IS THE NOMINATE SUBSPECIES OF THE COMMON CROSSBILL LOXIA C. CURVIROSTRA POLYTYPIC?
I. MORPHOLOGICAL DIFFERENCES AMONG YEARS AT A SINGLE SITE

PIM EDELAAR1,* & KEES TERPSTRA2


The discovery of discrete vocal types of Common Crossbill in Western Europe opens the possibility that the nominate subspecies Loxia c. curvirostra in fact consists of a group of cryptic, vocally differentiated and reproductively isolated sibling species, reflecting a similar situation in North America. We compared measures of Common Crossbills collected at a single Dutch site by a single observer from 1983 to 2001. During 1983-84 – and to a lesser extent also in 1985-88 and 1992 – Common Crossbills had relatively long wings, low body masses, and deep but short bills compared to other years. Changes in methods or phenotypic flexibility of the measures do not explain these results. Biometric differences among years are likely linked to the proportional abundance of different populations of Common Crossbills at the catching site. A difference between years in the relationship between wing length and bill depth supports the idea that differences between populations are the result of selection, not neutral differentiation. These results are consistent with the variable presence of multiple, at times sympatrically occurring, cryptic species.

Keywords: Loxia curvirostra - cryptic speciation - resource selection – biometry - vocalisation

1Department of Biology MSC 3AF, New Mexico State University, Las Cruces NM 88003, USA; 2Stenenkomweg 7, 8162 NT Epe, The Netherlands; *Present address: Absveen 29, 2211 EX Noordwijkerhout, The Netherlands; E-mail W.M.C.Edelaar@umail.leidenuniv.nl

INTRODUCTION

Recent work has shown that in North America the Common or Red Crossbill Loxia curvirostra is comprised of at least nine different taxa. These taxa differ in their vocalisations, in biometry, in distribution (Groth 1993) and in resource specialisation (Benkman 1993, 1999). Some of these taxa are regularly found breeding sympatrically, but hybridisation seems rare (Benkman 1993; Groth 1993). Therefore, these taxa most likely represent reproductively isolated species. This differentiation of crossbills was already indicated by the description of several subspecies within North America, based on morphological traits only. However, authors disagreed on the characteristics of these subspecies and on their distribution since collections at the same location but in different years yielded very different kinds of crossbills (reviewed in Groth 1993). It is now becoming clear that the occurrence of different types/species of crossbills depends on the resources available at a particular time, and that breeding and wintering distributions of the types/species may vary widely among years.

Discrete vocal types of Common Crossbills
have recently also been identified in Western Europe, based on the so-called flight calls and excitement calls (Clouet & Joachim 1996; Robb 2000; Summers et al. 2002; Edelaar et al. 2003). In contrast to the American vocal types, very little is known about the distribution, morphology, ecology, and evolutionary status of the European vocal types. However, the discovery of vocal types in Western Europe opens the possibility that the nominate subspecies *Loxia c. curvirostra* is actually polytypic (i.e. comprised of several subspecies with little geographic overlap in breeding distribution), or even comprised of a group of cryptic sibling species (with overlapping breeding distributions). This paper, dealing with morphological differences among years of Common Crossbills caught at the same site, is a first attempt to address these issues.

Several studies have described the biometry of Common Crossbills in Western Europe. In line with the presence of morphologically different populations, some studies report differences between years (Davis 1964; Herremans 1988; Marquiss & Rae 2002), and differences also exist between studies performed on different locations (Clouet & Joachim 1996; Cramp & Perrins 1994; Summers et al. 1996). Such biometric differences in time and space are indicative of the presence of different populations, but alternative explanations exist. Measures may vary due to phenotypic flexibility, e.g. different degrees of wear (Benkman 1993). Also, the observed differences are usually subtle, and differences in measuring methods among observers (Summers et al. 1996) and even between years by the same observer are likely to explain at least part if not most of these differences. Another drawback of the published studies is that they mostly deal with birds caught during irruption years, thus biasing against any resident local birds that may have a quite different morphology (Marquiss 1980).

