**Weak geographical structure in sperm morphology across
the range of two willow warbler *Phylloscopus trochilus*
subspecies in Scandinavia**

Journal:	<i>Journal of Avian Biology</i>
Manuscript ID	JAV-00981.R1
Wiley - Manuscript type:	Article
Date Submitted by the Author:	03-Mar-2016
Complete List of Authors:	Støstad, Hanna; University of Oslo, Natural History Museum Rekdal, Silje; University of Oslo, Natural History Museum Kleven, Oddmund; Norwegian Institute for Nature Research; University of Oslo, Natural History Museum Laskemoen, Terje; University of Oslo, Natural History Museum Marthinsen, Gunnhild; University of Oslo, Natural History Museum Johnsen, Arild; University of Oslo, Natural History Museum Lifjeld, Jan T.; University of Oslo, Natural History Museum
Keywords:	Sperm length, trait divergence, migratypes
Abstract:	<p>Sperm morphology is highly diversified among species and at higher taxonomic levels. In birds, there is also increasing evidence of geographical differentiation in sperm traits within species, especially in those with strong sperm competition. Geographical divergences in sperm traits might imply the formation of a reproductive barrier in a speciation process. Here we study sperm morphology variation of willow warblers <i>Phylloscopus trochilus</i> in a geographical context in Scandinavia, across the range of two subspecies that are differentiated in certain genetic markers, morphology and migratory routes. We describe geographical patterns in genotypes (two previously described single-nucleotide polymorphism (SNP) markers and four polymorphic microsatellites); in wing length, tarsus length and body mass; and in sperm traits of 330 male willow warblers sampled at 33 localities across Norway (58° N – 69° N). Birds were on average larger and longer-winged in the north (spp. <i>acredula</i>) than in the south (spp. <i>trochilus</i>), and showed a sigmoid change in the SNP allele frequencies and morphology around 65° N. We found no evidence of genetic structuring in the microsatellites. There was no geographical variation in sperm traits across Norway, except that sperm heads were on average longer in the south. Sperm head length was also associated with the two SNP markers, with longer sperm heads for the southern alleles, and midpiece length was weakly associated with one of the SNP markers. Similar among-male variances in total sperm length among the 33 sampling sites indicate uniform levels of sperm competition across Norway. We conclude that sperm morphology remains a rather undifferentiated trait between the two willow warbler subspecies in Scandinavia, which is consistent with a pattern of a shallow genetic divergence. This indicates that sperm</p>

	morphology is not a reproductive barrier maintaining the narrow hybrid zone.

SCHOLARONE™
Manuscripts

For Review Only

1 **Weak geographical structure in sperm morphology across the range**
2 **of two willow warbler *Phylloscopus trochilus* subspecies in**
3 **Scandinavia**

4
5 Hanna N. Støstad^{1*}, Silje L. Rekdal^{1,2}, Oddmund Kleven^{1,3,4}, Terje Laskemoen^{1,5}, Gunnhild
6 Marthinsen^{1,6}, Arild Johnsen^{1,7}, Jan T. Lifjeld^{1,8}

