Testing for parallel allochronic isolation in lake–stream stickleback

D. HANSON, R. D. H. BARRETT & A. P. HENDRY

Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada

Abstract

The evolution of reproductive isolation (RI) is a critical step shaping progress towards speciation. In the context of ecological speciation, a critical question is the extent to which specific reproductive barriers important to RI evolve rapidly and predictably in response to environmental differences. Only reproductive barriers with these properties (importance, rapidity, predictability) will drive the diversification of species that are cohesively structured by environment type. One candidate barrier that might exhibit such properties is allochrony, whereby populations breed at different times. We studied six independent lake–stream population pairs of threespine stickleback (*Gasterosteus aculeatus* Linnaeus, 1758) that are known from genetic studies to show RI. However, the specific reproductive barriers driving this RI have proven elusive, leading to a ‘conundrum of missing reproductive isolation’. We here show that breeding times differ among some of the populations, but not in a consistent manner between lakes and streams. Moreover, the timing differences between lake and stream populations within each pair could account for only a small proportion of total RI measured with neutral genetic markers. Allochrony cannot solve the conundrum of missing reproductive isolation in lake–stream stickleback.

Introduction

Ecological speciation occurs when adaptation by different populations to different environments causes the evolution of reproductive barriers (Schluter, 2000; Nosil, 2012). A classic signature of this process occurs when populations adapted to a given environment are reproductively isolated from populations adapted to different environments but not from populations adapted to similar environments (Funk, 1998; Rundle et al., 2000; Nosil et al., 2002; McKinnon et al., 2004; Schwartz et al., 2010; Ostevik et al., 2012). In animals, studies exploring this ‘parallel ecological speciation’ have focussed on assortative mate choice: females from populations adapted to a given environment prefer to mate with males from populations adapted to similar environments over males from populations adapted to different environments (Rundle et al., 2000; Nosil et al., 2002; McKinnon et al., 2004). However, parallel ecological speciation can be considered in the context of any sort of reproductive barrier, with our focus here being allochronic (temporal) isolation.

Allochronic isolation occurs when groups of individuals (populations or species) breed at different times and therefore are less likely to interbreed (Coyne & Orr, 2004). This reproductive barrier has been frequently studied in plants, where populations with different ecologies (temperature, pollinators, etc.) often have different flowering times, which reduces the probability of intercrossing (Fox, 2003; Weis, 2005; Elzinga et al., 2007; Martin & Willis, 2007; Wagner et al., 2014; Weis et al., 2014). For invertebrates, allochronic isolation has been shown to be important in phytophagous insects that consume plants with different phenologies, leading the insects to also evolve different phenologies (Ragland et al., 2012; Stearns et al., 2013; Medina et al., 2014; Powell et al., 2014). For vertebrates, allochrony has been documented for sympatric seabirds that show repeated divergence in breeding times on each of many small islands (Friesen et al., 2007), and for blackcap
warblers (*Sylvia atricapilla*) with different migration routes that manifest different breeding times in sympatry (Rolshausen *et al.*, 2010). While most of the above examples document allochrony as a product of ecological selection, allochrony could also result from nonecological mechanisms. For example, genetic drift might result in allochrony that is unrelated to ecological differences between two populations (e.g. Devaux & Lande (2008)). In such cases, however, allochrony would not be expected to be parallel when studied in multiple pairs of populations. In short, allochronic reproductive barriers can make important contributions to reproductive isolation and, thereby, speciation in a number of taxa.

Merging these two ideas, parallel ecological speciation might occur through allochrony. Although tests for this possibility are rare, examples do exist. For instance, multiple, sympatric early-winter/late-winter population pairs of moths (*Imorois punctigera*) in different locations in Japan have repeatedly evolved parallel differences in reproductive timing that restrict gene flow (Yamamoto & Sota, 2012). These barriers evolved owing to harsh climactic conditions that dictate two suitable breeding times: early-winter and late-winter. In the present study, we explore whether similar parallel ecological speciation driven by allochrony might occur in lake and stream ecotypes of threespine stickleback (*Gasterosteus aculeatus*).

