doi:10.1111/j.1420-9101.2007.01328.x

Inferring local adaptation from Q_{ST}–F_{ST} comparisons: neutral genetic and quantitative trait variation in European populations of great snipe

S. A. SÆTHER,*,† P. FISKE,‡ J. A. KÅLÅS,‡ A. KURESOO,§ L. LUIGUJÕE,§ S. B. PIERTNEY,¶ T. SAHLMAN* & J. HÖGLUND*

*Department of Population Biology, EBC, Uppsala University, Uppsala, Sweden
†Netherlands Institute of Ecology (NIOO-KNAW), Centre for Terrestrial Ecology, Heteren, The Netherlands
‡Norwegian Institute for Nature Research, Trondheim, Norway
§Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu, Estonia
¶Department of Zoology, University of Aberdeen, Aberdeen, UK

Keywords:

conservation units;
divergent selection;
local adaptation;
microsatellite primers;
partial Mantel test;
Q_{ST};
repeatability as heritability.

Abstract

We applied a phenotypic Q_{ST} (P_{ST}) vs. F_{ST} approach to study spatial variation in selection among great snipe (*Gallinago media*) populations in two regions of northern Europe. Morphological divergence between regions was high despite low differentiation in selectively neutral genetic markers, whereas populations within regions showed very little neutral divergence and trait differentiation. $Q_{ST} > F_{ST}$ was robust against altering assumptions about the additive genetic proportions of variance components. The homogenizing effect of gene flow (or a short time available for neutral divergence) has apparently been effectively counterbalanced by differential natural selection, although one trait showed some evidence of being under uniform stabilizing selection. Neutral markers can hence be misleading for identifying evolutionary significant units, and adopting the P_{ST}–F_{ST} approach might therefore be valuable when common garden experiments is not an option. We discuss the statistical difficulties of documenting uniform selection as opposed to divergent selection, and the need for estimating measurement error. Instead of only comparing overall Q_{ST} and F_{ST} values, we advocate the use of partial matrix permutation tests to analyse pairwise Q_{ST} differences among populations, while statistically controlling for neutral differentiation.

Introduction

It is a long-standing debate in evolutionary biology whether isolation is sufficient, necessary or only helpful for populations to diverge, and whether natural selection can generate divergence in the face of gene flow (Mayr, 1963; Endler, 1977). An increasing number of studies from a diverse range of taxa are highlighting morphological divergence among populations in the absence of differences at neutral genetic polymorphisms, suggesting that local adaptations can evolve over ecological timescales and/or in the absence of population isolation (e.g. Karhu *et al.*, 1996; Kinnison & Hendry, 2001; Piertney *et al.*, 2001; Wilding *et al.*, 2001; Koskinen *et al.*, 2002; Irwin *et al.*, 2005). Spatial variation in selection regimes might therefore be indicated by geographic variation in functional genes, morphology and/or life-history traits, which may prove independent of spatial patterns of neutral diversity (Ekblom *et al.*, 2007). However, if the extent of geographic differentiation is similar to that of selectively neutral genes, random genetic drift is sufficient for explaining the pattern without invoking differential selection in different areas. Therefore, it is

Correspondence: Stein Are Sæther, Netherlands Institute of Ecology (NIOO-KNAW), Centre for Terrestrial Ecology, PO Box 40, 6666 ZG Heteren, The Netherlands.
Tel.: +31 26 4791111; fax: +31 26 4723227;
e-mail: s.saether@nioo.knaw.nl

important to have a baseline of neutral genetic structure as a null model when making inferences about possible local adaptations in, e.g. quantitative traits.

An analysis of intraspecific geographic variation in selectively neutral genetic markers has provided insight into many aspects of population biology, including gene flow between populations, historical demographic events, colonization history and phylogeography, and the identification of conservation units (Avise, 1994, 2000). A growing number of studies are describing substantial levels of geographic structure in bird species, brought about by limited gene flow caused by processes, such as natal philopatry and territoriality (Avise & Ball, 1991; Piertney *et al.*, 1998, 1999). Thus, it cannot be taken for granted that any species will show low levels of geographic structure simply because it is, or has recently been, continuously distributed.

Recently, a number of studies have compared F_{ST} and Q_{ST} estimates for the same populations to make inferences about local adaptation (Merilä & Crnokrak, 2001; McKay & Latta, 2002). F_{ST} (and related statistics) measures the extent of population structuring of genetic variation, and Q_{ST} (Spitze, 1993) is a precisely analogous measure of differentiation in quantitative genetic traits. This implies that when $F_{ST} = Q_{ST}$ there is no evidence for geographically varying natural selection, whereas if $Q_{ST} > F_{ST}$ drift-migration balance cannot explain the entire pattern, and if $Q_{ST} < F_{ST}$ there is evidence for uniform natural selection across populations (Rogers, 1986; Lande, 1992; Whitlock, 1999; Merilä & Crnokrak, 2001; McKay & Latta, 2002). Hendry (2002) draws attention to some potential problems with the assumption that Q_{ST} and F_{ST} should be equal under neutrality.

Q_{ST} is the among-population proportion of the total additive genetic variance, and not phenotypic variance, of a quantitative trait. It is therefore ideally measured in a randomized 'common garden' experimental design to exclude the effect on the trait of environmental differences between populations (with appropriate design to partition the within-population variance, and preferably also to remove nonadditive effects). However, for many organisms it is not possible or practically feasible to conduct laboratory rearing experiments, but phenotypic as well as molecular data may be available from several populations. Fortunately, using realistic assumptions about the additive genetic components of variation within and among populations in lieu of proper quantitative genetic data, it may still be possible to say something about the extent of geographical differentiation in quantitative traits when compared with what is expected from neutral geographic differentiation (Barrowclough, 1980; Rogers & Harpending, 1983; Prout & Barker, 1993; Spitze, 1993; Podolsky & Holtsford, 1995; Kremer *et al.*, 1997; Merilä, 1997; Smith *et al.*, 1997; Storz, 2002; Saint-Laurent *et al.*, 2003; Bernatchez, 2004; Roseman, 2004; Østbye *et al.*, 2005), provided that a sensitivity analysis is undertaken of the assumptions. We may call this the P_{ST}–F_{ST} approach (P symbolizing 'phenotypic'-Q_{ST}, or 'pseudo'-Q_{ST} if one so desires). Often, general patterns are very similar when comparing F_{ST} with Q_{ST} based either on phenotypic or genetic variance (Lynch *et al.*, 1999; Schluter, 2000). One improvement to the phenotypic estimation of Q_{ST} might be to derive conservative (minimum) estimates of Q_{ST}, by measuring the repeatability of suitable traits (e.g. annually re-grown traits) to obtain maximum estimates of the genetic within-population variance component. Note that both the P_{ST} approach and common garden Q_{ST} investigations are sensitive to deviations from purely additive gene action. Epistasis can potentially mask some effects of divergent selection (Whitlock, 1999), and nonadditive neutral gene action can sometimes result in patterns falsely suggesting selection, i.e. deviations from F_{ST} (López-Fanjul *et al.*, 2003). However, these problems are probably most relevant to traits, such as life history, rather than to morphological traits that typically show substantial additive genetic variance (Crnokrak & Roff, 1995; DeRose & Roff, 1999; Merilä & Crnokrak, 2001; López-Fanjul *et al.*, 2003).

In this study, we analyse geographical variation in microsatellites and morphological traits among distributional regions and among populations within these regions of the great snipe (*Gallinago media*), a migrating lekking shorebird. In western Europe (Scandinavia), the great snipe is a scarce inhabitant of earthworm-rich mountain fens around the tree line (Kålås *et al.*, 1997a), whereas in eastern Europe, the great snipe is patchily distributed predominately in lowland meadows subject to annual flooding (flood plains), eastwards of the Yenisey (Fig. 1; Gromadzka *et al.*, 1985; Tomkovich, 1992;

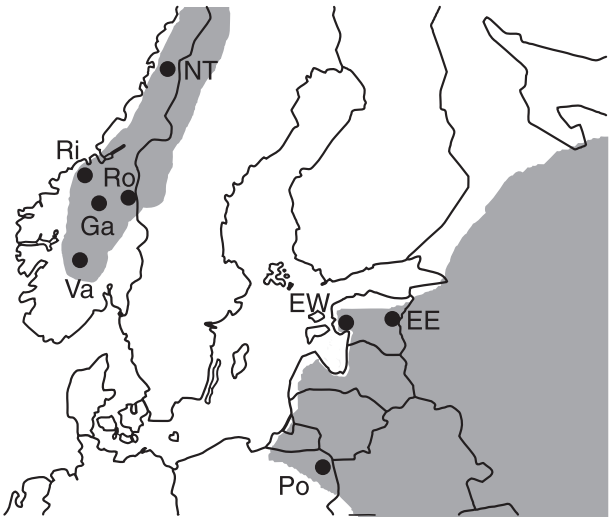

Fig. 1 Sampling locations for great snipe populations included in this study. Shaded areas roughly indicate the breeding distribution, but breeding populations occur very fragmented within these areas (especially in the westernmost part of the eastern region).

