Assortative flocking in crossbills and implications for ecological speciation

Julie W. Smith¹, Stephanie M. Sjoberg¹, Matthew C. Mueller¹ and Craig W. Benkman², *

¹Department of Biology, Pacific Lutheran University, Tacoma, WA 98447, USA
²Department of Zoology and Physiology, University of Wyoming, Laramie, WY 82071, USA

How reproductive isolation is related to divergent natural selection is a central question in speciation. Here, we focus on several ecologically specialized taxa or ‘call types’ of red crossbills (Loxia curvirostra complex), one of the few groups of birds providing much evidence for ecological speciation. Call types differ in bill sizes and feeding capabilities, and also differ in vocalizations, such that contact calls provide information on crossbill phenotype. We found that two call types of red crossbills were more likely to approach playbacks of their own call type than those of heterotypics, and that their propensity to approach heterotypics decreased with increasing divergence in bill size. Although call similarity also decreased with increasing divergence in bill size, comparisons of responses to familiar versus unfamiliar call types indicate that the decrease in the propensity to approach heterotypics with increasing divergence in bill size was a learned response, and not a by-product of calls diverging pleiotropically as bill size diverged. Because crossbills choose mates while in flocks, assortative flocking could lead indirectly to assortative mating as a by-product. These patterns of association therefore provide a mechanism by which increasing divergent selection can lead to increasing reproductive isolation.

Keywords: Loxia curvirostra; Pacific Northwest; playbacks; public information; reproductive isolation

1. INTRODUCTION

A central goal of evolutionary biology is to understand the importance of various processes driving adaptive radiations [1–3]. Two processes that have attracted considerable attention are ecologically based divergent selection that causes phenotypic divergence between populations, and the evolution of reproductive isolation that isolates them [4,5]. These two processes are integral to speciation and adaptive radiation, and they are often linked [6,7]. In particular, the magnitude of divergent selection has been found to be positively associated with the degree of reproductive isolation [8,9]. In some cases, the mechanism causing this association is straightforward. For example, when divergent selection causes sticklebacks (Gasterosteus aculeatus) to diverge in body size, reproductive isolation increases because mate choice is influenced by body size [10]. However, the mechanism(s) by which reproductive isolation arises is only rarely known [4,5].

When reproductive isolation is causally linked to ecologically based divergent selection it is termed ecological speciation [6]. Although ecological speciation may be common in some taxa [6,8,9], there is very little direct evidence for ecological speciation in birds, largely because most species are so old that alternative non-ecological mechanisms are hard to reject [5,11]. Three exceptions are found in seed-eating birds: Darwin’s ground finches (Geospiza spp. [2]), crossbills (Loxia spp. [12]) and the buntings of Tristan da Cunha (Nesospiza spp. [13]). Darwin’s ground finches are especially well studied and the mechanisms by which reproductive isolation arises are well described and classic [2].

One might expect that similar processes lead to ecological divergence and reproductive isolation in seed-eating passerine birds. However, reproductive isolation arises in numerous ways [4,5], and the mechanisms by which increasing divergent selection leads to increasing reproductive isolation could vary depending on subtle differences in behaviour and ecology. Here, we focus on a possible mechanism by which assortative flocking by phenotype potentially links divergent ecological selection with reproductive isolation in crossbills. This mechanism could have contributed to the continental radiation of over 140 species of cardueline finches including crossbills.