**Lake–stream stickleback**

Lake and stream populations of threespine stickleback are an excellent system for studying parallel evolution because many such pairs evolved independently after the end of the Pleistocene glaciation – on Vancouver Island (Lavin & Mcphail, 1993; Hendry & Taylor, 2004; Berner *et al.*, 2009; Keauffer *et al.*, 2012), elsewhere in British Columbia (Moodie, 1972; Reimchen *et al.*, 1985; Deagle *et al.*, 1996; Thompson *et al.*, 1997) and around the world (Aguirre, 2009; Berner *et al.*, 2010; Lucek *et al.*, 2010; Ravinet *et al.*, 2013). Studies of these populations have revealed strong lake–stream parallelism in some traits, especially body depth, but also strong lake–stream nonparallelism in other traits, especially defensive armour. Studies of parallelism in reproductive barriers have not yet been attempted.

Parapatric lake and stream stickleback typically show reproductive isolation as indicated by restricted gene flow at genetic markers (Thompson *et al.*, 1997; Hendry & Taylor, 2004; Berner *et al.*, 2009; Keauffer *et al.*, 2012; Roesti *et al.*, 2012; Ravinet *et al.*, 2013). However, studies seeking the specific barriers restricting gene flow (e.g. selection against migrants, assortative mating, habitat choice, etc.) have thus far failed to find a strong and consistent candidate (Hendry *et al.*, 2002; Bolnick *et al.*, 2009; Raeymaekers *et al.*, 2010; Räsänen *et al.*, 2012; Räsänen & Hendry, 2014). This discrepancy between high genetic isolation (low gene flow) and the apparent absence of strong reproductive barriers suggests the ‘conundrum of missing reproductive isolation’ (Räsänen *et al.*, 2012). The present study was motivated by the idea that allochrony (different reproductive timing of lake and stream populations) might help solve this conundrum.

Although the biology of reproductive timing (e.g. sensitivity to photoperiod, temperature, and body size) has been well studied in stickleback (Yeates-Burghart *et al.*, 2009; Borg, 2010; O’Brien *et al.*, 2012), how timing varies among populations has only rarely been considered. In one study that did so, stickleback from different latitudes that experience different photoperiods showed genetic differences in reproductive timing, with northern fish maturing under shorter day lengths (Yeates-Burghart *et al.*, 2009). In the context of parapatric/sympatric populations, however, we must ask to what extent populations experiencing a similar photoperiod show different reproductive timing. In this regard, Hagen (1967) found that sympatric anadromous and freshwater stickleback populations each had a 4-month breeding season; yet those seasons overlapped by only 1 month. Allochrony thus could be an important contributor to ecological speciation in stickleback; yet no study has systematically examined its importance across multiple population pairs.

In the lake–stream context, we can see several reasons why reproductive timing might differ among populations. Depth and flow differences can cause temperature differences between lake and stream, which is known to have a strong effect on sexual maturation (Borg, 1982; Borg & Veen, 1982; Sokolowska & Kulczykowska, 2009). In addition, lakes and streams in temperate regions often differ in the timing of ice thaw and in subsequent temperature profiles (Ashton, 1986). As a result, the timing of availability of breeding sites and prey can differ between lakes and streams (Jensen *et al.*, 2007; Prowse *et al.*, 2011), which might favour different reproductive timing for stickleback. To consider this possibility, as well as its implications for ecological speciation, we address five questions. First, to what extent does reproductive timing differ among stickleback populations? Second, to what extent are timing differences parallel between lake and stream populations across different watersheds? Third, how much reproductive isolation would be expected to result from the timing differences observed among populations? Fourth, does reproductive isolation occur to a greater degree between populations in different environments (lakes vs. streams) than between populations in similar environments (lakes vs. lakes, streams vs. streams)? Fifth, is any of the variation in reproductive timing associated with the genetic, morphological or ecological differences among populations – as would be expected if allochrony was important in driving progress towards ecological speciation?
Materials and methods