Kuresoo & Leibak, 1994). Hence, habitat differences may predict different selection regimes in the two disjunct regions, whereas differences among populations within the regions are more likely to be affected by drift and migration alone.

Previously, these two distributional regions were more closely, if not completely, connected. Up until the mid-19th century, great snipe were breeding over large parts of the lowlands also of western Europe (Germany, Denmark, southern part of the Scandinavian Peninsula and Finland), occupying a similar habitat as is still present in eastern Europe. The great snipe is now extinct in lowland western Europe, mainly because of the extensive man-made transformation of suitable habitat for agricultural purposes. The remaining western population (in the Scandinavian mountains) is estimated to be in the range of 10 000–30 000 males at present (Gjershaug *et al.*, 1994). Although there is little information about the size of the eastern population, it is clear that it has recently declined (Gromadzka *et al.*, 1985; Panchenko, 1985; Tomkovich, 1992; Kuresoo & Leibak, 1994). The species is currently classified as 'Near Threatened' at a global level (BirdLife International, 2000).

Here, we first analyse the patterns of genetic structure derived from microsatellite DNA. Second, we compare variation in morphological traits (body size measures and a secondary sexual trait) among great snipe populations, to the variation expected under neutrality. If gene flow is limited, we may expect adaptation to local conditions if selection is sufficiently strong. Such local adaptation may, on the other hand, be swamped by extensive migration between populations. If the geographical structure of quantitative traits were more pronounced than the geographical structure of neutral genetic markers, differential natural selection among populations, rather than drift alone, would have to be invoked to explain differences between populations.

Our approach was to compare F_{ST} with a pseudo-Q_{ST} measure derived from estimates of within- and between-population variance components of phenotypic traits. This approach is largely similar to those of Merilä (1997), Storz (2002) and Saint-Laurent *et al.* (2003). As quantitative trait data were purely phenotypic, assumptions about the additive genetic components of variance had to be made to be able to directly compare the magnitude of Q_{ST} and F_{ST} estimates, and we performed sensitivity analyses of these assumptions. For traits re-grown annually, we measured in one population the between-year repeatability to obtain a maximum estimate of the genetic proportion of the within-population variance component of the trait (Falconer & Mackay, 1996; Lynch & Walsh, 1998; but see Dohm, 2002 and Discussion) to derive a conservative (minimum) estimate of Q_{ST}. We took various steps to ensure that measurement error and measurer bias did not inflate our estimates.

Materials and methods

Collection of samples and morphological measurements

We caught great snipe with mist nets on leks distributed from Nordland, Norway, in the north to Biebrza, Poland in the south (Table 1, Fig. 1). Morphological measures of tarsus (true tibio tarsus length), total head (bill plus head), bill (to end of skin), bill to nostrils, wing length (maximum flattened) and tail white (the length of white on the outermost tail feather) were done according to Höglund *et al.* (1990b). A few 1-year-old birds had not yet moulted their juvenile tail feathers, and measurements of these feathers were excluded. All measurements were taken with digital callipers to the nearest 0.1 mm except wing length, which was measured to the nearest 1.0 mm using rulers with a riveted right angle stop. We used mean values of traits measured more than once for the same individual, and mean values of bilateral traits (tarsus, wing length and tail white), avoiding pseudoreplication and lowering the influence of measurement error.

All birds were measured by JAK except in Gåvålia (JAK 86%, PFI 8%, SAS 6%), Nord-Trøndelag/Nordland (PFI 100%) and Poland (SAS 100%). The noise introduced by differences between these observers was generally small, as inferred from a sample of birds measured independently by several persons in the same year (Table 2). Measurement error was low (generally 1–4%) and repeatability high both within and among observers (Table 2), apart from wing length, which was excluded from geographical analyses. However, slight systematic differences (Appendix A1) between observers could still potentially bias Q_{ST} estimates, in particular as two populations were not measured by JAK. These problems were overcome first, by the exclusion of wing length in geographical analyses, second by adjusting the

Table 1 Sampling locations for great snipe included in this study.

Population	Locality	Country	Geographic location	*n*	Year
NT	Namsskogan	Norway	64°53′N 13°12′E	13	1995
	Røyrvik	Norway	64°55′N 13°28′E		1995
	Hattfjelldal	Norway	65°40′N 13°48′E		1996
Ro	Røros	Norway	62°43′N 11°30′E	30	1996
Ga	Gåvålia	Norway	62°17′N 09°36′E	22	1993–1996
Ri	Rindal	Norway	63°09′N 09°17′E	12	1995–1996
Va	Valdres	Norway	60°53′N 09°01′E	18	1996
Po	Biebrza	Poland	53°31′N 22°38′E	12	1994
EE	Kärevere/Tartu	Estonia	58°25′N 26°31′E	47	1996, 1998
EW	Roude	Estonia	58°43′N 23°50′E	16	1998

Within each locality, birds were sampled at 1–10 leks. Sample sizes for microsatellite DNA are indicated (samples were larger for morphological traits). In the NT population, samples were pooled from the three locations.

Table 2 Repeatabilities and measurement errors of morphological traits of great snipe males measured on different dates in the same year in the Gåvålia population, either by the same person (JAK; within-observer repeatability), or by two or three persons (JAK, PF and SAS; between-observer repeatability) (all values of $P < 0.001$).

Trait	R*	F	d.f.	ME† (%)
Within observer				
Total head	0.961	50.21	46,47	3.91
Bill	0.897	18.37	46,47	10.33
Bill to nostrils	0.963	53.07	46,47	3.71
Mean tarsus	0.963	52.59	46,47	3.77
Mean wing	0.821	10.20	45,46	17.87
Mean tail white	0.990	197.35	47,48	1.01
PC1‡	0.975	80.18	46,47	2.47
Between observers				
Total head	0.963	56.08	39,45	3.71
Bill	0.940	34.29	40,46	5.99
Bill to nostrils	0.969	66.99	40,46	3.11
Mean tarsus	0.882	16.87	40,46	11.83
Mean wing	0.806	9.75	36,41	19.40
Mean tail white	0.964	58.37	42,48	3.55
PC1‡	0.992	261.08	39,45	0.79

*Repeatability.
†Measurement error.
‡First principal component, accounting for 95.27% of the variation in total head, bill and bill to nostrils in this sample.

tarsus measures of PFI to those of JAK by subtracting the mean difference of pairwise measurements between PFI and JAK (0.44 mm), and third by using a principal component measure of bill measures with high repeatability that did not show systematic difference between measurers. The first axis of a principal component analysis (conducted separately for males and females) of total head, bill and bill to nostrils captured 91.36% of the variation in these traits, which were all highly correlated. As great snipe males are much more likely to be caught on leks than are females, sample size for females was small in most populations, and females were hence excluded from most morphological analyses because there are systematic size differences between the sexes (Höglund *et al.*, 1990b).

As Q_{ST} and F_{ST} are attributes specific to the populations and cohorts compared, finding evidence for spatial structure might alternatively be an artefact of temporal variation if populations are not sampled at the same time. Our genetical data were collected over a short time span (Table 1). Morphological data were collected over a longer time span in one of the populations (GA, 1986–2003), rendering temporal effects more problematic. General linear mixed models revealed very low effects of year sampled in this population (data not shown), as well as of age (except for wing length), justifying pooling data from different years and ages. Age was estimated at first capture as either 1-year old or older using feather wear (Sæther *et al.*, 1994). Limiting the analyses to the same years as for genetical data or to measures of more than 1-year-old males had only minor effects on estimates of variance components (overall Q_{ST} values changed in the third or fourth decimal) and did not change any conclusions (data not shown). We therefore chose to include all years, and use mean values of individuals irrespective of age. We chose not to present detailed analyses of geographical variation in mean wing length as this trait showed large measurement error within season (both within and between observers), as well as age-related variation (and we did not have accurate age estimates for all birds in all populations to remove this effect).