Crossbills have diversified in response to divergent selection favouring specialization on alternative adaptive peaks characterized by different seed resources [14,15], which in the case of crossbills are represented by seeds in the cones of different species of conifers [14–18]. Because of this specialization on different conifers, habitat isolation is an important reproductive isolating barrier for crossbills [12] as different species of conifers are not completely intermingled. However, habitat isolation is incomplete because many taxa feed on a variety of conifers depending on the relative availability of seeds, which varies seasonally [19], and several taxa can occur together and breed in the same area [12, 20–24]. Behavioural reproductive isolation however is extremely high when multiple taxa breed in the same habitat [12]. Assortative pairing by vocalizations is one mechanism contributing to behavioural isolation [25], but other behavioural mechanisms are possible. A role for ecology is suggested by the trend for an increase in behavioural isolation with increasing morphological divergence [12].
Morphological diversity is most evident in the red or common crossbill (L. curvirostra) complex, which was initially partitioned into multiple subspecies [17,26]. More recently, however, red crossbills have been categorized by their distinctive contact calls [20,27]; contact calls (figure 1) are used in a variety of situations—for example, to coordinate activities such as departures by foraging flocks from trees [28]. Within North America, 10 taxa or call types are recognized [20,29,30]. Although the mean bill characteristics of the different call types reside near or on the adaptive peaks for foraging on alternative conifers [14,15], many individuals can be categorized only by differences in their vocalizations [20,23] because of considerable size overlap between the similarly plumaged call types [20]. The large size overlap between call types is not because different call types are highly variable morphologically, as might arise from hybridization [31,32]. For example, the coefficients of variation of bill depth for call types 1 and 2 [33; see also 20] are less than those for five of six species of finch-like Hawaiian honeycreepers (Drepanididae) [32], a group that shared a common ancestor with the cardueline finches ca 6 Ma [34] and a group in which hybridization is probably rare [32]. The key issue therefore is to understand how high levels of reproductive isolation [12,21,23,35,36] have evolved, especially when sympatric forms overlap extensively in size and are so similar in appearance.

One mechanism that could contribute to reproductive isolation between populations of crossbills diverging to exploit alternative conifers is assortative flocking by trophic phenotype. Crossbills use the foraging success of flock mates to more rapidly assess patch quality [28]; such information is termed public information, which is defined as the cues inadvertently produced by others that provide information about the environment [37] and even to choose mates [38]. Because the value of using the foraging success of flock mates depends on the similarity between the feeding capabilities of flock mates [28] and feeding capabilities increasingly differ as bill size diverges [14], assortative flocking by bill size is favoured [28]. Assortative flocking in turn could lead to assortative pairing and reproductive isolation as a by-product because crossbills, such as other cardueline finches, begin to pair while in flocks [12,17,39,40]. This could provide a mechanism by which (behavioural) reproductive isolation increases with increasing divergent selection. Here, we use playbacks to test the prediction that discrimination against other call types increases with increasing divergence in bill size (depth), which is a measure of divergent selection [14,15].

Alternatively, assortative flocking by vocalizations could lead to increasing reproductive isolation with increasing divergent selection if vocalizations diverged as a by-product of divergence in bill size [41,42] or body size [5,43,44]; bill depth and body mass are allometrically related among call types [45]. We compare the similarity of contact calls to determine whether call similarity decreases with increasing divergence in bill size and whether the responses of crossbills are predicted better by call similarity than by bill size differences.

2. MATERIAL AND METHODS
To determine whether crossbills preferentially associate with members of their own call type and whether they show graded responses relative to differences in bill size and call similarity, we quantified the responses of free-flying crossbills to recordings of contact calls from four call types broadcast from a field speaker during 3–14 August 2009 in the Okanogan-Wenatchee National Forest near Lake Wenatchee, Washington. This site was chosen because two call types, types 2 and 4, were very numerous and breeding in a mixed Douglas fir (Pseudotsuga menziesii)—ponderosa pine (Pinus ponderosa) forest.

We used Raven Pro v. 1.2 (Cornell Laboratory of Ornithology, Ithaca, NY, USA) to copy and paste digitally recorded contact calls to synthesize recordings representing monotypic flocks of types 2, 4, 5 [20] and 9 [29]. Each synthesized recording was made using the contact calls of a single individual (figure 1), and calls from 10 individuals of each call type were used to make 10 synthesized recordings per call type. The individuals were chosen based on the clarity of their recordings that we had made over the past 15 years; we did not attempt to capture the variation within a call type. All synthesized recordings were standardized for duration, call rate and amplitude; 30 s of calling at a rate of 88 calls per 30 s followed by a 15 s pause.