We studied paired lake and stream populations of threespine stickleback from each of six independent watersheds on Vancouver Island, British Columbia, Canada (Table S1). These watersheds, and the specific study sites within each (henceforth ‘populations’), were chosen to coincide with previous work on reproductive barriers and parallel evolution in this system (Hendry et al., 2002; Berner et al., 2009; Bolnick et al., 2009; Raeymaekers et al., 2010; Kaeuffer et al., 2012; Räsänen et al., 2012). Importantly, previous genetic studies have confirmed that lake–stream divergence within each watershed likely occurred independently following separate colonization by marine ancestors (Hendry & Taylor, 2004; Berner et al., 2009; Kaeuffer et al., 2012; Roesti et al., 2012, 2014).

Each population was visited at 2-week intervals over a period of 12 weeks, starting at 6 May 2013 and ending at 27 July 2013. Sampling carried out before May 6 resulted in no breeding males or gravid females, so we are confident that we encompassed the start of the breeding season. Similarly, sampling carried out after July 27 resulted in no gravid females, so the end of the breeding season was also included. On each visit, between 13 and 59 unbaited minnow traps were placed across a diversity of microhabitats, with trap deployment lasting 2–9 h. (This variation in trap number and deployment time matched variation in the effort required to capture the maximum number of fish we could process in a day.) Young-of-the-year fish were released immediately after capture and were not included in the analysis. Adult fish were visually scored for sexual maturity (Fig. 1) in a manner consistent with previous work (Milinski & Bakker, 1990; McKinnon et al., 2004; Boughman et al., 2005; Boughman, 2007). Specifically, males were assigned a score between 1 and 3, with 3 being the most colourful (both area and intensity) and 1 being the least colourful (just enough to identify as a male). Colour was evaluated both as red on the throat, mouth, snout, and operculum area, as well as blue in the eye, posterior flank, and in the area around the testes. The area and intensity that would give a maximum score (3) varied between watersheds, as some populations are less colourful than others. Females were scored on a different 1–3 scale: (1) evidence of gravidity but no eggs visible, (2) cloaca everted with visible eggs but eggs not running and (3) cloaca everted with visible running eggs (light squeezing of the abdomen was used to assess the latter). Fish that could not be identified as either male or female were counted as ‘unknown’. All scoring was performed by one author (DH).

Reproductive timing

Denoting $T$ as the total number of fish caught for a particular population (e.g. Misty Stream) during a particular sampling period and $N$ as the number of mature fish (males and females with a score greater than one) for the same population/period, we calculated the percentage of mature fish for that population and sampling period ($\% M = N/T$). To quantify the uncertainty in these estimates, we used a Bayesian approach to construct a likelihood for $\% M$. Taking a uniform prior for $\% M$, this likelihood is given by:

$$L(\% M | T, N) = B(N, T, \% M) (T + 1),$$

where $B(N, T, \% M)$ is the probability mass function for a binomial distribution:

$$B(N, T, \% M) = (T \text{ Choose } N) (\% M)^N (1 - \% M)^{T - N}.$$

These likelihoods allowed us to specify 95% confidence intervals for $\% M$.

We calculated a maturity-weighted mean breeding date ($D$) for each population, which gives the mean breeding date of the entire population over the entire sampling period. Each sampling date was assigned a day of the year ($d$) starting with 1 for January 1; for example, July 17 was $d = 197$. Denoting the estimate of $\% M$ for a particular population/day $d$ as $\% M_d$, we define the maturity-weighted mean breeding date $D$ for a population as

$$D = \frac{\sum d (\% M_d)}{\sum \% M_d},$$

which is directly equivalent to the average date of capture of reproductive individuals. These calculations of $D$ were based on samples of $\% M$ drawn from the likelihood function for each sampling day (described above as $L(\% M | T, N)$).

