P. W. & Gowing, H. (2000) Age and growth methods and state averages. In *Manual of fisheries survey methods II: with periodic updates* (Schneider, J. C. ed.), pp. 1-14. Michigan Department of Natural Resources, Fisheries Special Report 25, Ann Arbor.

Schnell, N. K., Konstantinidis P. & Johnson, G. D. (2016). High-proof ethanol fixation of larval and juvenile fishes for clearing and double staining. Copeia **104(3)**, 617–622. doi.org/10.1643/CI-15-382

Secor, D. H., Dean, J. M. & Campana, S. E. (Eds.) (1995) *Recent developments in fish otolith research.* Columbia: University of South Carolina Press.

Singkam, A. R. & MacColl, A. D. C. (2018) Otolith development in wild populations of stickleback: Jones & Hynes method does not apply to most populations. *Journal of Fish Biology* **93**, 272–281. doi.org/10.1111/jfb.13687

Sotola, V. A, Maynard, G. A., Hayes-Pontius, E. M., Mihuc, T.B., Malchoff, M. H. & Marsden, J. E. (2014) Precision and bias of using opercles as compared to otoliths, dorsal spines, and scales to estimate ages of largemouth and smallmouth bass. *Northeastern Naturalist* **21**, 565–573.

Tiller, I. V. (1972) Vozrast i rost trekhigloy kolyushki *Gasterosteus aculeatus* L. ozera Dal'nego. *Izvestiya TINRO* **82**, 219–225. In Russian.

Van der Meulen, D. E, West, R. J. & Gray, C. A. (2013) An assessment of otoliths, dorsal spines and scales to age the long-finned gurnard, *Lepidotrigla argus*, Ogilby, 1910 (Family: Triglidae). *Journal of Applied Ichthyology* **29**, 815–824.

Von Hippel, F.A., Smayda, L. C., Zimmerman, C. E. & Bell, M. A. (2013) Validation of daily growth increments in otoliths to age threespine stickleback (*Gasterosteus aculeatus*). *Evolutionary Ecology Research* **15**, 947–957.

Watkins, C. J, Ross, T. J., Hardy, R. S. & Quist, M. C. (2015) Precision of hard structures used to estimate age of mountain whitefish (*Prosopium williamsoni*). *Western North American Naturalist* **75**, 1–7.

Witten, P. E. & Hall, B. K. (2015) Teleost skeletal plasticity: modulation, adaptation, and remodelling. *Copeia* **103**, 727–739.

Wootton, R. (1984) *A functional biology of sticklebacks.* Croom Helm, London & Sydney.

Wright, P. J. & Huntingford, F. A. (1993) Daily growth increments in the otoliths of the three-spined stickleback, *Gasterosteus aculeatus* L. *Journal of Fish Biology* **42**, 65–77.

Yates, J. R., Watkins, C. J. & Quist, M. C. (2016) Evaluation of hard structures used to estimate age of common carp. *Northwest Science* **90**, 195–205. doi.org/10.3955/046.090.0210

Yershov, P. & Sukhotin, A. (2015) Age and growth of marine three-spined stickleback in the White Sea 50 years after a population collapse. *Polar Biology* **38**, 1813–1823. doi.org/10.1007/s00300-015-1743–1747

Zhu, X., Wastle, R. J., Howland, K. L., Leonard, D. J., Mann, S., Carmichael, T. J. & Tallman, R. F. (2015) A comparison of three anatomical structures for estimating age in a slow-growing subarctic population of lake whitefish. *North American Journal of Fisheries Management* **35**, 262–270.

Zymonas, N. D. & McMahon, T. E. (2009) Comparison of pelvic fin rays, scales and otoliths for estimating age and growth of bull trout, *Salvelinus confluentus*. *Fisheries Management and Ecology* **16**, 155–164.

Table captions

Table 1. Summary of linear mixed model equations used for estimation of repeatabilities of age scoring.

Table 2. Technical repeatability of age estimates in the three-spined stickleback (*Gasterosteus aculeatus*) for different structure types and readers. The analyses are based on the original dataset. Highest repeatabilities for each reader are indicated in bold-face. Column headings refer to abbreviated reader identities.