We present an analysis of the biometry of Common Crossbills caught during 19 consecutive years on the same site and measured by the same observer. Hence this unique dataset overcomes the limitation of previous studies regarding irruptive years and the problem of multiple observers. It allows us to test whether different morphological types of Common Crossbills occur within the range of the nominate subspecies *Loxia c. curvirostra*, and whether these are differentially represented among years at a single site. In doing so, we assume that individuals of putatively resident subspecies from outside the currently accepted range of *L. c. curvirostra* (occurring in northwestern Africa, the Balearic Islands, Corsica, Cyprus, far eastern Siberia: Cramp & Perrins 1994) do not occur at the study site, or at least not in significant numbers.

**METHODS**

837 Birds were caught and banded under license in the northeastern part of the Veluwe, The Netherlands (52°23′N, 5°55′E) during 1983-2001. The site is situated in an area of predominantly dry Scots Pine *Pinus sylvestris* plantations, although a small (one square km) plot of mature Douglas-fir *Pseudotsuga menziesii* is nearby. Individuals were attracted to caged Common Crossbills and the presence of small ponds for drinking and bathing, and were caught in mist nets upon approach (predominantly before drinking or bathing). The types of measurements taken varied between years, and thus some useful measures such as bill width could not be included in the comparison without sacrificing data from most other years. We decided to limit the analysis to wing length (maximum stretched, to nearest mm), bill length (from beginning of skull to tip of upper mandible, to nearest 0.1 mm), bill depth (callipers placed at start of feathering on forehead and perpendicular to cutting edge, to nearest 0.1 mm) and body mass (to nearest gram), as these were measured in most of the years (Svensson 1992). All measures were taken with the same wing rule, callipers and pesola throughout the whole study. Birds were sexed and aged using the criteria in Svensson (1992). Five ageing categories were used: 1st calendar year, 2nd calendar year, after 1st calendar year, after 2nd calendar year and full-grown birds. Birds that were not sexed were excluded from analysis. This also
excludes all newly fledged and potentially still growing juveniles that did not yet show body moult: biometric measures of birds with some adult feathers have reached their final size (Groth 1993).

Common Crossbills are known to be sexually dimorphic, and juveniles are sometimes smaller than adults (especially in wing length and body mass). Therefore we corrected all individual measures for age and sex based on the effect sizes of the factors sex and age combined in an ANOVA of each of the four log-transformed variables (all years combined). The effect of sex was always highly significant ($P < 0.0001$). The effect of age was only close to significance for bill depth ($F_{4,831} = 2.10, P = 0.079$) and significant for body mass ($F_{4,831} = 2.43, P = 0.046$), but not for wing length ($P = 0.47$) and bill length ($P = 0.91$). We still corrected all variables for both sex and age in order to keep all manipulations of the data consistent (just like Marquiss & Rae (2002)). Furthermore, correcting for non-significant factors removes existing small biases that could not be detected statistically, without introducing statistically significant new biases. We visually examined probability distributions of these corrected data to test for normality and outliers: the data conformed to normality, and no outliers were identified.

In The Netherlands crossbill influxes normally start in June. If local food conditions are good, birds may stay until the next spring. Once conifer seeds are shed in April/May, most birds disappear again. We thus divided the records in ‘immigration-years’, running from 1 June to 31 May the next year, and identify each immigration-year by the year of arrival. We tested for immigration-year differences in biometry by comparing PCA scores of all birds. For this, we first corrected for immigration-year differences in each of the four biometric variables and then calculated four principal components for the combined dataset (a ‘generalised’ Common Crossbill). Again, the $\chi^2$ probability distribution of $t^2$-values showed no outliers (Johnson & Wichern 1988). Next, we calculated PCA scores for each of the birds based on the factor scores of each of the four variables, using the data that was corrected for age and sex but not for immigration-year. Finally, we tested for immigration-year differences of PCA scores by ANOVA.

Any differences found could be caused by differences in measurement; i.e. the biometric methods, not the Common Crossbills, differ among immigration-years. If a change in methods occurs, it would most likely happen only once or would change gradually (yielding a trend without reversals). We therefore present the average for each of the four measures plotted against immigration-year for visual inspection of such a pattern.