Allochrony

We tested for allochrony among all possible population pairs by comparing estimates for $\% M$ over time, where $\% M$ was calculated for each of six sampling periods (each period encompassing a 2-week sampling interval). For each population pair, we first calculated Pianka’s overlap index (Overlap) (Pianka, 1974), which ranges from zero (no overlap = total allochronic isolation) to 1 (complete overlap = no allochronic isolation). This index is defined as

$$\text{Overlap} = \frac{\sum_i (\% M_i \times \% M_j)}{\sum_i (\% M_i^2 \times (\% M_j^2)^{-1}},$$

where $i$ and $j$ are the two populations being compared and $p$ is the sampling period, which ranges from 1 to 6. This index is appropriate for our context because resource overlap is a frequently used metric for a number of reproductive barriers (Ramsey et al., 2003; Martin & Willis, 2007; Scopece et al., 2007; Sobel &
Streisfeld, 2015), including allochrony (Husband & Schemske, 2000; Kay, 2006), and Pianka’s index has previously been used precisely in this context (Wright & Calderon, 1995; Lobo et al., 2003; Herreras-Diego et al., 2006). To quantify the uncertainty in these estimates, we generated samples of %M from the likelihood function described above and recomputed Overlap for each sample, obtaining the probability distribution Pr(Overlap | observed) based on 10,000 Overlap estimates. The reported Overlap value for each lake–stream pair is the mean of this distribution, with the standard deviation providing a measure of uncertainty. Subtracting this Overlap value from one provides a measure of prezygotic reproductive isolation attributed to allochrony (RI = 1 – Overlap).

We next used a Monte Carlo simulation to compare these observed Overlap distributions to the distributions expected under the null hypothesis that the two populations in each comparison were indistinguishable in breeding time (i.e. Overlap = 1). For each simulation and pair of populations within a sampling period, we randomly reassigned mature fish between the two

Fig. 1 Representative fish from the Robert’s Stream population showing maturity scores: (a) female = 3, (b) female = 2, (c) female = 1, (d) male = 3, (e) male = 2, (f) male = 1, and (g) unknown.
populations (keeping the total number of fish \(T\) fixed for each pair) and recalculated \(\text{Overlap}\). Repeating this procedure over 10,000 simulations, we obtained a probability distribution \(\text{Pr}(\text{Overlap\null})\) for \(\text{Overlap}\) under the null hypothesis of no difference.

To quantify the agreement/discrepancy between the distribution of observed \(\text{Overlap}\) values and the distribution expected under the null hypothesis, we use the Hellinger distance \(H_{\text{Dist}}\) given by

\[
H_{\text{Dist}} = \frac{1}{\sqrt{2}} \sqrt{\sum_{\text{Overlap}} \left( \sqrt{\text{Pr}(\text{Overlap\null})} \text{observed} - \sqrt{\text{Pr}(\text{Overlap\null})} \text{null} \right)^2}
\]

(Warren et al., 2008). \(H_{\text{Dist}}\) ranges from 0 to 1 and is closest to 1 when the observed and null hypothesis distributions are farthest apart, indicating that the populations deviate from the null hypothesis of no difference in breeding date (and thus exhibit allochrony). Conversely, an \(H_{\text{Dist}}\) value close to 0 indicates that the observed and null hypothesis distributions are close together and the populations are indistinguishable with regard to reproductive timing (no allochrony).

We next compared \(\text{Overlap}\) and \(H_{\text{Dist}}\) values among various types of population pairs: parapatric lake–stream (lake vs. stream within each watershed, e.g. Misty Lake vs. Misty Stream), allopatric lake–stream (e.g. Misty Lake vs. Boot Stream), allopatric stream–stream (e.g. Misty Stream vs. Boot Stream) and allopatric lake–lake combinations (e.g. Misty Lake vs. Boot Lake). These comparisons allowed an assessment of the compatibility (here defined by reproductive timing) of populations from the same environment but different locations, one of the putative signatures of parallel speciation (Funk, 1998; Rundle et al., 2000; Nosil et al., 2002; Ostevik et al., 2012). These comparisons were primarily heuristic because the nonindependence of the data between groups disallows statistical inference to be made.