Microsatellite genotyping

Five hypervariable tetranucleotide microsatellites were isolated using an enrichment protocol similar to that of Piertney & Dallas (1997) and Piertney *et al.* (1998). Individuals were genotyped at these loci (SNIPE B2, 3, B5, 12, 20; primers described in Appendix B1). The 10-μL PCR mixture contained approximately 10 ng of DNA, 1 μL of 10× buffer without $MgCl_2$ (MBI Fermentas, Ontario, Canada), 1 μL of 25 mM $MgCl_2$, 1 μL of 2.5 mM dNTPs, 0.5 μL of each of 10 μM forward and reverse primer, 0.25 units of Taq-polymerase (MBI Fermentas) and 5 μL of ddH_2O. PCRs were performed on a GeneAmp PCR System 9600 (Perkin Elmer, Waltham, MA, USA) using the following conditions: initial denaturation 94 °C for 3 min; c cycles of denaturation at 94 °C for 30 s, annealing at a °C for 30 s, extension at 72 °C for 40 s (c = number of cycles, a = annealing temperature); then a final extension step at 72 °C for 2 min. For locus SNIPE B2, $c = 30$ and $a = 51$; SNIPE 3 and SNIPE 20, $c = 32$, $a = 52$; SNIPE B5, $c = 31$, $a = 51$; SNIPE 12, $c = 30$, $a = 56$.

Aliquots (~3.5 μL) of the PCR products were separated on denaturing 6% polyacrylamide gels (Sambrook *et al.*, 1989). PCR products with shorter fragment sizes (140–200 bp; loci B2, 3 and 12) were run on the gels for at least 1 h, whereas PCR products with longer fragment sizes (300–350 bp, loci B5 and 20) were run for at least 1.5 h. After electrophoresis, the PCR products were visualized by silver staining (Sambrook *et al.*, 1989). Individuals were assigned genotypes by comparison with a standard set of samples of known allele size. The microsatellite sequences obtained in this study are deposited at GenBank under the accession numbers, AY363298–AY363302.

Data analyses

Microsatellite DNA

None of the females was heterozygous for the microsatellite loci, SNIPE 3 and SNIPE 12. These are therefore probably located at the Z chromosome and data from females for these loci were excluded from the following analyses. Each locus in each population was tested for deviations from Hardy–Weinberg equilibrium, and the

probability of deviations from Hardy–Weinberg equilibrium for all loci combined in each population was calculated according to Fisher's method for combining probabilities (Sokal & Rohlf, 1981). The presence of linkage disequilibrium was also tested for each pair of loci in each population. These tests were done using Genepop on the Web (http://genepop.curtin.edu.au; Raymond & Rousset, 1995). Each locus was also tested for the proportion of multistep mutations vs. single-step mutations with the program MISAT (Nielsen, 1997).

To check for evidence of recent bottlenecks we used the program Bottleneck 1.2.02 (Cornuet & Luikart, 1996). We chose to use a Wilcoxon test under the assumptions that all loci fit the stepwise mutation model, or that all loci fit a two-phased mutation model with the proportion of multistep mutations found by the program MISAT.

The genetic structuring of populations was examined by a hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) computed with Arlequin 1.1 (Schneider *et al.*, 1997). Variance was partitioned between eastern (Estonian and Polish populations) and western (Norwegian) populations, between populations nested within these two groups, and among individuals within populations. Pairwise population differentiation was calculated based on F_{ST} (Weir & Cockerham, 1984) using Genetix 4.05 (Belkhir *et al.*, 1996). We also estimated R_{ST} (Slatkin, 1995), and did analyses using both F_{ST} and R_{ST} to ensure that conclusions did not depend on the choice of differentiation statistic. R_{ST} is a measure of genetic differentiation based on the stepwise mutation model, and is often more appropriate for microsatellites as differentiation might be underestimated by F_{ST} if mutations create allelic homoplasy and mutation rate is high relative to the migration rate. However, if mutation rates are low relative to migration rates F_{ST} can be expected to provide more accurate estimates of genetic differentiation than R_{ST} (Slatkin, 1995). As an estimator of R_{ST} we used Goodman's unbiased ρ. R_{ST} was calculated using RstCalc 2.2 (Goodman, 1997) after standardizing allele sizes to a global mean of zero and unit standard deviation, and after averaging variance components over loci. *P*-values of global R_{ST} estimates over all populations or over regions were obtained by permutation tests, and approximate 95% confidence intervals by the range of the central 95% of 1000 bootstrap estimates.

Note that differentiation estimates below zero are most likely because of sampling variation (and not because alleles from different populations actually are more similar to each other than the alleles within the same population). The best estimate for negative values would therefore be zero. We did not adjust negative pairwise estimates to zero, as this would have created a bias when sampling variation of positive estimates is not similarly adjusted.

Isolation by distance was tested with Mantel tests assuming a linear relationship between pairwise values of $F_{ST}/(1 - F_{ST})$ and the natural logarithm of geographic distances (km) between all population pairs (Rousset, 1997). Geographic distances were calculated following the Earth's curvature, using the GeoDistances module in R 4.0 (Casgrain & Legendre, 2001).

Quantitative traits

Repeatability [$\mathrm{var}_{between}/(\mathrm{var}_{between}+\mathrm{var}_{error})$], measurement error (1 − repeatability) and pairwise Q_{ST} values were estimated using the VARCOMP procedure in SPSS 11, applying the ANOVA (type III sum of squares) approach. Maximum likelihood-based estimates were very similar (data not shown). Repeatability as a maximum estimate of heritability is also reported corrected for the separately estimated within-observer measurement error (Lynch & Walsh, 1998) as

$$\frac{\mathrm{repeatability}_{between-year}}{\mathrm{repeatability}_{within-observer}}.$$

Q_{ST} was estimated as

$$\frac{(g)\mathrm{var}_{population}}{(g)\mathrm{var}_{population} + 2(h^2)\mathrm{var}_{error}}$$

where g is the assumed additive genetic proportion of differences between populations, h^2 (narrow-sense heritability) is the assumed additive genetic proportion of differences between individuals within populations, $\mathrm{var}_{population}$ is the observed between-population variance component and var_{error} is the observed within-population variance component. A sensitivity analysis was performed, simulating different values of g and h^2 (including the corrected between-year repeatability of tail white), but the pairwise Q_{ST} values used are those obtained assuming $g = 1$ and $h^2 = 0.5$, unless otherwise stated. The advantage of using these particular assumptions is that significance testing of the estimate (or rather the between-population component) can then be conducted by standard methods of analysis of variance.

Nested analyses, partitioning the variation in morphology among regions and among populations within regions, were performed using procedure GLM in SPSS 11 and associated variance components estimated with procedure VARCOMP. Simple and partial matrix permutation tests were performed using R 4.0 (Casgrain & Legendre, 2001). Statistical analysis of whether pairwise Q_{ST} values involving populations in different regions were larger than those within regions were conducted by calculating the standardized Mantel statistic (r_M) between a distance matrix **A** of pairwise Q_{ST} values and a matrix **B** of kind of comparison (within or between regions). Matrix permutation on **A** 10 000 times was then used to obtain a randomization *P*-value of the null hypothesis of no difference. To test if among-region Q_{ST} values were larger than within-region values when controlling for neutral genetic variation, a distance matrix **C** of pairwise F_{ST} or R_{ST} values was also constructed and the partial r_M (Smouse *et al.*, 1986)

computed between **A** and **B** while controlling for **C**. A randomization *P*-value was obtained by matrix permutation of **A**, holding **B** and **C** constant. Matrix permutation of **B** instead of **A** while controlling for **C** gave the same conclusions (results not shown).

Results

Microsatellite DNA

The number of alleles at the five loci ranged from two to 19 (SNIPE B2 = 13 alleles; SNIPE 3 = 10 alleles; SNIPE B5 = 19 alleles; SNIPE 12 = 10 alleles; SNIPE 20 = 2 alleles). The mean number of alleles per locus was between 5.6 and 9.0 in the populations (Appendix B2). In the western Estonian population locus SNIPE 20 was monomorphic, whereas all other populations were polymorphic for all loci. We found statistically significant deviations (heterozygote deficiencies) from Hardy–Weinberg equilibrium for one or more loci in five of the eight examined populations (data not shown). After combining probabilities for all loci in each population, four populations (eastern Estonia, Rindal, Gåvålia and Røros) showed evidence of deviations from Hardy–Weinberg equilibrium ($P < 0.05$). However, none of these remain significant after Bonferroni adjustment of the α level for the number of populations. There was no consistent linkage disequilibrium between pairs of loci across populations.

The maximum likelihood tests for proportion of multistep mutations (pmm) vs. single-step mutations indicated a very low proportion of multistep mutations (pmm ≤ 0.01 for all loci), indicating that the use of *R*-statistics is appropriate (Nielsen, 1997). There was no evidence of a recent bottleneck in any population (data not shown) although the power of these tests is low with only five loci. The observed heterozygosity did not deviate from what could be expected under a strict stepwise mutation model, or a two-phased mutation model assuming that 5% of the variation in allele size is attributable to an infinite allele model, and 95% to a stepwise mutation model.