7
8
9 ¹Natural History Museum, University of Oslo, 0318 Oslo, Norway

10 ²s.l.rekdal@nhm.uio.no

11 ³Norwegian Institute for Nature Research, NO-7485 Trondheim, Norway

12 ⁴oddmund.kleven@nina.no

13 ⁵terje.laskemoen@nhm.uio.no

14 ⁶g.m.marthinsen@nhm.uio.no

15 ⁷arild.johnsen@nhm.uio.no

16 ⁸j.t.lifjeld@nhm.uio.no

17
18
19 *Corresponding author. Email: h.n.stostad@nhm.uio.no
20

21 **Abstract**

22 Sperm morphology is highly diversified among species and at higher taxonomic levels. In birds, there
23 is also increasing evidence of geographical differentiation in sperm traits within species, especially in
24 those with strong sperm competition. Geographical divergences in sperm traits might imply the
25 formation of a reproductive barrier in a speciation process. Here we study sperm morphology
26 variation of willow warblers *Phylloscopus trochilus* in a geographical context in Scandinavia, across
27 the range of two subspecies that are differentiated in certain genetic markers, morphology and
28 migratory routes. We describe geographical patterns in genotypes (two previously described single-
29 nucleotide polymorphism (SNP) markers and four polymorphic microsatellites); in wing length, tarsus
30 length and body mass; and in sperm traits of 330 male willow warblers sampled at 33 localities across
31 Norway (58° N – 69° N). Birds were on average larger and longer-winged in the north (spp. *acredula*)
32 than in the south (spp. *trochilus*), and showed a sigmoid change in the SNP allele frequencies and
33 body morphology around 65° N. We found no evidence of genetic structuring in the microsatellites.
34 There was no geographical variation in sperm traits across Norway, except that sperm heads were on
35 average longer in the south. Sperm head length was also associated with the two SNP markers, with
36 longer sperm heads for the southern alleles, and midpiece length was weakly associated with one of
37 the SNP markers. Similar among-male variances in total sperm length among the 33 sampling sites
38 indicate uniform levels of sperm competition across Norway. We conclude that sperm morphology
39 remains a rather undifferentiated trait between the two willow warbler subspecies in Scandinavia,
40 which is consistent with a pattern of a shallow genetic divergence. This indicates that sperm
41 morphology is not a reproductive barrier maintaining the narrow hybrid zone.

42
43 **Keywords:** Sperm length, trait divergence, migratypes, sperm competition, speciation

Sperm cells exhibit a striking amount of variation across species and higher taxa (Briskie and Montgomerie 1992; Gage 1994; Morrow and Gage 2000; Immler and Birkhead 2007; Birkhead et al. 2009). The evolution of sperm morphology depends to a large extent on interactions between the male and the female, as it is essential for males to match their sperm cells to the anatomy and physiology of the reproductive tract of females of the same species (Gomendio and Roldan 1993; Pitnick et al. 2003; Higginson et al. 2012). Thus, divergent evolution in sperm morphology can be a forerunner to reproductive isolation and speciation.

Passerine birds are an excellent study system for the evolution of sperm competition and sperm morphology, due to the fact that they are generally promiscuous and therefore potentially subject to post-copulatory sexual selection (Griffith et al. 2002). Interspecific variation in sperm length is fairly high, with mean lengths ranging from about 42 μm to 285 μm among 196 studied passerine species (Immler et al. 2011). Recent comparative studies of passerine birds have shown that high evolutionary rates in sperm traits are associated with high levels of sperm competition (Rowe et al. 2015a). In populations with competition for fertilisations, optimal sperm morphology is likely to be important, and so sperm competition presumably acts as a force of stabilizing selection on sperm length (Calhim et al. 2007, Kleven et al. 2008, Lifjeld et al. 2010). Sperm competition appears to minimise variation in sperm length among males in a population (Kleven et al. 2008), which means that the among-male coefficient of variation (CV_{bm}) is a good predictor for sperm competition in a population (Lifjeld et al. 2010; Laskemoen et al. 2013a). ~~Thus post-copulatory selection on sperm length appears to result in high interspecific variation but low intraspecific variation. Thus we see two effects of post-copulatory sexual selection on sperm morphology: stabilising selection appears to result in low variation within a population, but over evolutionary time, species tend to diverge due to the fast evolutionary rates of these traits.~~

Intraspecific, geographical differences in sperm morphology are known from birds (e.g. Lüpold et al. 2011; Schmoll and Kleven 2011; Laskemoen et al. 2013a; Hogner et al. 2013; Rowe et al. 2015b) and other taxa (reviewed in Pitnick et al. 2009). Hogner et al. (2013) found differences in sperm morphology among European subspecies of bluethroats *Luscinia svecica*, possibly indicating early stages of speciation. Laskemoen et al. (2013a) found similar patterns in the barn swallow *Hirundo rustica*, as well as an among-population correlation between sperm CV_{bm} and extrapair paternity rates. Lüpold et al. (2011) showed a more large-scale pattern for the red-winged blackbird *Agelaius phoeniceus*, with a gradual increase of sperm length from southwest to northeast of the breeding range in North America. Schmoll and Kleven (2011) found significant differences in sperm length of coal tits *Periparus ater* between Germany and Norway, which belong to two different subspecies (Pentzold et al. 2013). Rowe et al. (2015b) also showed small, but significant differences in sperm