**Potential predictors of variation in allochrony**

To explore whether timing differences are associated with genetic, morphological or ecological differences between lake–stream pairs, we used previously reported data from Kaeuffer et al. (2012), wherein the same twelve populations were examined. Genetic differences were represented as the average \(F_{st}\) of six putatively selected microsatellites and the average \(F_{st}\) of six putatively neutral microsatellites. Morphological differences were represented as the mean of six \(P_{st}\) values: the first relative warp of body shape, as well as gill raker number, gill raker length, plate number, and dorsal and pelvic spine length (see Kaeuffer et al. (2012) for details). Ecological differences were represented as the mean of three \(E_{st}\) values: (i) proportion of limnetic prey in stickleback stomachs, (ii) trophic position (determined by \(\delta^{15}\text{N}\)) and (iii) the relative importance of different sources of primary production (determined by \(\delta^{13}\text{C}\)). For each of these four measures (average selected \(F_{st}\), average neutral \(F_{st}\), average \(P_{st}\) and average \(E_{st}\)), we calculated Pearson correlation coefficients between the measure and both \(\text{Overlap}\) and \(H_{\text{Dist}}\) across the six watersheds.

**Results**

**Reproductive timing**

Mean breeding date (\(D\)) varied among populations from a minimum of 158.4 (June 7) for Misty Stream to a maximum of 175.7 (June 24) for Pye Stream (Fig. 2). Some of this variation occurred between lake and stream populations within each parapatric pair, most strongly so for Village Bay (lake fish tended to breed later than stream fish). However, few of the differences among populations were dramatic and the direction of the difference was not consistent between lake and stream habitats (i.e. no parallelism). For instance, the estimated mean breeding date was earlier for lake fish than for stream fish in three systems, whereas the opposite was true in the other three systems. Moreover, the mean of the mean breeding dates across all lakes was 168.3 ± 4.5 (June 17) and across all streams was 165.9 ± 7.3 (June 14).

**Allochrony**

Among all possible population pairs, Pianka’s \(\text{Overlap}\) ranged from an intermediate level (0.59 ± 0.07 for Pye Stream vs. Village Bay Stream) to almost complete

![Fig. 2 Mean breeding date (D) where breeding date ranges from 1 (January 1) to 365 (December 31) by habitat type and watershed.](image-url)
overlap (0.94 ± 0.04 for Beaver Lake vs. Misty Lake) (Fig. S1). For parapatric pairs, mean Overlap was 0.82 ± 0.11, with the lowest overlap for Village Bay (0.61 ± 0.09) and the highest overlap for Beaver (0.92 ± 0.04) (Fig. 3, Table 1). Thus, parapatric reproductive isolation based on allochrony ranged from 39% for Village Bay to 8% for Beaver. When comparing these Overlap values to simulated values under the expectation of no differentiation, the greatest deviation was seen in Village Bay (H_{DIST} = 0.65) and the lowest deviations were in Beaver (H_{DIST} = 0.16) and Robert’s (H_{DIST} = 0.11) (Fig. 3, Table 1).

Lake–stream parallelism in allochrony was very low. First, as noted above, some pairs showed much greater Overlap and smaller H_{DIST} values than did others. Second, the partial allochrony that was present in some pairs was in different directions (lake earlier vs. stream earlier). Third, no differences were evident between the four possible classes of population pairs: parapatric lake–stream, allopatric lake–stream, allopatric lake–lake and allopatric stream–stream (Fig. 4, Table 2).

Potential predictors of variation in allochrony

None of the correlations between measures of differentiation and Overlap or H_{DIST} were statistically significant (Table 3). The largest correlation between Overlap and any of the morphological, ecological or genetic measures of differentiation was with F_{ST} of neutral markers (r = 0.54). That is, as Overlap values increase, indicating less differentiation in reproductive timing, F_{ST} of neutral markers increases, indicating more genetic differentiation. The smallest coefficient, 0.09, was between Overlap and mean E_{ST}. H_{DIST} showed the same pattern, with the largest coefficient of −0.38 with the F_{ST} of neutral markers, and the smallest of −0.18 with mean E_{ST}.