An analysis of molecular variance confirmed that great snipe are weakly structured into one western and one eastern group. There seems to be no variation among populations within these groups, and most of the variance is accounted for within populations (Table 3). Pairwise R_{ST} and F_{ST} estimates indicate low, but significant, population differentiation between the eastern Estonian population and most Norwegian populations (Table 4). There was a strong correlation between pairwise F_{ST} and R_{ST} estimates (Mantel test, $r_M = 0.859$, $P < 0.001$).

The global estimates of divergence over all populations were (±bootstrap 95% confidence limits): $R_{ST} = 0.059$ (±0.041, $P < 0.001$) and $F_{ST} = 0.026$ (±0.026, $P = 0.007$). The estimates over regions were: $R_{ST} = 0.051$ (±0.037, $P < 0.001$) and $F_{ST} = 0.018$ (±0.022, $P = 0.012$). The higher estimates of R_{ST} than of F_{ST} may indicate that stepwise-like mutations rather than drift alone have contributed to the differentiation.

Table 4 Population pairwise R_{ST} estimates from microsatellite variation (above diagonal) and pairwise estimates of Weir–Cockerham F_{ST} (below diagonal).

	Eastern region			Western region				
	Po	EE	EW	Ri	Ga	NT	Ro	Va
Po		**0.096**	−0.036	*0.029*	***0.162***	***0.082***	***0.068***	*0.058*
EE	**0.055**		**0.095**	***0.086***	*0.032*	***0.104***	***0.092***	***0.084***
EW	−0.016	**0.052**		*0.092*	***0.192***	*0.131*	***0.109***	*0.102*
Ri	*0.027*	***0.049***	*0.033*		0.052	−0.005	0.003	0.009
Ga	***0.090***	*0.011*	***0.104***	0.040		0.063	**0.073**	**0.107**
NT	*0.023*	***0.026***	*0.032*	0.000	0.019		−0.031	0.003
Ro	*0.009*	***0.031***	***0.024***	0.005	0.021	−0.018		−0.002
Va	*0.025*	***0.040***	*0.007*	0.007	0.040	−0.018	−0.018	

Significant values are highlighted in bold, and comparisons involving populations in different regions are highlighted in italics. Population abbreviations as in Table 1.

Table 3 Hierarchical analysis of molecular variance (AMOVA) for eight great snipe populations categorized into two regions (Norway and Estonia/Poland) (a) based on weighted average *R* over five microsatellite loci, (b) based on microsatellite allele frequencies.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	*P*-value	R_{ST}
(a)						
Among regions	1	51.72	0.263	2.41	0.053	0.024
Among populations within regions	6	51.81	−0.055	−0.5	0.46	−0.005
Within populations	332	3559.05	10.72	97.51		
Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	*P*-value	F_{ST}
(b)						
Among regions	1	2.655	0.017	1.72	0.15	0.017
Among populations within regions	6	0.73	−0.023	−2.33	0.99	−0.023
Within populations	332	335.69	1.01	100.6		

Isolation by distance

We found some indication of a linear isolation-by-distance pattern when comparing population pairwise genetic and geographic distances. The pattern using Goodman's unbiased ρ as an estimator of R_{ST} (Mantel $r = 0.356$, $P = 0.048$) is shown in Fig. 2a. The Weir–Cockerham F_{ST} estimator indicated a somewhat weaker pattern (Mantel $r = 0.224$, $P = 0.17$; Fig. 2b). This isolation-by-distance effect, although small, may explain the small differentiation between regions, but it is hard to analyse whether there is an additional effect of region as there was no overlap in geographic distance.

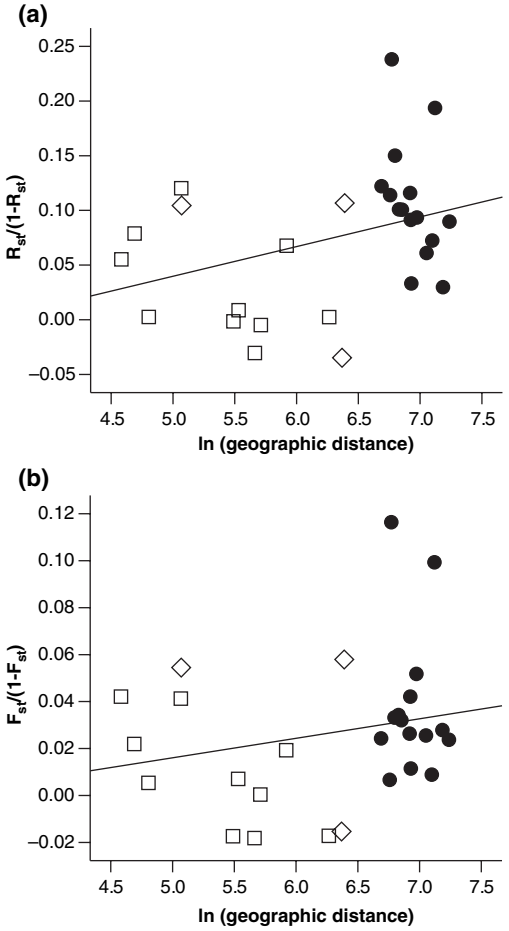

Fig. 2 Neutral genetic divergence increases somewhat with geographic distance, as shown in these isolation-by-distance plots using either R_{ST} (a) or F_{ST} (b) against the natural logarithm of distance in kilometres along the curvature of the Earth. Comparisons of populations in different regions are indicated by solid symbols, and comparisons of populations within the same region are indicated by open symbols (diamonds in the east and squares in the west). Also shown is the least squares linear regression.

Quantitative traits

Annually re-grown traits (tail white and wing length) showed moderate to high repeatability between years, indicating substantial heritability of these traits (Table 5). After correcting for measurement error, repeatability estimates were of similar magnitude (~0.8) for both traits, although we found somewhat lower values for tail white in females. Limiting the analysis to adults (to remove potential noise introduced by age-related variation) only marginally increased repeatability of tail white, but increased the repeatability of wing length to 0.95 (Table 5).

Populations nested within region had only little influence on trait values, but different traits showed striking variation in the degree of differentiation among regions (Appendix C1). Tarsus length and amount of white on tail showed very strong divergence among regions (Appendix C), whereas a composite measure of bill length (PC1) showed only very weak (but significant) differentiation. The divergence in tail white and tarsus appeared to be entirely independent of each other. Although there was a slight overall correlation between the two measures ($r = 0.102$, $P < 0.001$, $n = 1502$ males) this was an artefact of both traits differing among regions: within region, there was no correlation (Fig. 3).

Pairwise Q_{ST} estimates (Appendix D1) between populations show that comparisons involving populations in different regions often had large (and significant) Q_{ST} values of tail white and tarsus, whereas comparisons within regions were small (often negative) and nonsignificant. No pairwise Q_{ST} values for bill were significant. Matrix permutation tests confirmed that Q_{ST} values were significantly larger between regions than within regions both for tail white ($r_M = 0.788$, $P = 0.018$) and tarsus

Table 5 Between-year repeatabilities of annually re-grown traits measured on the same individuals in two or more years irrespective of age, and in two or more years as adults only.

Trait	R*	R'†	F	d.f.
(Males)				
Tail white	0.823	0.832	13.53	382, 647
Wing length	0.649	0.790	6.01	296, 510
(Females)				
Tail white	0.720	0.728	6.83	93, 119
Wing length	0.670	0.816	5.59	93, 119
(Adult males)				
Tail white	0.829	0.837	13.78	266, 438
Wing length	0.776	0.945	10.22	194, 323
(Adult females)				
Tail white	0.689	0.696	5.98	67, 85
Wing length	0.729	0.888	7.61	66, 84

*Repeatability.
†Repeatability adjusted for within-observer measurement error, 1.009% for tail white and 17.868% for wing length.
All $P < 0.001$.

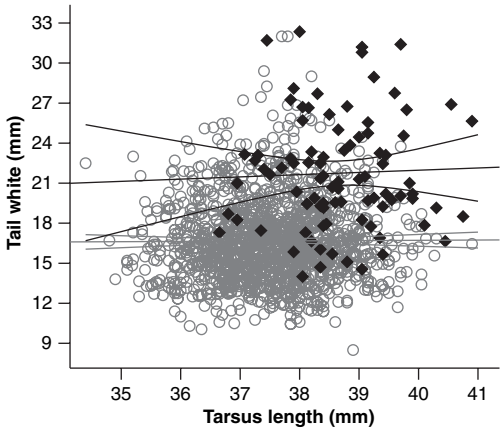

Fig. 3 Both tarsus length and tail white is larger in the eastern (black diamonds) than in the western region (open grey circles). The overall correlation between the two traits is an artefact of these differences. Within regions there is no relationship (least squares regression lines and 95% confidence limits shown, east $r^2 = 0.001$, $n = 88$; west $r^2 < 0.001$, $n = 1416$).

($r_M = 0.736$, $P = 0.018$), but not for bill ($r_M = 0.158$, $P = 0.218$).