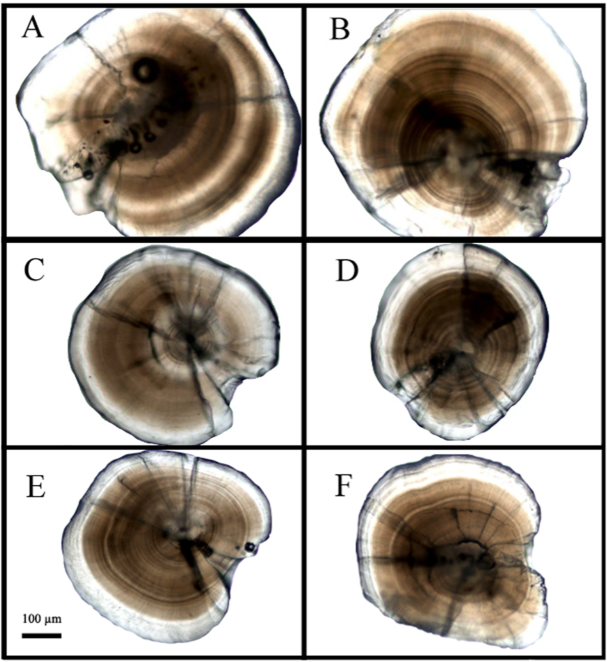
Table 3. Comparison of three aging structures of the three-spined stickleback (*Gasterosteus aculeatus*), in terms of their convenience for treatment and analysis. Characterizations are somewhat subjective based on experiences and results gained from the current study.

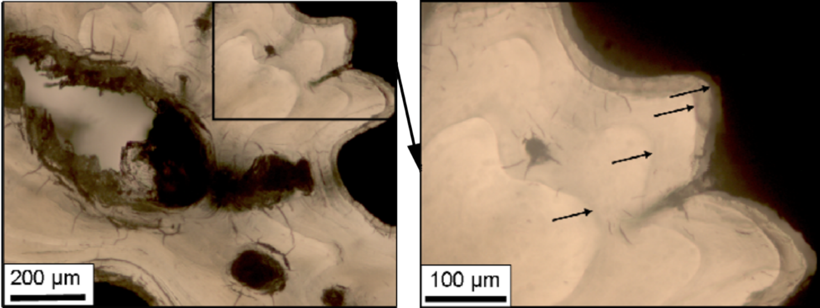
Figure legends

Fig. 1. Images of otoliths used for age estimation in three-spined stickleback. A&B – otoliths of a focal fish in which age is to be identified (2+ in this case), C&D – otoliths of marine yearlings (the largest and smallest in the sample), E&F – otoliths of sticklebacks reared for 11 months in the pond. Light is transmitted from the bottom, hence opaque zones are dark, and transparent zones are white.

Fig. 2. Images of pelvic spine cross-sections used for age estimation in three-spined sticklebacks.

Fig. 3. Distribution of age groups (in %) in the sample of three-spined sticklebacks from the White Sea, based on analysis of three types of aging structures: otoliths, gill covers, and spines. Fish age (years): white – 1, light grey – 2, grey – 3, dark grey – 4, black – 5.


A
B
C
D
E
F
100 µm


200 µm
100 µm

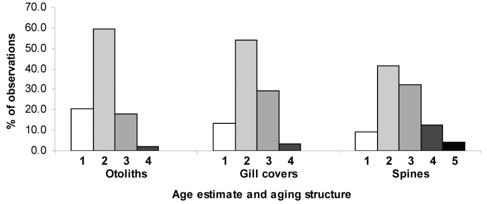

70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
% of observations
1
2
3
4
Otoliths
1
2
3
4
Gill covers
1
2
3
4
5
Spines
Age estimate and aging structure

Table 1. Technical repeatability of age estimates in the three-spined stickleback (*Gasterosteus aculeatus*) for different structure types and readers. The analyses are based on the original dataset. Highest repeatabilities for each reader in bold-face.

	All readers	**AYu†**	**DL†**	**KN†**	**MI**	**PG**	**SV†**	**TA†**	**TI**
All structures	0.57^{E2}	0.48^{E5}	0.74^{E5}	0.86^{E5}	0.61^{E5}	0.59^{E5}		0.62^{E5}	0.61^{E5}
Otoliths	0.69^{E4}	**0.80**^{E6}	0.81^{E6}	**0.88**^{E6}	0.74^{E6}	0.72^{E6}	0.67^{E6}	0.52^{E6}	0.73^{E6}
Gill covers	0.53^{E4}	0.37^{E6}	0.61^{E6}		0.51^{E6}	**0.73**^{E6}			0.70^{E6}
Spines	0.63^{E4}	0.68^{E6}	**0.83**^{E6}	0.77^{E6}	**0.78**^{E6}	0.67^{E6}		**0.75**^{E6}	**0.91**^{E6}

†AYu, DL: Gill covers analysed twice, **KN** and **TA**: Gill covers analysed only once, **SV**: Gill covers and Spines not analysed. Other readers analysed each structure thrice.