Discussion

For the five questions posed in the Introduction, we conclude the following. First, stickleback populations on Vancouver Island differ in reproductive timing to a degree that ranges from moderate to minimal. Second, the timing differences that occur between lake and stream populations are not parallel; that is, no consistent trend exists for lake populations to breed earlier or later than stream populations, whether in parapatry or allopatry. Third, although allochrony can generate some reproductive isolation (RI) between lake and stream stickleback, its contribution to overall reproductive isolation is expected to be weak in most instances. Fourth, allochronic isolation between lake and stream pairs was not greater than that between stream-stream or lake–lake pairs. Fifth, the degree of allochrony was not correlated with measures of genetic, morphological or ecological differentiation.

Before discussing biological explanations and implications, we need to consider potential methodological limitations. First, our measure of reproductive timing might not accurately reflect actual reproductive timing, whether due to sampling error or methodological issues. For example, our allochrony measure might be biased towards overestimating the amount of overlap because males can exhibit breeding coloration both before and after females are available (while building/defending nests and while defending young). Indeed, it would be beneficial to assess other indicators of reproductive activity, such as reproductive hormone levels and the presence of male nests (Jakobsson et al., 1999; Barber et al., 2001; Mayer et al., 2004). Yet our metric was similar to those that have revealed timing differences in other populations (Hagen, 1967; Shimada et al., 2011), and so we have no reason to suspect it would inappropriate in our specific case. Second, we examined populations in nature rather than in a common-garden environment, which is a typical starting point for studies of adaptive divergence and which should be most directly relevant to RI in nature. Finally, if distinct subpopulations exist within lakes or streams, it would be informative to assess within-lake or within-stream allochrony. This would allow for a comparison with levels of allochrony between lakes and streams, to determine whether allochrony is driven by differences between in the environment types or rather nonecological mechanisms (e.g. genetic drift). Overall, while more work could certainly be performed on this system, our methods seem sufficient to encourage our biological interpretation of the outcomes.

Whither allochrony?

Allochrony was generally weak, suggesting little scope for it to resolve the ‘conundrum of reproductive isolation’ (Räsänen et al., 2012) in lake–stream stickleback. Indeed, the mean Overlap index for parapatric lake–stream pairs was 0.82 ± 0.11, meaning that RI due to allochrony would be only 0.18. Although similarly low values have been reported for reproductive barriers in some comparisons of other taxa (Coyne & Orr, 1989, 1997; Mendelson, 2003; Ramsey et al., 2003; Martin & Willis, 2007) and for other barriers in stickleback (Nosil et al., 2005), mean values in those systems tend to be considerably higher. Thus, it seems that allochrony, although it might sometimes make a contribution to reproductive isolation in lake–stream stickleback, is so weak as to not warrant further attention as the main barrier promoting ecological speciation.

Allochrony is low in parapatric stickleback likely as a result of considerable temporal overlap in the environmental cues determining reproductive timing (e.g. photoperiod, temperature, flow, prey availability). In addition, the relatively long breeding season (several months) dictates that divergence at the beginning or
Fig. 3 Histograms of observed Pianka values (Overlap) for lake and stream populations (solid line) and simulated single populations (dashed line) in left panels and %M in right panels (open = lake, closed = stream). Vertical line in left panels represents Pianka value (Overlap) calculated on the data not taking sampling error into consideration (see Methods).
end of the breeding season are likely to be swamped by considerable overlap in the middle of the breeding season. A similar pattern was documented in humming-bird-pollinated ginger (*Costus*) plants, which have a breeding season that extends from March to August. Although flowering periods did not overlap at the beginning or end of the season, they did in the middle, resulting in nearly zero RI due to allochrony (Kay, 2006). In the case of *Costus*, other components of prezygotic isolation contributed substantially to RI, including geographical distance (RI = 0.478), habitat isolation (RI = 0.438), and mechanical isolation (RI = 0.769) (Kay, 2006). We are still searching for these ‘other components’ in lake–stream stickleback, as will be discussed below.