If differences in methodology can be sufficiently excluded, we remain with the question why Common Crossbills have a different morphology among immigration-years. To distinguish between differences due to neutral evolution and due to selection, we compared two groups of immigration-years when Common Crossbills differed most, and plotted the two traits that are least affected by phenotypic flexibility (wing length and bill depth, see Discussion). Assuming that the positive correlation between these two traits (see Results) has a genetic basis, neutral changes in biometry between populations would be expected to occur along this correlation. However, if populations differ in biometry away from this correlation, it is less likely that this is due to neutral evolution, and selection is the more parsimonious explanation (Schluter 2000). We used ANCOVA to test whether significant changes in intercept and slope of the relationship between bill depth and wing length are present.

RESULTS

Using the age-, sex- and immigration-year-corrected data, the first principal component describes variation in size as judged by the large positive loadings of all four traits (table 1). The other principal components represent shape variation, with different traits being important in different principal components. Low PC2 scores identify
birds with relatively long bills and low body masses, low PC3 scores identify birds with relatively long wings, and low PC4 scores identify birds with relatively deep, short bills and low body mass.

Differences among immigration-years in individual scores for each of the four principal components are highly significant (ANOVA - PC1: $F_{18,818} = 17.4$, $P < 0.001$; PC2: $F_{18,818} = 10.7$, $P < 0.001$; PC3: $F_{18,818} = 18.3$, $P < 0.001$; PC4:

**Table 1.** Factor scores for each of four biometric traits of Common Crossbills caught at a single Dutch site (data corrected for age, sex and immigration-year differences). Important factor scores (based on their absolute deviation from zero and standard error) are underlined. Eigenvalues and percentage of variance explained by each principal component are listed at the bottom of the table.

<table>
<thead>
<tr>
<th>Trait</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>wing length</td>
<td>0.64</td>
<td>0.24</td>
<td>-0.73</td>
<td>-0.06</td>
</tr>
<tr>
<td>bill depth</td>
<td>0.73</td>
<td>-0.16</td>
<td>0.29</td>
<td>-0.60</td>
</tr>
<tr>
<td>bill length</td>
<td>0.64</td>
<td>-0.64</td>
<td>0.02</td>
<td>0.43</td>
</tr>
<tr>
<td>body mass</td>
<td>0.63</td>
<td>0.59</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>eigenvalue</td>
<td>1.74</td>
<td>0.84</td>
<td>0.76</td>
<td>0.65</td>
</tr>
<tr>
<td>% variance</td>
<td>43.6</td>
<td>21.0</td>
<td>19.1</td>
<td>16.3</td>
</tr>
</tbody>
</table>

**Fig. 1.** Variation among immigration-years in the morphology of Common Crossbills caught at a single Dutch site. Plotted are annual means ± SE of scores for four principal components based on log-transformed age- and sex-corrected measures of wing length, bill depth, bill length and body mass.
Judging by the $F$-values of the ANOVA’s (equivalent to the $R^2$-values), differences among immigration-years are not mostly caused by general size differences (PC1) as was true for the immigration-year-corrected data, but mostly by the shape-component PC4 (deep but short bills and low body mass).

In order to visualise any patterns in biometric differences among immigration-years, we plotted the scores of each of the four principle components against immigration-year (Fig. 1). A first glance at the plotted data reveals that the immigration-year-effects are roughly similar for all four principal components: in each of the four sub-plots the immigration-years 1983 and 1984 stand out as having low scores. Birds caught in those first two immigration-years can be described as small Common Crossbills with a relatively long wing, a low mass, and a deep but short bill. A more detailed comparison, taking into account the standard error around the means, reveals that (especially for PC1 and PC3), the same is true for the immigration-years 1985-88 and 1992.

Figure 2 depicts the averages per immigration-year for each of the four measures. Significant year-to-year differences occur (wing length: $F_{18,818} = 2.81, P < 0.001$; bill depth: $F_{18,818} = 2.23, P = 0.002$; bill length: $F_{18,818} = 13.85, P < 0.001$, body mass: $F_{18,818} = 26.9, P < 0.001$), but not in a very orderly fashion: measures frequently increase and decrease as time progresses.