The different call types diverge in average bill size (mean bill depths for each call type (mean of the two sexes) are 8.75, 9.34, 9.56 and 9.79 mm for types 4, 5, 2 and 9, respectively [14,29]). Types 2, 4 and 5 are nomadic and overlap broadly in distribution in the Pacific Northwest [20,46,47], whereas type 9 is restricted to two small mountain ranges in southern Idaho, the South Hills and Albion Mountains [29]. We assume that the type 2 and 4 crossbills in our study had
encountered type 5; type 5 was present at the study site although uncommon. However, because the South Hills and Albion Mountains represent such a small area (70 km² of lodgepole pine forest) within the vastly larger ranges of types 2, 4 and 5, most individuals of types 2, 4 and 5 are unlikely to encounter a type 9 crossbill during their lifetime. Consequently, the responses of types 2 and 4 to the recordings of type 9 are assumed to represent responses to an unfamiliar call type. On the basis of genetic data and the palaeoecology of the conifers on which each crossbill specializes, the four call types examined probably diverged recently [47] so that differences in responses to different call types are unlikely to be caused by differences in time since divergence.

Forty trials were conducted, 10 trials for each call type. A trial consisted of broadcasting one synthesized recording of contact calls of a given call type for 150 min from a Saul Mineroff Electronics amplified field speaker (Saul Mineroff Electronics, Inc., Elmont, NY, USA) in fairly open areas adjacent to mature forests. Open areas were chosen because crossbills are more likely to land near the speaker than in forests, and we could observe crossbills from a distance. Ten locations were chosen so that they were separated by at least 250 m, with the exception of one location that was less than 100 m away from another location. We could not eliminate the possibility that the same flock flew over more than once. However, crossbills were very common in the area and fly long distances so that it is likely that many different flocks flew over during the trials. One trial was conducted in the morning (between 06.00 and 08.30) and one in the evening (17.00–19.30) for two successive days for a total of four trials at each location. The morning and evening were selected because crossbills were more active at these times than during midday. Each call type was played at each location, and the order that the call types were played was alternated so that each call type was played with equal frequency in the morning and evening.

During a trial, we recorded whether crossbills landed near the speaker or continued to fly over. We assume that landing nearby represents a greater propensity for association and flocking than if the crossbills continued to fly over because crossbills are highly social and forage in flocks. Flocks that landed in response to the recordings (n = 364) landed within 25 m of the speaker. The contact calls of all flocks of crossbills that either landed or flew over the decoys were recorded using a Sennheiser directional microphone (Old Lyme, CT, USA) and a Marantz PMD-660 or PMD-670 digital recorder (Aurora, IL, USA). In cases where a flock flew over while a flock that had landed still remained perched in trees near the speaker, we recorded the overhead flock by turning our backs to the perched birds and directing the directional microphone at the flock flying over but away from the perched birds. The directionality of the microphone creates recordings with obvious differences in amplitude, making it possible to distinguish between the two flocks. Although the occurrence of a wild flock near the speakers probably acts to increase the tendency of other flocks to respond by landing, it should not bias the results. We made spectrograms of the recorded contact calls using Raven Pro v. 1.2 and assigned each vocalization to a call type as in Smith & Benkman [12]. Even though some flocks may have included more than one call type, we did not record vocalizations of more than one call type from any of the 663 flocks. We assumed that the recorded contact calls represent the only call type in a flock in the analyses, and all flocks were assigned to one call type.

We used the cross-correlation method in Raven Pro v. 1.4 (Cornell Laboratory of Ornithology) to estimate vocal similarity for the contact calls of 20 individuals from each call type, including the calls from the 40 individuals (10 from each call type) used to make synthesized recordings for the experiments. We used the mean cross-correlation value for within and between call type comparisons in the analyses. The cross-correlation method slides two spectrograms past each other, compares their overall similarity with respect to frequency and time, and outputs normalized values of similarity with values from zero to one. A similarity value of zero represents no overlap, whereas a value of one represents complete overlap. To create spectrograms, the sampling rate was set to 44 100 Hz and the sample size option set to 16 bits. Then, we produced spectrograms using the following settings: window type Hann, 248 Hz filter bandwidth, 5.8 ms frame size, frequency grid spacing 172 Hz (discrete Fourier transform size 256 samples) and time grid hop size 128 samples (50% overlap). When producing each spectrogram, the clipping level was adjusted to eliminate extraneous background noise.