length between two subspecies of the long-tailed finch *Poephila acuticauda*. On the other hand, the Azores bullfinch *Pyrrhula murina* and the Eurasian bullfinch *P. pyrrhula* provide an example of two sister species with undifferentiated sperm morphology (Lifjeld et al. 2013). Notably, the bluethroat (Johnsen and Lifjeld 2003), barn swallow (Laskemoen et al. 2013a), red-winged blackbird (Gibbs et al. 1990; Gray 1996) and coal tit (Schmoll et al. 2003) are all known to have relatively high levels of sperm competition, whereas the bullfinches appear to have relatively low or absent sperm competition (Birkhead et al. 2006; Lifjeld et al. 2013). It is therefore possible that the increased selection pressure from sperm competition leads to faster sperm evolution and thus more divergences among populations, which makes comparative sperm morphology in subspecies complexes with high sperm competition an interesting topic for speciation research.

The willow warbler *Phylloscopus trochilus* is one of the most numerous bird species in Scandinavia, and its distribution extends across the northern Palearctic from the British Isles to eastern Siberia (BirdLife International 2013). In Scandinavia, there are two subspecies; *trochilus* in the south and *acredula* in the north. They meet in a narrow contact zone between 62°N and 63°N (Chamberlain et al. 2000; Bensch et al. 1999; 2002; 2009), although the detailed mapping of the contact zone has mainly been carried out in Sweden, with few sampling points in Norway. The two subspecies represent two distinct migratory phenotypes, or “migratypes”; *trochilus* migrates to the southwest, whereas *acredula* migrates via a south-southeast route (Hedenström and Pettersson 1987). They also spend the winter in different regions of sub-Saharan Africa (Chamberlain et al. 2000). It is thought that the two migratypes reflect the colonization of Scandinavia from two directions after the last ice age; *trochilus* from the south and *acredula* from the north (Bensch et al. 1999; 2002; 2009). The migratypes differ genetically at the bi-allelic WW2 locus, which might be linked to genes associated with migratory behaviour (Bensch et al. 2009). The narrow contact zone could in this way be maintained through selection against hybrids with maladaptive migration direction (Bensch et al. 2002; 2009; Liedvogel et al. 2014). Further, geographic variation is also demonstrated at the WW1 locus (Bensch et al. 2002; Lundberg et al. 2011; Larson et al. 2014). Bensch et al. (2002) suggested that this is a noncoding SNP, which Lundberg et al. (2011) located in a genomic region of about 2.5 Mb that is differentiated between northern and southern willow warblers in Sweden. As the geographic pattern of the WW1 marker is more consistent with habitat and climatic factors than migration routes, selective forces related to environmental variables might act upon genomic regions that are linked to the WW1 locus (Lundberg et al. 2011, Larson et al. 2014).in allele frequencies is also found at the WW1 locus (Bensch et al. 2002; Larson et al. 2014). As this pattern is more consistent with habitat and climatic factors than migration routes, Larson et al. (2014) suggested that selective forces related to environmental variables might act upon the WW1 locus or linked genomic

112 | ~~regions~~. In addition, there are also phenotypic divergences between the subspecies, i.e. in wing
 113 | length, body mass and plumage colouration (Fonstad and Hogstad 1981; Bensch et al. 1999). On the
 114 | other hand, there seems to be a lack of genetic differentiation at neutral loci (Bensch et al. 1999;
 115 | 2002; 2009), which might indicate a rather recent divergence and/or high levels of gene flow
 116 | between the two subspecies. This was further elaborated by Lundberg et al. (2013), who found that
 117 | out of 85 000 SNPs obtained from brain-derived cDNA, only 55 were highly differentiated between
 118 | trochilus and acredula, and they clustered in two chromosome regions.