E- equation number from Materials and Methods section used to obtain given estimate.

Table 2. Comparison of three aging structures of the three-spined stickleback (*Gasterosteus aculeatus*), in terms of their convenience for treatment and analysis

	Otoliths	**Gill covers**	**Spine sections**
Extraction	Difficult	Easy	Easy
Pre-analysis treatment	Easy	Easy	Difficult
Age reading	Easy	Difficult	Moderate
Exchange of scoring materials	Easy	Moderate	Easy
Intra-reader repeatability	High	Low	High
Inter-reader repeatability	High	High	Very low

Table S1. Sample sizes (number of individuals) in the study together with number of repeated age measurements (in parentheses) per reader (columns) and structure type (rows).

	AY	**DL**	**KN**	**MI**	**PG**	**SV**	**TA**	**TI**
Otoliths	48 (3)	48 (3)	48 (3)	48 (3)	48 (3)	48 (3)	48 (3)	47 (3)
Gill covers	48 (2)	48 (2)	32 (1)	32 (3)	48 (2)		32 (1)	32 (3)
Spines	31 (3)	31 (3)	31 (3)	31 (3)	31 (3)		31 (3)	22 (3)

Table S2. Technical repeatability of age estimates for different structure types and readers as estimated on the basis of a randomly selected subset of data, where each reader and structure had the same number of measurements**.**

	All readers	**AY**	**DL**	**KN**	**MI**	**PG**	**SV***	**TA**	**TI**
All structures	0.46	0.66	0.65	0.81	0.54	0.64		0.41	0.30
Otoliths	**0.68**	**0.83**	0.78	**0.84**	0.77	**0.75**		0.55	0.57
Gill covers	0.51	0.52	0.48		0.49	0.68			0.58
Spines	0.64	0.62	**0.86**	0.79	**0.81**	0.67		**0.74**	**0.93**

*Not in use here

Table S3. Age estimates of the three-spined stickleback from different parts of their distribution range with use of three types of reference structures: otoliths, gill covers and pelvic spines.

Locality	**Structures**	**Age**								**Reference**
	Otoliths	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
Wirral peninsula (Cheshire, NW England)		+	+	+	+	–	–	–	–	Jones & Hynes, 1950
Kodiak Island, Alaska		+	+	+	–	–	–	–	–	Greenbank & Nelson, 1959
Kandalaksha Bay, the White Sea		+	+	+	+	–	–	–	–	Mukhomediarov, 1966
Lake Dal’nee, Kamchatka		+	+	+	–	–	–	–	–	Tiller, 1972
Lake Frongoch (Great Britain)		+	+	–	–	–	–	–	–	Allen & Wootton, 1982
Upland Irish Reservoir System		+	+	+	+	–	–	–	–	Dauod *et al.*, 1985
L'Isle Verte, Quebec, Canada		+	+	–	–	–	–	–	–	Dufresne *et al.*, 1990
Utkholok River, NW Kamchatka		–	–	+	+	+	–	–	–	Pichugin *et al.*, 2008
Lake Constance and tributaries, Central Europe		+	+	+	–	–	–	–	–	Moser *et al.*, 2012
The northern and central parts of the Baltic Sea		+	+	+	+	–	–	–	–	Odelström *et al.*, 2014
The Bothnian Sea and Central Baltic Sea		+	+	+	+	–	–	–	–	Bergström *et al.*, 2015
Solovetsky Islands, the White Sea		+	+	+	+	+	–	–	–	Golovin *et al.*, 2015
	Gill covers									
Kandalaksha Bay, the White Sea		–	–	+	+	+	–	–	–	Ivanova *et al.*, 2007
Nonindigenous, southeast		+	+	+	–	–	–	–	–	Patimar *et al.*, 2010

Caspian Sea, Iran									
Kandalaksha Bay, the White Sea	+	+	+	+	+	–	–	–	Yershov & Sukhotin, 2015
Pelvic spines									
Drizzle Lake, Haida Gwaii, British Columbia, Canada	+	+	+	–	+	+	+	+	Reimchen, 1992
13 localities, Haida Gwaii, British Columbia, Canada	+	+	+	+	+	+	–	–	Gambling & Reimchen, 2012
Fennoscandia	+	+	+	+	+	+	–	–	DeFaveri & Merilä, 2013 (both otoliths and spines were examined)