Q_{ST}–F_{ST} comparisons

Partial matrix permutations of pairwise values showed that, even when controlling for the neutral genetic structure (or geographic distance), Q_{ST} for both tail white and tarsus (but not bill) were higher for comparisons between populations in different regions than comparisons within regions (Table 6; Fig. 4). These results should be very robust to varying the assumptions of the proportion of additive genetic variance, because the relative difference between pairwise Q_{ST} values in a matrix does not change much by varying the g and h^2 parameters of that matrix. This was confirmed by partial matrix permutation tests using simulated Q_{ST} matrices for

Table 6 Partial matrix permutation tests of the relationship between pairwise differentiation in quantitative traits (**A**) and the kind of comparison (**B**, populations within or among regions), while controlling for pairwise differentiation in neutral genes or geographic distance (**C**).

Matrix **A**	Matrix **C**	r_M (**AB·C**)*	*P*-value
$Q_{ST(tail\ white)}$	F_{ST}	0.749	0.015
$Q_{ST(tail\ white)}$	R_{ST}	0.702	0.015
$Q_{ST(tail\ white)}$	ln(distance)	0.672	0.003
$Q_{ST(tarsus)}$	F_{ST}	0.691	0.013
$Q_{ST(tarsus)}$	R_{ST}	0.685	0.006
$Q_{ST(tarsus)}$	ln(distance)	0.347	0.062
$Q_{ST(PC1)}$	F_{ST}	0.176	0.130
$Q_{ST(PC1)}$	R_{ST}	0.078	0.311
$Q_{ST(PC1)}$	ln(distance)	−0.103	0.312

*Partial Mantel statistic.

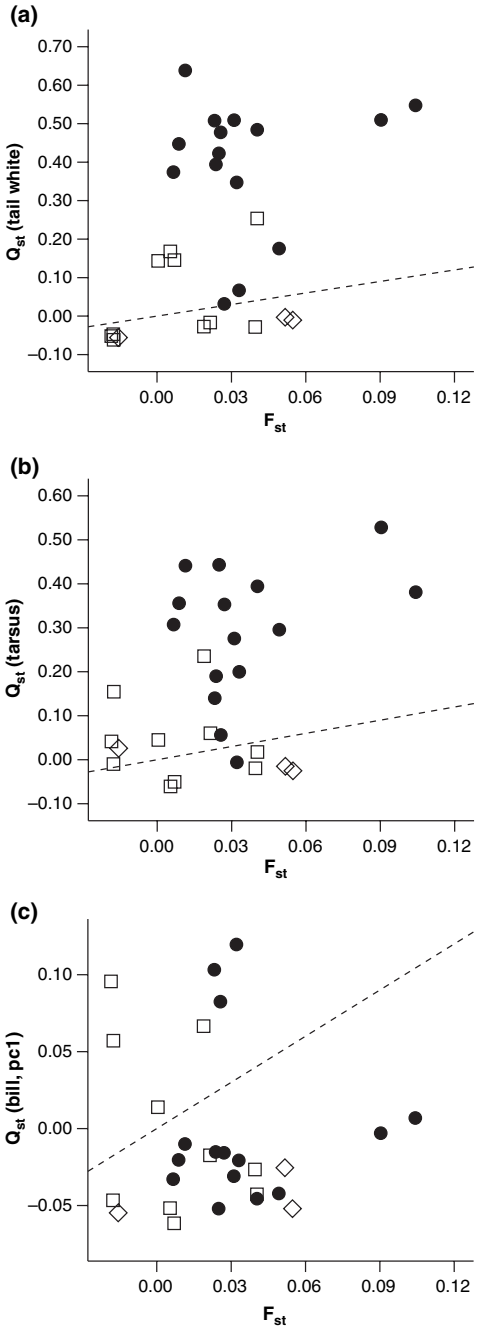

Fig. 4 Population differentiation in quantitative traits (Q_{ST}, assuming $g = 1$ and $h^2 = 0.5$) in relation to neutral genetic differentiation (F_{ST}). For both tail white (a) and tarsus (b), comparisons among populations in different regions (solid symbols) are larger than expected from the neutral differentiation whereas this is not the case for comparisons within the regions (open symbols, diamonds in east and squares in west). A different pattern is found for bill length (PC1) where Q_{ST} values among regions do not differ from within regions and are not larger than expected from neutral variation (c). Note the different scales on the Q_{ST} axes. Dashed lines are expectations if $Q_{ST} = F_{ST}$.

a range of g and h^2, and the observed R_{ST} and F_{ST} matrices (data not shown).

The overall divergence between regions in bill (Q_{ST} = 0.01) was lower than, or similar to, the values of F_{ST} and R_{ST} (0.018 and 0.051, respectively, see above), whereas tail white and tarsus showed much stronger differentiation (Q_{ST} = 0.568 and 0.416 respectively). The problematic wing length measure showed no indication of a divergence deviating from neutral expectations (Q_{ST} = 0.043).

We recalculated Q_{ST} values for different assumptions about heritability (h^2, 0.25, 0.5, corrected between-year repeatability of tail white, and 1.0) and the magnitude of the additive genetic proportion of the between-population variance component (g, 0.05–1.0). This exercise showed that the conclusions are not sensitive to varying g and h^2 even outside realistic parameter space (Fig. 5). Exceptionally small additive genetic proportion of the between-population variance have to be invoked to arrive at Q_{ST} values comparable with neutral markers for tarsus (Fig. 5b) and in particular tail white (Fig. 5a). For the composite measure of bill length most simulations yielded lower Q_{ST} values than expected from microsatellite differentiation, but these were often within the bootstrapped 95% confidence intervals of the neutral differentiation (Fig. 5c).

For females, overall Q_{ST} between regions (n = 513–523 in western region and 16 in eastern region) resembled male values ($Q_{ST(tail\ white)} = 0.545$, $Q_{ST(tarsus)} = 0.454$, $Q_{ST(pc1\ bill)} = -0.007$), but low sample size in most populations precluded calculation of pairwise values.

Discussion

This study has highlighted that neutral genetic differentiation is not sufficient to explain geographic differentiation in some quantitative traits in great snipe, and suggests local adaptation to different habitats despite high gene flow. The two habitats coincide with distributional regions, but the trait divergence cannot be explained as an artefact of isolation by distance.

Assumptions

Any comparison of Q_{ST} and F_{ST} to infer spatial variation in adaptations requires that these two measures are comparable and unbiased. Different ways of calculating subdivision for microsatellites (R_{ST} and Weir–Cockerham F_{ST}) gave similar results. Critical assumptions behind the Q_{ST} estimates were investigated to ensure that conclusions did not depend on inflated values because of uncertainties about heritability and additive genetic proportion of differences between populations. Sensitivity analyses revealed that conclusions were very robust to variation in these parameters. It thus appears that a quantitative genetic common garden rearing scheme of great snipe to arrive at these conclusions would have

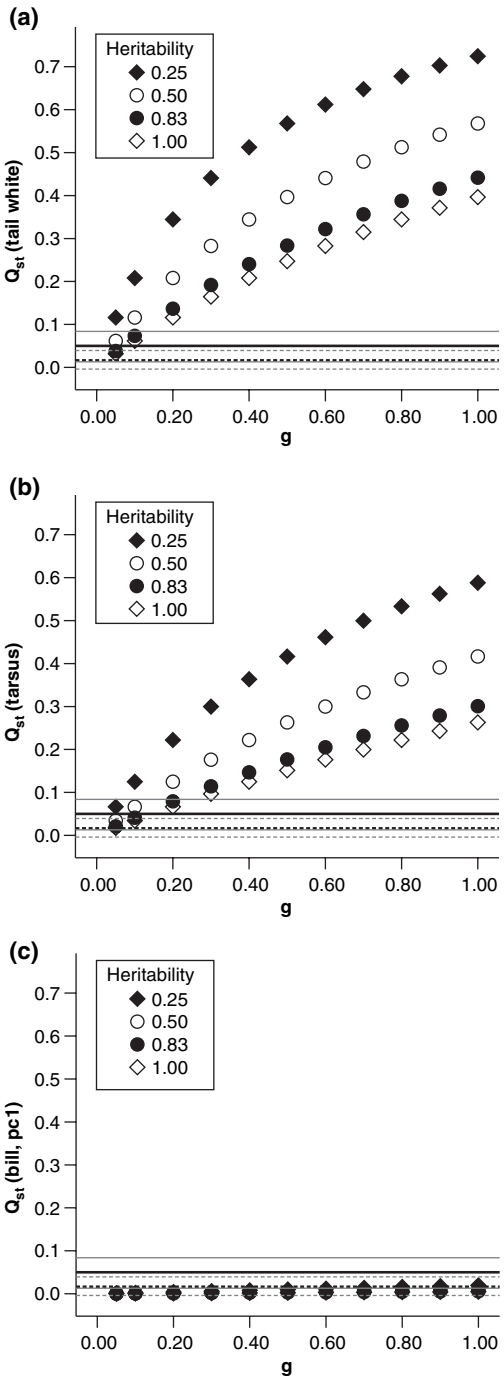

Fig. 5 Q_{ST} sensitivity plots of varying the additive genetic proportion of between-population (g) and within-population (heritability) variance components. Q_{ST} values are calculated treating the two regions as two populations. Estimates of neutral divergence is shown as horizontal lines ±95% bootstrap confidence limits (bold solid line for R_{ST}; dashed line for Weir–Cockerham F_{ST}). Q_{ST} for (a) tail white and (b) tarsus is considerably larger than neutral genetic divergence for most parameter space, whereas Q_{ST} for (c) bill (PC1) is similar or smaller.