We tested crossbill responses by comparing the mean proportion of crossbill flocks of types 2 and 4 that landed in response to playbacks of the different call types. We started by using a generalized linear model (GLM) with an identity link function to determine whether crossbills had a greater propensity to land (the response variable) in relation to vocalizations of their own call type (homotypic) than to vocalizations of other call types (heterotypic). We included all the one-way effects in this model (bill depth difference, responding call types (types 2 and 4), vocalization similarity and familiar call type or not), and an interaction between bill depth difference and call type familiarity because we predicted a decline in landing response with increasing differences in bill depth for familiar but not unfamiliar call types. We did not include other interactions because we did not expect them. We then re-ran the model focusing solely on heterotypic comparisons (excluding data for responses to homotypic call types), and removed variables that were insignificant in the full model. All analyses were conducted using JMP v. 9.0.1 (SAS Institute Inc., Cary, NC). The data are available at: http://dx.doi.org/10.5061/dryad.r6c36.

3. RESULTS

Crossbills were about twice as likely to land in response to the vocalizations of homotypic as heterotypic call types (figure 2; full model in table 1). Other significant effects were bill depth difference and the interaction between bill depth difference, and whether the playback call type was familiar or not to the responding call type (table 1). Responding call type and vocal similarity did not have significant effects on the landing responses of the crossbills (p > 0.85) when accounting for the other variables (table 1). When comparing just the responses to vocalizations of heterotypic call types and excluding variables that were not significant in the full model, the effect of bill depth difference and the earlier-mentioned interaction remained significant (heterotypic model in table 1). These results are evident in figure 2, which indicates that the probability that crossbills land declines with
increasing divergence in bill depth in response to familiar but not unfamiliar heterotypic call types. In addition, once bill depth difference is taken into account, there is no evidence of an effect of responding call type on the response to heterotypic call types (table 1). Although we did not detect any evidence that differences in vocalizations influenced the behaviour of crossbills when bill depth differences were included in the analysis (table 1), vocalizations may still diverge as a consequence of divergence in bill depth (and body mass). Vocalizations tended to diverge between call types with increasing divergence in bill depth (figure 3; including only heterotypic comparisons: $r = -0.80$, $p = 0.03$). However, much of the variation among the different call type comparisons was related to the distinctive vocalizations of the smallest billed call type (type 4; figure 1) rather than to bill size differences (figure 3). In particular, vocalization similarity increased with increasing bill size differences between call types when we included the effect of making comparisons between type 4 or not (GLM, $\chi^2 = 29.0$, $p < 0.0001$; the effect of type 4 was significant ($p < 0.0001$) and vocalization similarity increased with increasing bill depth difference (slope = 0.056, s.e. = 0.020, $p = 0.02$)).

### 4. DISCUSSION

During our study, type 2 crossbills were observed foraging exclusively on ponderosa pine, and type 4 fed exclusively on Douglas fir so that foraging and flying flocks were completely assortative. However, mixed call type flocks, especially flocks foraging in trees, occur in other locations and other years. For example, types 2 and 4 can be found foraging together on the same Douglas fir trees where ponderosa pine is absent (C. W. Benkman 1984–1985, personal observation). Although in this situation, the different call types usually fly off separately so that flocks in flight tend to be more assortative than those foraging in trees (C. W. Benkman 1984–1985, personal observation). Consequently, we believe our measures of flocking propensity are representative of flying flocks joining foraging flocks when conifer species are not partitioned, which is also when the potential for hybridization is greatest. Next, we discuss the results and implications of our experiments.

Crossbills responded preferentially to the vocalizations of their own call type. Additionally, the use of recordings of contact calls in our experiment demonstrates that free-flying crossbills were responding to the vocalizations and not to other variables such as plumage or morphology. This is consistent with the experiments with captive crossbills (types 2 and 9) showing that females preferentially associate with males of their own call type even though there were no consistent differences in bill size or coloration between males during the experiments [25]. Similarly, a recent study showed that two call types (types 3 and 4) in captivity called more frequently in response to the recordings of vocalizations (contact calls) of their own call type than to those of heterotypics [48]. Our results indicate that vocalizations play an important role in assortative flocking, and are consistent with the hypothesis that selection has favoured distinct vocalizations for phenotypic recognition (i.e. acting as a marker trait) [25,28].