To test for neutral versus selected evolution, we created two groups of Common Crossbills that differ most (based on fig. 1): those from 1983-84, and those from 1989-2001 excluding 1992. Figure 3 shows the individual values for wing length and bill depth. Bill depth increases with increasing wing length (ANCOVA - $F_{1,640} = 32.2, P < 0.001$), and both intercept and slope of the two regression lines differ (group-effect: $F_{1,640}$...
DISCUSSION

The measures of Common Crossbills caught at the same site and measured by the same observer differed significantly among immigration-years. Such differences can be interpreted in several ways, and we will discuss these below.

First, perhaps the differences are caused by differences in measurement; i.e. the biometric methods, not the Common Crossbills, differed among immigration-years. We find this unlikely for a number of reasons. In the time period presented here, reversals in the averages of the measurements occur frequently (see Fig. 2) instead of only once or gradually. Additionally, a limited sample measured by the second author in 1971 and 1973 shows that these Common Crossbills were similar to those in the early 1990s (excluding 1992), so Common Crossbills were unlikely to be measured differently in 1983-85 due to inexperience or a change in methodology. Furthermore, clear differences between immigration-years are apparent in all measures, even those that are not easily measured differently among years (e.g. body mass). We therefore reject this explanation for the differences among immigration-years in Common Crossbill measurements.

Second, immigration-year differences may be due to phenotypic flexibility in the traits. This explanation seems supported by Fig. 2, as mostly bill length and body mass seem to vary among immigration-years. Bill length is known to be influenced by variation in wear: when crossbills feed on closed cones or cones with thicker and stronger scales, wear increases and bill length decreases (Benkman 1993). Body mass is inherently variable, largely depending on storage of fat. If this explanation is true, then bill length and body mass should be higher when food is abundant and easily obtained: in April/May when the old cones open and expose their seeds. However, in immigration-years when both bill length and body mass were low (1983-85), bill length was but body mass was not higher in those months (not shown). In fact, the highest monthly mean of body mass in those three years was still lower than the lowest monthly mean in years when birds were heavier (1989-2001). Comparison of our average body mass during 1983-85 with body masses collected during the 1963 invasion in Switzerland also showed that our birds had unusually low body masses: whereas the average of our birds was just below 33 grams, only 0.2% (2 out of 1008) of the Swiss birds had a body mass lower than 33 grams (Newton 1970). The Swiss birds were caught on migration in the early morning, before drinking and foraging, yet they weighed on average 39.8 grams, i.e. more than 20% heavier. We thus conclude that the variability in body mass reflects substantial variation in structural size, not just different amounts of fat. Furthermore, not only bill length and body mass, but also wing length and bill depth showed differ-
ences among immigration-years (Fig. 2), which are less easily (wing length) or not at all (bill depth) explained by wear/phenotypic flexibility. Therefore, other factors besides phenotypic flexibility are needed to explain the large variability in the measured traits among immigration-years.

Third, differences among immigration-years at a site may arise due to selective settlement or survival of Common Crossbills. If so, we would expect the bill morphology in some immigration-years to reflect adaptation to feeding on Scots pine (the local dominant tree). Other pine specialists such as Parrot Crossbill *L. pytyopsittacus*, Scottish Crossbill *L. scotica*, and Mediterranean subspecies of the Common Crossbill are all identified by relatively deep and more decurved bills (Cramp & Perrins 1994) compared to the bill of *L. c. curvirostra* which mostly utilises Norway Spruce *Picea abies*. Birds in 1983-88 and 1992 are characterised by deeper but shorter (so possibly more decurved) bills than in the remaining immigration-years, indicating that in 1983-88 and 1992 Common Crossbills indeed may have been more adapted to foraging on Scots pine. Common Crossbill numbers in The Netherlands fluctuate from a few thousand to several million between years (Bijlsma et al. 2001; Linnartz 2002) and such fluctuations are obviously not caused by selective mortality alone. Thus it is likely that in immigration-years with low numbers in The Netherlands feeding conditions elsewhere were better for Common Crossbills. If the suitability of resources is related to a particular morphology, this can lead to selective settlement. Such assortment of phenotype to resource has been observed in crossbills (Summers et al. 1996; Marquiss & Rae 2002). Perhaps a pine-adapted small resident population occurs at the study site, which is in many immigration-years swamped by invading Common Crossbills with different morphologies from other areas. This is in line with the observation that these putative resident birds (dominating catches in 1983-84, and to a lesser extent in 1985-88 and 1992) basically stand out for all four principal components instead of for some principal components in some immigration-years and other principal components in other immigration-years. However, why these birds have relatively long wings is currently not understood. And actually the pattern in body mass is rather opposite to expectation: this putative ‘pine-specialist’ has a low body mass, whereas other pine-adapted crossbill taxa with deeper, more decurved bills typically have higher body masses (Cramp & Perrins 1994).