Although variation among populations in reproductive timing (which would determine allochrony) was generally weak, some differences were certainly evident (Fig. 2, Fig. 3). For instance, the mean breeding time for lake fish was approximately 14 days later than for stream fish in the Village Bay system (Fig. 2). Yet similar differences were not evident in other systems, being much weaker in some (Beaver and Robert’s) (Table 1, Fig. 2) and reversed in others (Pye, Beaver, and Robert’s) (Fig. 2, Fig. 3). This lack of parallelism in timing suggests that the environmental parameters determining stickleback breeding time do not diverge in a predictable fashion for lakes vs. streams. Given that photoperiod is the same (within less than 10 min) for all breeding sites, some other (nonparallel) parameters

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Overlap</th>
<th>$H_{\text{Dist}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beaver</td>
<td>0.92 ± 0.04</td>
<td>0.16</td>
</tr>
<tr>
<td>Boot</td>
<td>0.85 ± 0.06</td>
<td>0.31</td>
</tr>
<tr>
<td>Misty</td>
<td>0.82 ± 0.08</td>
<td>0.21</td>
</tr>
<tr>
<td>Pye</td>
<td>0.79 ± 0.07</td>
<td>0.36</td>
</tr>
<tr>
<td>Robert’s</td>
<td>0.90 ± 0.05</td>
<td>0.11</td>
</tr>
<tr>
<td>Village</td>
<td>0.61 ± 0.09</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Fig. 4 Boxplots showing Overlap (top) and $H_{\text{Dist}}$ (bottom) values for all contrast types. Bottom and top of the boxes represent the first and third quartiles, and the band inside the box represents the median. Ends of the whiskers represent the most extreme data point that is no more than 1.5 times the interquartile range. Outliers beyond 1.5 times the interquartile range are shown with circles.
must be contributing to any timing differences. Such cues could include the timing of ice-out (as a function of the timing and magnitude of spring runoff), temperature (in combination with water depth), water clarity, productivity, or a combination of these or other factors. As one example, variability in the timing of ice-out (‘breakup’) is likely greater among streams and among lakes than between these two habitat types on average (Magnuson et al., 2000). Additionally, temperature differences between lakes and streams might differ among watersheds due to variation in depth, flow and latitude (Misty and Beaver lakes are separated from the other four watersheds by approximately 0.5 degrees of latitude). Explicitly testing this hypothesis would require detailed data on the timing of these environmental parameters experienced by our study populations.

Where now for lake–stream reproductive isolation?

Although our study did not resolve the ‘conundrum of missing reproductive isolation’ in lake–stream stickleback (Räsänen et al., 2012), it provides further insight into the problem. Given that many isolating mechanisms, now including allochrony, have been tested and found to be weak, it becomes increasingly likely that the strong differentiation seen in genetic markers (Thompson et al., 1997; Hendry & Taylor, 2004; Berner et al., 2009; Kaeuffer et al., 2012; Roesti et al., 2012; Ravinet et al., 2013) is the result of several relatively weak barriers working in combination sequentially and asymmetrically in ways that vary from watershed to watershed. Indeed, tests of barriers in isolation of all others may be unrepresentative of processes occurring in nature (Ramsey et al., 2003). Multiple barriers are certainly evident in other stickleback systems, such as assortative mating and selection against hybrids in the benthic-limnetic pairs (McKinnon & Rundle, 2002; Hendry et al., 2009). Yet the lake–stream system remains unique in that these same barriers seem much weaker. For instance, in lake–stream stickleback, tests of assortative mating (Räsänen et al., 2012), selection against migrants (Raeymaekers et al., 2010; Räsänen et al., 2012; Räsänen & Hendry, 2014) and habitat choice (Hendry et al., 2002) have all yielded weak and asymmetrical outcomes. As this list grows, future work would do well to formally consider all components of RI together. This approach is generally rare, but can provide valuable insights into the speciation process (Ramsey et al., 2003; Nosil, 2007).

References