Although vocalizations tended to diverge with increasing divergence in bill depth (figure 3), much of the variation between call types was related to the distinctive vocalizations of the smallest billed call type (type 4) rather than to bill size differences. When we included both an effect of bill size difference and whether the comparison involved type 4 or not, we found that call similarity did not decrease with increasing bill depth difference. This result indicates that call divergence was not simply a correlated effect of bill (or body) size divergence. Similarly, different call types overlap extensively in bill, and body size yet differ in the structure of their contact calls [20,29]. Furthermore, divergence in bill and body size occurs in the absence of consistent differences in peak sound energy of their contact calls. For example, the main energy of contact calls is similar among North American call types (mostly between 3.5 and 4 kHz; mean bill depths range from 8.1 to 11.0 mm [20]) and overlaps that for the largest crossbill in the world (parrot crossbill *L. pytopsittacus*; 3.7 kHz [27]; 13.23 mm bill depth [23]). These results are not surprising, given that the structure of crossbill vocalizations (figure 1) is unlike that of taxa for which such correlated
Table 1. Summary of the two generalized linear models testing for effects of different variables on the probability that flocks of crossbills landed in response to playbacks of the contact calls of different red crossbill taxa (call types). (The response variable was the mean probability of landing for each combination of variables based on the responses of 663 flocks. Significant $p$-values are in bold.)

<table>
<thead>
<tr>
<th>variables</th>
<th>full model</th>
<th></th>
<th>heterotypic model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>d.f.</td>
<td>$p$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>homotypic versus heterotypic</td>
<td>4.47</td>
<td>1</td>
<td>0.034</td>
<td>—</td>
</tr>
<tr>
<td>bill depth difference</td>
<td>6.46</td>
<td>1</td>
<td>0.011</td>
<td>13.43</td>
</tr>
<tr>
<td>responding call type</td>
<td>0.03</td>
<td>1</td>
<td>0.85</td>
<td>1.96</td>
</tr>
<tr>
<td>call similarity</td>
<td>0.03</td>
<td>1</td>
<td>0.87</td>
<td>—</td>
</tr>
<tr>
<td>familiar call type or not</td>
<td>1.96</td>
<td>1</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>bill depth difference × call type familiarity</td>
<td>11.77</td>
<td>1</td>
<td>0.004</td>
<td>9.11</td>
</tr>
<tr>
<td>overall model</td>
<td>35.45</td>
<td>6</td>
<td>&lt;0.0001</td>
<td>14.60</td>
</tr>
</tbody>
</table>

Figure 3. Vocal similarity (mean cross-correlation value) is highest for homotypic comparisons (black-filled circles) and decreases with increasing divergence in bill depth. Heterotypic comparisons between non-type 4 crossbills and type 4 crossbills are designated with different symbols (empty and grey-filled circles, respectively).

effects have been detected [41; but see 49], and contact calls, unlike sexually selected songs, are unlikely to be under selection to reach performance limits so that vocalization structure is correlated with bill morphology [43].

Selection may favour only some minimal level of vocal distinctiveness between co-occurring groups of crossbills that is sufficient for recognition [49], with the variation in response to heterotypic vocalizations being based on learned recognition mechanisms [5,50]. The observed level of divergence in call structure between call types could have arisen if individuals alter their vocalizations away from confusing signals (e.g. vocalizations of other call types) in response to benefits from accurate recognition [49,51]. This mechanism is consistent with the divergence in call structure within adult, type 9 individuals across years from the calls of the commonest co-occurring call type (type 2) [52]. Although the shift was slight, it was significant, and such a shift in vocalizations within a generation would then be passed down to the subsequent generation because crossbills imitate the calls of their parents [20,52,53]. This process could then repeat itself in subsequent generations.

We predicted that if public information use was an important mechanism promoting assortative flocking [28], then the extent to which crossbills distinguish between their own call type and that of others would increase with increases in the average difference between the bill sizes of the two call types. This is exactly what we found (figure 2 and table 1). The contrast between the graded responses to familiar heterotypics in relation to differences in bill depth, and the absence or near absence of such a response to unfamiliar heterotypics (figure 2 and table 1) further suggests that the graded response to familiar heterotypic call types is probably learned (see Irwin & Price [50]).