Fourth, a decoy of a particular vocal type may selectively attract crossbills of the same vocal type. Thus in theory differences in biometry among immigration-years could be unrelated to Common Crossbill numbers or local food resources, but due to the kind of decoy used in particular immigration-years. However, turnover of decoys was rather large, and new decoys were always taken from the local pool of birds present. Therefore the birds that were caught probably reflect the presence of particular birds in the area quite well. Yet, if the catching of only a selection of all crossbills present due to the use of decoys of a particular vocal type does explain the annual differences in biometry, we have to invoke that the vocal types indeed do differ in biometry. This conclusion would support that the nominate subspecies is in fact polytypic.

So why would Common Crossbills in different areas/years have a different morphology? Given the accumulative body of recent work, this is most likely a reflection of adaptation to particular resources (different species or types of conifers). However, populations may also differ historically due to neutral evolution caused by processes such as founder events, population bottlenecks and genetic drift (Lande 1980). A detailed comparison of the different kinds of Common Crossbills caught in this study shows that it is unlikely that neutral evolution has caused the divergence of populations. When comparing the two groups of Common Crossbills that differed most (1983-84 versus 1989-2001, excluding 1992), both intercept and slope of the relationship between bill depth and wing length differed significantly. Substantial shifts in intercept and especially slope are strong indications that population differentiation is due to selection, since not only overall size but also shape differs. It should be
noted that such a difference in slope also strongly rules out differences in biometry between immigration-years due to changes in measurement techniques and due to phenotypic flexibility (unless there has been some interactive effect, which is difficult to imagine).

Additional support for a selective basis for population differentiation comes from genetic studies. Piertney et al. (2001) investigated divergence between Common, Scottish and Parrot Crossbills, using neutral, highly polymorphic microsatellite and mitochondrial markers, and with samples originating from Great Britain (all taxa) and the Pyrenees (only Common Crossbills). No evidence was found for differentiation in these rapidly evolving markers. If this result of no neutral genetic differentiation between Common Crossbills from Scotland and the Pyrenees extends to Common Crossbill populations from the European mainland (likely, given the more continuous habitat instead of an island situation), then any genetically based phenotypic differentiation must be due to selection, since neutral evolution did not result in divergence for neutral markers (Schluter 2000).

Overall, we show that Common Crossbills caught at the same site and measured by the same observer differed significantly between immigration-years in their morphology. We ruled out or provided strong evidence against these differences being due to changes in methods, phenotypic flexibility of the measures, or neutral evolution of traits. The most parsimonious explanation for the differences in morphology among immigration-years is therefore that selection has caused differences between Common Crossbills populations within the range of the nominate subspecies. Given the enormous fluctuation in numbers in The Netherlands it is likely that there are populations in Europe that differ in morphology due to selection, and that these populations occur at the study site in different proportions in any immigration-year (cf. Davis 1964; Herremans 1988; Knox 1992; Marquiss & Rae 2002). Whether these populations are discrete and have particular (overlapping?) geographical distributions is unknown at the moment. This current lack of knowledge means that an ecological understanding of the selective forces that have differentiated these populations is limited to extrapolations from what is known about morphological adaptation to resources. Such an exercise indicated that one of the populations seems more pine-adapted, at least in relation to bill measures.