been unnecessary. Moreover, measurement error was in general low and would not be expected to cause overestimation of Q_{ST} values. The same person took most measures, but the deviating values of tarsus length in the NT population measured by another person suggest that our steps to eliminate observer bias were not fully adequate for this trait in this population (as indicated by the high within-region and low between-region pairwise Q_{ST} values involving this population, Appendix D1b). For the other traits, and for all other populations, there were no detectable systematic between-observer effects that could bias results. Excluding measures of tarsus in the NT population would have increased between-region Q_{ST} estimates, but for the sake of being conservative we chose to keep it. It is important for any Q_{ST} study to estimate measurement error (both within and among observers), and take steps to ensure that conclusions are not biased, e.g. because different persons measured traits in different populations. We urge all future studies of Q_{ST} to include repeated independent measures of a subsample for this purpose.

Pairwise comparisons – avoiding some assumptions

By comparing pairwise Q_{ST} values among populations either in the same or in different regions, and controlling for the neutral divergence among the same populations using partial matrix permutations, we could avoid relying on specific assumptions about g and h^2 for our conclusions to hold. This is because – even if the absolute values of Q_{ST} may, e.g. be overestimated – the difference in relative magnitude between pairwise comparisons either within or among the units of interest (regions, in our case) is less affected by these assumptions. We suggest that adopting this pairwise approach, together with a sensitivity analysis of varying g and h^2 for the overall Q_{ST}–F_{ST} comparison, allows for robust conclusions to be drawn using purely phenotypic data in lieu of common garden experiments. This highlights the advantage of sampling several subpopulations to do a more reliable analysis of local adaptation than only obtaining an overall estimate of Q_{ST}.

Neutral differentiation

We found that great snipe populations are weakly structured across northern Europe. Microsatellite DNA markers detected a genetic division between western and eastern populations (Norwegian and Estonian/Polish samples respectively). This weak neutral genetic differentiation between the two regions might simply be an isolation-by-distance effect. Although it is possible that there is also an effect of region (see Fig. 2), it would be very hard to say, based on contemporary genetic variation, whether such an effect is because of the recent separation or because of more ancient restriction of gene flow between populations in different habitats. The recent fragmentation of the distribution – which is probably mostly because of habitat changes in the lowlands induced by humans during the 19th century (Kålås *et al.*, 1997a) – has geographically separated the remaining lowland populations from the western mountain populations, but we cannot detect any genetic signature of this separation. In view of that, analyses of museum specimens from the now-extinct populations in lowland western Europe would be interesting.

Quantitative trait differentiation

Different traits showed different patterns of Q_{ST} compared with F_{ST}. Birds from the eastern and western regions differed substantially more in both tail white and tarsus than expected from neutral genetic differentiation, but did not in bill length (see also Kålås *et al.*, 1997b for morphological variation).

Divergent selection

Eastern birds had whiter tails than western birds. The amount of white in the tail has probably been subjected to sexual selection (Höglund *et al.*, 1990a; Sæther *et al.*, 2000). As birds in Poland and Estonia display at lower latitudes and earlier in the season (J.A. Kålås, S.A. Sæther, A. Kuresoo, L. Luigujoe, unpublished data), birds from these localities perform their displays under considerably darker light conditions. Hence, more extensive white in the eastern populations might be because of requirements of a more conspicuous signal there. It is thus possible that the difference in tail white between the regions represents a local adaptation to light conditions (in all populations males display during the night). Furthermore, it is more likely that western populations have evolved less white tails, rather than that eastern populations have become more white. This is because great snipe probably must have colonized Scandinavia from the south (-east), rather than vice versa, after the last glaciation. If so, our results suggest that it is the cost of maintaining extensive white in the western populations (because of, e.g. predation, see Höglund *et al.*, 1992) – rather than the benefit of more white tails in the eastern populations – that has shifted the trade-off balance and is the ultimate cause of the differentiation, but further studies are needed to confirm this. Interestingly, also females had more white tails in the eastern populations.

Tarsus length also differed more than expected from neutral markers between regions, and showed little variation between populations within the regions. This could perhaps be because of the habitat differences between the western and eastern populations. In the east, great snipe occur largely in sites subject to annual flooding early in the breeding season (when males display at leks) and we may speculate that this has led to natural selection for longer legs than in the mountain populations. Unfortunately, we did not have samples

from large parts of the very eastern distribution of great snipe. In particular, it would be interesting to compare the Scandinavian mountain populations with northern Russian ones that occur in similar habitat (Morozov, 1994). Our prediction is that those birds, despite being located further away, should have trait values more similar to Scandinavian than to Polish and Estonian birds.

Detecting uniform selection: limitations of the approach

The weak divergence among regions in pc1 (bill length) corresponded to a pattern expected from neutral differentiation, or was possibly lower. Optimal bill length in great snipe is likely to be influenced by the depth at which earthworms (their main food) occur, and the birds prefer habitat with a suitable balance between easier soil penetrability (wetter areas) and earthworms occurring closer to the surface (drier areas) (Løfaldli *et al.*, 1992). Perhaps earthworms are sufficiently available at the same soil depth in the two regions to prevent divergence, or perhaps conditions at overwintering grounds in Africa are more important. We cannot exclude that there is in fact similar stabilizing selection on bill length in both regions. Given the low values of genetic population differentiation in this study, it would be very hard to statistically document $Q_{ST} < F_{ST}$ for these populations to convincingly show stabilizing selection on any trait across environments, although the overall region analyses suggest so (Fig. 5c). Given also the unknown magnitude of g and h^2, and the difficulties involved in calculating standard errors for ratios, this illustrates one important limitation of the P_{ST} vs. F_{ST} method: by using this approach it would often be much harder to find evidence for uniform stabilizing selection than to find evidence for differential selection.

However, this shortcoming is also often shared by Q_{ST}–F_{ST} comparisons involving common garden experiments. It is instructive to note that the few documented cases of $Q_{ST} < F_{ST}$ (Merilä & Crnokrak, 2001; McKay & Latta, 2002; Edmands & Harrison, 2003) often show large population differentiation in neutral genes. Also, as Hendry (2002) pointed out, when F_{ST} is approaching unity, it will be hard to show that Q_{ST} is even larger. A related problem is that the maximum value of F_{ST} is in practice often less than unity (because of mutation), and that Q_{ST} may be less constrained from reaching its maximum value under neutrality. Relative measures of between-population divergence, such as F_{ST}, is heavily affected by the within-population diversity and may therefore be poor measures of divergence for loci with high diversity such as microsatellites (Charlesworth, 1998; Hedrick, 1999). Any factor affecting the difference in within-population diversity, such as different degrees of inbreeding or demographic histories of bottlenecks, could therefore potentially result in different values of F_{ST} even if the absolute levels of divergence are similar (Charlesworth *et al.*, 1997), and it seems likely that Q_{ST} estimates will not be affected in a similar way. These problems must be traded against the straightforwardness of comparing dimensionless estimates of divergence at marker loci and quantitative traits.

The Q_{ST}–F_{ST} approach may therefore be most useful: (a) for providing indirect evidence of divergent, rather than uniform selection; and (b) in situations with low-to-moderate neutral subdivision (because of gene flow or relatively recent isolation), rather than for populations separated a very long time ago.

Repeatability as maximum heritability in P_{ST} analyses

Without additive genetic variance there can be no evolutionary response to selection on a trait, hence Q_{ST} may remain low despite different selection 'pressures' being present. It may therefore be important to establish if a trait under study is heritable. If analysed carefully, repeatability of a trait suitable for such an analysis might be considered an upper limit on how much additive genetic variation is present for the trait within a population (Falconer & Mackay, 1996; Lynch & Walsh, 1998). It is well acknowledged that the true narrow-sense heritability may be substantially lower than this upper limit, but using repeatability as an estimate of heritability in the calculations of Q_{ST} may provide conservative (low) estimates of Q_{ST} (i.e. conservative in the context of showing $Q_{ST} > F_{ST}$). However, it is important that the repeatability estimate is not downward biased for this reasoning to hold true. The worry is that repeated measurements of a trait might not be comparable, and that measurement error will deflate estimates. These two problems, and a solution, may be illustrated by our measures of mean wing length.