A critical step however is how a correlation arises between the marker trait and the ecological trait within a crossbill population experiencing divergent selection for foraging on alternative conifers. Call structure varies among individuals within a call type [20,52], and variation among individuals is used by crossbills to recognize individuals, including their mate [54]. Crossbills often even modify the fine structure of their contact calls to imitate their mates [21,52,54,55], although contact calls continue to act to identify the call type of an individual's parents, except in rare cases [52]. Once a variant call arises, over time and by chance (e.g. cultural drift) it might tend to be associated with, for example, large-billed crossbills within the population. Assuming crossbills have a propensity to flock assortatively by call structure, larger individuals with this variant call would be at an advantage because they would tend to flock with other large-billed crossbills and have increased foraging efficiency [28]. Larger-billed individuals with the original call would be at a disadvantage because the original call is associated with both large and small individuals. The advantage of assortative flocking would be especially great when morphological differences are large such as when the population relies on two (or more) species of conifers and experiences divergent selection favouring different sizes of bills [14,15]. In this situation, both frequency- and density-dependent selection should increase bill size variation [56].

The increased feeding efficiencies of large-billed individuals with the variant call should translate into reproductive advantages and an increase in the number of large-billed offspring with variant calls because bill size is highly heritable in crossbills [23], and crossbills imitate the contact calls of their parents [20,52,53]. As the variant call increases among large-billed individuals, the original call will eventually be associated with mostly or only small-billed individuals, thereby increasing their foraging efficiency and fitness. In time, fitness will
stabilize between the two nascent call types that differ in bill size, flock assortively and differ in conifer use. Reproductive isolation between call types will ensue because of habitat differences [12], crossbills choose mates within flocks [12,17], females preferentially pair with males of their own call type [29] and because of song divergence between call types [29] perhaps resulting from divergence in bill and body size [42]. Given the similarity of the above-hypothesized process to Dieckmann & Dobeli’s [57] model for sympatric speciation, call divergence and speciation in crossbills may not have required a period of allopatry, although initial allopatry or differences in habitat preferences would facilitate the process.

5. CONCLUSIONS
Assortative flocking by phenotype may have been critical in enabling the rapid diversification of crossbills. Speciation often requires over 1 million years in birds [5]; yet many populations of crossbills appear to have progressed far in the speciation process within the last 10,000 years [12,47]. It is conceivable that public-information-based assortative flocking has increased the cohesiveness of the different call types and has been critical in promoting the behavioural and ecological distinctiveness of call types beyond what divergent selection for foraging on alternative conifers alone has favoured. In the absence of selection favouring assortative flocking for resource assessment, the crossbill complex (outside of isolated areas where crossbills rely on a single species of conifer) may instead have been simply one variable panmictic population, because the alternative conifers are not sufficiently distinct to favour discrete morphological groups as found in most sexually reproducing species [4].

Our experimental results, combined with the observation that crossbills choose mates within flocks are consistent with studies that demonstrate that the degree of reproductive isolation is positively associated with the magnitude of divergent selection [8,9], including studies of crossbills [12], and provides an additional mechanism promoting ecological speciation [6]. Such processes may also be important in the ecology and evolution of morphologically different vocal types within other species of cardueline finches [40,58]. If the conditions giving rise to the correlation between divergent selection and assortative flocking in crossbills are general to other cardueline finches, as it has been suggested by Price [5].

The experiment meets the conditions of the Pacific Lutheran University Institutional Animal Care and Use Committee.

We thank Murdock Charitable Trust Life Sciences Grant and Pacific Lutheran University Division of Natural Sciences Summer Research Programme for providing support for the fieldwork. P. Edelaar, T. Hahn, D. Irwin, T. Parchman, T. Price, L. Snowberg and anonymous reviewers provided very helpful comments on the manuscript.

REFERENCES
23 Summers, R. W., Dawson, R. J. G. & Phillips, R. E. 2007 Assortative mating and patterns of inheritance indicate