The evidence so far identifies that biometric differences due to selection do exist between populations that have different abundances among immigration-years at the catching site. It is tempting to relate the occurrence of such populations to the newly discovered vocal types. This interpretation is aided by the observation that the proportional abundance of the different vocal types in The Netherlands also varies among years (Robb 2000; Weber 1972 for fluctuations in abundance of vocal types). Unfortunately, no calls were recorded of the birds caught in this study, so such a relationship could not be tested, and there are no published studies of the biometry of the European vocal types so far. Thus the obvious next step is to test whether the identified vocal types differ in biometry, indicating ecological differentiation. Recent undertakings have yielded a large combined dataset of over 500 Common Crossbills of which both morphology and vocalisation is known (Kees Terpstra, Kees van Eerde, Erik Maassen, ringing group Nebularia-Westenschouwen unpubl. data), so this will become clearer in the near future. Ecological differentiation is aided by reduced gene flow between populations that are adapting to different resources (Schluter 2001), as seems to be the case for the morphologically and ecologically differentiated North American vocal types. Thus another point of investigation should be whether the European vocal types mate assortatively when sympatric, indicating reproductive isolation and hence evolutionary independence. A growing dataset is being collected on the assortative mating of sympatrically occurring vocal types of the Common Crossbill in order to address this question. Strong assortative mating between the vocal types would be the clearest evidence that Loxia c. curvirostra is actually comprised of a number of cryptic differentiated species. Overall, the results presented in this
paper are not inconsistent with the presence of multiple cryptic, selection driven and at times sympatrically occurring species. As such, this paper provides the first step in identifying if discrete, morphologically specialised populations exist within the geographic distribution of *Loxia c. curvirostra* as it is currently defined.

**ACKNOWLEDGEMENTS**

Lively discussions with Craig Benkman, Magnus Robb, Adam Siepielski, Ron Summers and many others have shaped the ideas for this paper. Useful comments from Craig Benkman, Magnus Robb, Ron Summers, Kees van Eerde, Marc Herremans and Rob Bijlsma on earlier drafts have improved the quality of the paper. This study was financially supported by a TALENT stipend to PE by the Netherlands Organization for Scientific Research (NWO).

**REFERENCES**


**SAMENVATTING**

Uit recent onderzoek blijkt dat Kruisbekken *Loxia curvirostra* zijn in te delen in aparte, discrete groepen op basis van hun geluiden (zogenaamde geluidstypen). In Noord-Amerika is vastgesteld dat dergelijke geluidstypen hoogstwaarschijnlijk aparte soorten zijn, met grotendeels overlappende verspreidingsgebieden. Onlangs zijn dergelijke geluidstypen ook in West-Europa vastgesteld. Het is dus mogelijk dat de bij ons voorkomende nominaatvorm *L. c. curvirostra* ook uit meerdere, reproductief geïsoleerde populaties bestaat. In dit artikel laten we zien dat er duidelijke en consistente verschillen tussen jaren bestaan in afmetingen van Kruisbekken die tussen 1983 en 2001 op dezelfde locatie door dezelfde persoon gevangen en gemeten zijn. Gedurende
1983-84, en in mindere mate gedurende 1985-88 en 1992, hadden de Kruisbekken relatief lange vleugels, lage lichaamsgewichten en hoge, maar korte snavels. De verschillen tussen de jaren kunnen niet goed worden verklaard door een verandering in meetmethode of door verschillen in bijvoorbeeld de slijtage van de vleugelpennen of de hoeveelheid lichaamsvet. De verschillen in biometrie tussen de jaren zijn het best te verklaren door een variabele toestroom van Kruisbekken met verschillende afmetingen. Het verband tussen vleugellengte en snavelhoogte is verschillend tussen jaren. Dit is een sterke aanwijzing dat de verschillen tussen de Kruisbekken veroorzaakt worden door selectie en niet door neutrale evolutie. De gevonden verschillen in biometrie tussen jaren zijn dus overeenkomstig met de variabele aanwezigheid van verschillende kruisbekkenpopulaties (mogelijk soorten) in Nederland.

Corresponding editor: Rob G. Bijsma
Received 18 April 2003, accepted 28 November 2003