At first sight, wing length may appear to show low repeatability compared with tail white. However, after correcting for the substantial measurement error, wing length shows very similar between-year repeatability to tail white (Table 5). (The relatively large measurement error of wing length is probably not only because of low accuracy of measurements, but may also be an effect of wing feather wear during the breeding season, Sæther *et al.*, 1994.) Moreover, unlike tail feathers, 1-year-old great snipe have not yet moulted their wing feathers (Sæther *et al.*, 1994), whereas adult birds have fresh wings that are on average longer. Hence, after further restricting the analyses to known adult males (Table 5) the among-year repeatability estimate turns out to be extremely high (0.94), suggesting a potentially large genetic component of the wing length variation among adult individuals. This highlights the importance of using comparable measurements when calculating repeatability and to take measurement error into account. If not, repeatability might in fact potentially underestimate heritability (Widemo & Sæther, 1999; Dohm, 2002) and thus overestimate Q_{ST} instead of providing a conservative estimate.

General patterns of Q_{ST} vs. F_{ST} in natural populations

In our study, different quantitative traits showed very different patterns of divergence. There is no reason, in general, to expect traits in linkage equilibrium to show a correlation of Q_{ST} values beyond what is expected from the neutral genetic differentiation. Q_{ST} is therefore an attribute of the particular trait, and cannot be said to characterize the populations as such, unlike F_{ST} ideally would. The pattern observed of Q_{ST} usually exceeding F_{ST} in published studies (Merilä & Crnokrak, 2001; McKay & Latta, 2002) is therefore heavily dependent on the particular traits that happen to have been studied. This pattern is likely to be affected by biases toward traits showing high Q_{ST} because spatial variation might have motivated the study in the first place, and biases against publishing low Q_{ST} values because such traits might be deemed uninteresting or because of the statistical difficulties of rejecting the null model when F_{ST} is low. Thus, it is not easy to say anything in general from studies comparing F_{ST} and Q_{ST} about whether natural selection has a predominately diversifying or homogenizing effect on metapopulations, apart from the fact that both occur. A more fruitful approach might be to compare patterns emerging from different kinds of traits, such as those involved in premating isolation vs. other traits (e.g. Butlin & Tregenza, 1998). Our study indicates that both a sexual signal (tail white) and some, but not all, morphological traits show larger differentiation than expected from neutral loci.

Implications for conservation genetics

Our results may have practical implications for conservation biology. Although great snipe populations are very weakly differentiated at neutral loci, adaptive genetic differentiation (as measured by P_{ST} here and also in MHC divergence by Ekblom *et al*., 2007) makes it clear that the eastern and western regions might need to be treated as separate conservation units. Such units are often defined using divergence in neutral markers alone, but our results support the view that this approach risks failing to identify ecologically important genetic differences among populations (e.g. Karhu *et al*., 1996; Butlin & Tregenza, 1998; Hedrick, 1999; Crandall *et al*., 2000; Fraser & Bernatchez, 2001; Pearman, 2001; Reed & Frankham, 2001; McKay & Latta, 2002; Stockwell *et al*., 2003; Hansson & Richardson, 2005). It is not at all clear whether one can predict ecologically or evolutionary important differences among natural populations from neutral divergence, or indeed if there is a general correlation between F_{ST} and Q_{ST} (Merilä & Crnokrak, 2001; Reed & Frankham, 2001; Crnokrak & Merilä, 2002; Latta & McKay, 2002; McKay & Latta, 2002; Ekblom *et al*., 2007). Ultimately, the maintenance of ecologically meaningful and adaptively significant genetic diversity should be the primary goal in conservation genetics, and not the maintenance of neutral variation. Although not a cure-all, adopting the P_{ST}–F_{ST} approach may open up avenues for putting the tools of neutral genetic variation into their proper organismal context in a whole new set of natural populations and species, which are otherwise unavailable to quantitative genetic analysis. Many organisms of conservation concern presumably fall into this category, and it may sometimes be important to go beyond neutral genetic variation because adaptive population divergence may have evolved in the face of gene flow.

Acknowledgments

Thanks to Sten Svartaas, Henrik Brøseth and Marek Borkowski for help with the fieldwork, Gunilla Olsson and Reija Dufva for assistance in the laboratory, and to Martin Carlsson, John Dallas, Johan Dannewitz and two reviewers for helpful comments. Financial support was received from the University of Trondheim and the Research Council of Norway (to SAS and PF), the Norwegian Institute for Nature Research (to JAK) and the Swedish Research Council (to JH).

References

Avise, J.C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Avise, J.C. 2000. *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise, J.C. & Ball, R.M. Jr. 1991. Mitochondrial DNA and avian microevolution. *Proc. Int. Ornithol. Congr.* **20**: 514–524.

Barrowclough, G.F. 1980. Genetic and phenotypic differentiation in a wood warbler (genus *Dendroica*) hybrid zone. *Auk* **97**: 655–668.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. 1996. *GENETIX 4.05, Logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bernatchez, L. 2004. Ecological theory of adaptive radiation: an empirical assessment from Coregonine fishes (Salmoniformes). In: *Evolution Illuminated: Salmon and their Relatives* (A. P. Hendry & S. C. Stearns, eds), pp. 175–207. Oxford University Press, Oxford.

BirdLife International 2000. *Threatened Birds of the Word*. BirdLife International, Cambridge, UK.

Butlin, R.K. & Tregenza, T. 1998. Levels of genetic polymorphism: marker loci versus quantitative traits. *Philos. Trans. R. Soc. Lond. B* **353**: 187–198.

Casgrain, P. & Legendre, P. 2001. *The R Package for Multivariate and Spatial Analysis*, version 4.0 d6 – User's Manual. Département de sciences biologiques, Université de Montréal [www document]. URL http://www.bio.umontreal.ca/Casgrain/en/labo/R/.

Charlesworth, B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* **15**: 538–543.

Charlesworth, B., Nordborg, M. & Charlesworth, D. 1997. The effects of local selection, balanced polymorphism and

background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* **70**: 155–174.
Cornuet, J.M. & Luikart, G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.
Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M. & Wayne, R.K. 2000. Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* **15**: 290–295.
Crnokrak, P. & Merilä, J. 2002. Genetic population divergence: markers and traits. *Trends Ecol. Evol.* **17**: 501.
Crnokrak, P. & Roff, D.A. 1995. Dominance variance: associations with selection and fitness. *Heredity* **75**: 530–540.
DeRose, M.A. & Roff, D.A. 1999. A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution* **53**: 1288–1292.
Dohm, M.R. 2002. Repeatability estimates do not always set an upper limit to heritability. *Funct. Ecol.* **16**: 273–280.
Edmands, S. & Harrison, J.S. 2003. Molecular and quantitative trait variation within and among populations of the intertidal copepod *Tigriopus californicus*. *Evolution* **57**: 2277–2285.
Ekblom, R., Sæther, S.A., Jacobsson, P., Fiske, P., Sahlman, T., Grahn, M., Kålås, J.A. & Höglund, J. 2007. Spatial pattern of MHC class II variation in the great snipe (*Galllinago media*). *Mol. Ecol.* (in press).
Endler, J.A. 1977. *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.
Excoffier, L., Smouse, P. & Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.
Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*, 4th edn. Harlow, UK.
Fraser, D.L. & Bernatchez, L. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Mol. Ecol.* **10**: 2741–2752.
Gjershaug, J.O., Thingstad, P.G., Eldøy, S. & Byrkjeland, S. 1994. *Norsk Fugleatlas*. Norsk Ornitologisk Forening, Klæbu, Norway.
Goodman, S.J. 1997. R_{ST}Calc: a collection of programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Mol. Ecol.* **6**: 881–885.
Gromadzka, J., Stawarczyk, T. & Tomialoje, L. 1985. Breeding waders in Poland. *Wader Study Group Bull.* **43**: 29–33.
Hansson, B. & Richardson, D.S. 2005. Genetic variation in two endangered *Acrocephalus* species compared to a widespread congener: estimates based on functional and random loci. *Anim. Conserv.* **8**: 83–90.
Hedrick, P.W. 1999. Perspective: highly variable genetic loci and their interpretation in evolution and conservation. *Evolution* **53**: 313–318.
Hendry, A.P. 2002. QST > = ≠ < FST? *Trends Ecol. Evol.* **17**: 502.
Höglund, J., Eriksson, M. & Lindell, L.E. 1990a. Females of the lek-breeding great snipe, *Gallinago media*, prefer males with white tails. *Anim. Behav.* **40**: 23–32.
Höglund, J., Kålås, J.A. & Løfaldli, L. 1990b. Sexual dimorphism in the lekking great snipe. *Ornis Scand.* **21**: 1–6.
Höglund, J., Kålås, J.A. & Fiske, P. 1992. The costs of secondary sexual characters in the lekking great snipe (*Gallinago media*). *Behav. Ecol. Sociobiol.* **30**: 309–315.
Irwin, D., Bensch, S., Irwin, J.H. & Price, T.D. 2005. Speciation by distance in a ring species. *Science* **307**: 414–416.
Kålås, J.A., Fiske, P. & Höglund, J. 1997a. Food supply and breeding occurrences: the West-European population of the lekking great snipe (Aves). *J. Biogeogr.* **24**: 213–221.
Kålås, J.A., Kuresoo, A., Luigujoe, L. & Svartaas, S.L. 1997b. Morphometrical comparisons between Estonian and Norwegian great snipe (*Gallinago media*). *Proc. Estonian Acad. Sci. Biol. Ecol.* **46**: 115–122.
Karhu, A., Hurme, P., Karjalainen, M., Karvonen, P., Karkkainen, K., Neale, D. & Savolainen, O. 1996. Do molecular markers reflect patterns of differentiation in adaptive traits of conifers? *Theor. Appl. Genet.* **93**: 215–221.
Kinnison, M.T. & Hendry, A.P. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* **112**–-113: 145–164.
Koskinen, M.T., Haugen, T.O. & Primmer, C.R. 2002. Contemporary fisherian life-history evolution in small salmonid populations. *Nature* **419**: 826–830.
Kremer, A., Zanetto, A. & Ducousso, A. 1997. Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics* **145**: 1229–1241.
Kuresoo, A. & Leibak, E. 1994. Breeding status of snipes in Estonia and in the eastern Baltic region. In: *Fourth European Woodcock and Snipe Workshop* (H. Kalchreuter, ed.), pp. 81–84. IWRB, Slimbridge, Gloucester, UK.
Løfaldli, L., Kålås, J.A. & Fiske, P. 1992. Habitat selection and diet of Great Snipe *Gallinago media* during breeding. *Ibis* **134**: 35–43.
Lande, R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and recolonization. *Evolution* **46**: 381–389.
Latta, R.G. & McKay, J.K. 2002. Genetic population divergence: markers and traits. *Trends Ecol. Evol.* **17**: 501–502.
López-Fanjul, C., Fernández, A. & Toro, M.A. 2003. The effect of neutral nonadditive gene action on the quantitative index of population divergence. *Genetics* **164**: 1627–1633.
Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, MA.
Lynch, M., Pfrender, M., Spitze, L., Lehman, N., Allen, D., Hicks, J., Latta, L., Ottene, M., Bogue, F. & Colbourne, J. 1999. The quantitative and molecular genetic architecture of subdivided species. *Evolution* **53**: 100–110.
Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge, MA.
McKay, J.K. & Latta, R.G. 2002. Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.* **17**: 285–291.
Merilä, J. 1997. Quantitative trait and allozyme divergence in the Greenfinch (*Carduelis chloris*, Aves: Fringillidae). *Biol. J. Linn. Soc.* **61**: 243–266.
Merilä, J. & Crnokrak, P. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* **14**: 892–903.
Morozov, V.V. 1994. Great snipe. In: *Birds in Europe: Their Conservation Status* (G. M. Tucker & M. F. Heath, eds), pp. 268–269. BirdLife International, Cambridge, UK.
Nielsen, R. 1997. A likelihood approach to microsatellite population samples. *Genetics* **146**: 711–716.
Østbye, K., Næsje, T.F., Bernatchez, L. & Sandlund, O.T. 2005. Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in lake Femund, Norway. *J. Evol. Biol.* **18**: 683–702.
Panchenko, V.G. 1985. Great Snipe. In: *Migration of Birds of Eastern Europe and Northern Asia, Gruiformes – Charadriformes*

(V. D. Ilijev, J. A. Viksne & C. A. Mickelson, eds), pp. 235–239. Academy of Sciences of the USSR, Moscow.

Pearman, P.B. 2001. Conservation value of independently evolving units: sacred cow or testable hypothesis? *Conserv. Biol.* **15**: 780–783.

Piertney, S.B. & Dallas, J.F. 1997. Isolation and characterisation of hypervariable microsatellites in the red grouse *Lagopus lagopus scoticus*. *Mol. Ecol.* **6**: 93–95.

Piertney, S.B., MacColl, A.D.C., Bacon, P.J. & Dallas, J.F. 1998. Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Mol. Ecol.* **7**: 1645–1654.

Piertney, S.B., MacColl, A.D.C., Moss, R., Lambin, X. & Dallas, J.F. 1999. Spatial distribution of genetic relatedness in a moorland population of red grouse. *Biol. J. Linn. Soc.* **68**: 317–331.

Piertney, S.B., Summers, R. & Marquiss, M. 2001. Microsatellite and mitochondrial DNA homogeneity among phenotypically diverse crossbill taxa in the UK. *Proc. R. Soc. Lond. B* **268**: 1511–1517.

Podolsky, R.H. & Holtsford, T.P. 1995. Population structure of morphological traits in *Clarkia dudleyana*. I. Comparison of F_{ST} between allozymes and morphological traits. *Genetics* **140**: 733–744.

Prout, T. & Barker, S.F. 1993. *F*-statistics in *Drosophila buzzatii*: selection, population size and inbreeding. *Genetics* **134**: 369–375.

Raymond, M. & Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity* **86**: 248–249.

Reed, D.H. & Frankham, R. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**: 1095–1103.

Rogers, A.R. 1986. Population differences in quantitative characters as opposed to gene frequencies. *Am. Nat.* **127**: 729–730.

Rogers, A.R. & Harpending, H.C. 1983. Population structure and quantitative characters. *Genetics* **105**: 985–1002.

Roseman, C.C. 2004. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. *Proc. Natl Acad. Sci. USA* **101**: 12824–12829.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219–1228.

Sæther, S.A., Kålås, J.A. & Fiske, P. 1994. Age determination of breeding shorebirds: quantification of feather wear in the lekking great snipe. *Condor* **96**: 959–972.

Sæther, S.A., Fiske, P., Kålås, J.A. & Gjul, J.M. 2000. Females of the lekking great snipe do not prefer males with whiter tails. *Anim. Behav.* **59**: 273–280.

Saint-Laurent, R., Legault, M. & Bernatchez, L. 2003. Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smalet ecotypes (*Osmerus mordax* Mitchill). *Mol. Ecol.* **12**: 315–330.

Sambrook, X., Fritsch, E.F. & Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Laboratory Press, New York.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Schneider, S., Kueffer, J.-M., Roessli, D. & Excoffier, L. 1997. ARLEQUIN ver 1.1 A software for population genetic data analysis[www document]. URL http://lgb.unige.ch/arlequin/.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457–462.

Smith, T.B., Wayne, R.K., Girman, D.J. & Bruford, M.W. 1997. A role for ecotones in generating rainforest biodiversity. *Science* **276**: 1855–1857.

Smouse, P.E., Long, J.C. & Sokal, R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* **35**: 627–632.

Sokal, R.R. & Rohlf, F.J. 1981. *Biometry: The Principles and Practice of Statistics in Biological Research*. W. H. Freeman and Company, New York.

Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* **135**: 367–374.

Stockwell, C.A., Hendry, A.P. & Kinnison, M.T. 2003. Contemporary evolution meets conservation biology. *Trends Ecol. Evol.* **18**: 94–101.

Storz, J.F. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Mol. Ecol.* **11**: 2537–2551.

Tomkovich, P.S. 1992. Breeding-range and population changes of waders in the former Soviet Union. *Br. Birds* **85**: 344–365.

Weir, B.S. & Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Whitlock, M.C. 1999. Neutral additive genetic variance in a metapopulation. *Genet. Res.* **74**: 215–221.

Widemo, F. & Sæther, S.A. 1999. Beauty is in the eye of the beholder: causes and consequences of variation in mating preferences. *Trends Ecol. Evol.* **14**: 26–31.

Wilding, C.S., Butlin, R.K. & Grahame, J. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J. Evol. Biol.* **14**: 611–619.

Supplementary material

The following supplementary material is available for this article:

Appendix A Analysis of systematic observer bias.

Appendix B Microsatellite primers, polymorphism and heterozygosity.

Appendix C Nested ANOVAS of morphological variation, mean values and frequency distributions.

Appendix D Pairwise Q_{ST} values, and relationship with R_{ST}.

This material is available as part of the online article from http://www.blackwell-synergy.com/doi/abs/10.1111/j.1420-9101.2007.01328.x

Received 20 January 2007; accepted 24 January 2007