120 | The willow warbler has sperm cells of the corkscrew-twisted shape which is typical for passerine
 121 | spermatozoa, but in the lower size range (Lifjeld et al. 2010, Immler et al. 2011), also when compared
 122 | with other members of the Phylloscopidae family (Supriya et al. ~~in-review~~2016). It has a socially
 123 | monogamous mating system with a high frequency of extra-pair paternity, ranging from 23% to 33%
 124 | extra-pair offspring in three studies in northern Europe (Bjørnstad and Lifjeld 1997; Fridolfsson et al.
 125 | 1997; Gil et al. 2007). Given the general tendency for sperm cells of such species to evolve fast (Rowe
 126 | et al. 2015a), one might expect to find differences in sperm morphology between the two subspecies
 127 | in Scandinavia. Recent comparative work on sperm evolution in passerine families has indicated that
 128 | evolutionary rates can be lineage-specific and may change over evolutionary time-scales
 129 | (Omotorigun et al. 2016; Supriya et al. ~~in-review~~2016). In fact, the latter study indicated that sperm
 130 | size differentiated relatively early in the evolutionary history of the Phylloscopidae, with subsequent
 131 | low rates of evolution. This would, therefore, predict low sperm differentiation between the two
 132 | willow warbler subspecies. Whether the two subspecies are actually differentiated in sperm
 133 | morphology has yet not been investigated.

134 | This study has two main aims. First, we wanted to describe the genotypic (~~WW1, WW2,~~
 135 | ~~microsatellites~~) and phenotypic (~~wing, tarsus, body mass~~) trait variation between the two subspecies
 136 | of willow warblers in Norway, and compare the intergradation zone to the one described in Sweden.
 137 | Second, we wanted to examine whether sperm morphology shows geographical variation and
 138 | especially if there is any differentiation in sperm traits between the two subspecies.

139

140 | Methods

141 | During the breeding season (May 18 – June 13) of 2008, we collected samples from 330 willow
 142 | warbler males, captured at 33 sites across Norway (Fig. 1), i.e. sampling 10 males at each site. Males
 143 | were caught in mist-nets using playback of willow warbler song. We obtained sperm samples using
 144 | cloacal massage (as described in Laskemoen et al. 2013b), collected the ejaculate in a microcapillary

145 and fixed it in a tube containing a 5% formaldehyde solution. We also measured wing length (n =
146 330), tarsus length (n = 300) and body mass (n = 328) along with a blood sample for DNA analysis (n =
147 330).

148 *Sperm measurements*

149 We were able to analyse 325 normal sperm samples. Five samples had no sperm cells or abnormal
150 sperm. For each sperm sample, a small aliquot of approximately 15 µl was applied on a microscope
151 slide, allowed to air-dry, and subsequently gently rinsed with distilled water and air-dried again. We
152 used a Leica DFC420 camera mounted on a Leica DM6000 B digital light microscope to obtain digital
153 images at magnifications of 160×. The morphometric measurements were conducted using Leica
154 Application Suite (version 2.6.0 R1). Head, midpiece, and tail (±0.1 µm) of ten intact spermatozoa per
155 male were measured by a single observer. Measurements by this observer have earlier been shown
156 to be highly repeatable (Laskemoen et al. 2013b). Flagellum length was calculated as the sum of
157 midpiece and tail length, and total length as the sum of all three sperm components. Measuring ten
158 sperm cells per male has been shown to give representative estimates of an individual's mean sperm
159 length (Laskemoen et al. 2007). We calculated the coefficient of among-male variation in total sperm
160 length as $CV_{bm} = SD/mean * (1+1/4n) * 100$, where n is the number of males. We used this sperm CV_{bm}
161 metric as an indicator of the level of sperm competition (Lifjeld et al. 2010).

162 *Genetic analyses*

163 Genomic DNA was extracted from blood using a commercial kit (E.Z.N.A. DNA extraction kit, Omega
164 Bio-Tek, Inc., Norcross, GA, USA). Individuals were typed at four polymorphic microsatellite loci
165 (Table 1). Loci were amplified with fluorescently labelled forward primers using multiplex polymerase
166 chain reaction (PCR). Multiplexing was performed with Qiagen multiplex PCR kit (Qiagen, Hilden,
167 Germany) following the manufacturer's protocol, but using a 10 µL reaction volume. Alleles were
168 separated using capillary electrophoresis on an ABI 3130xl Genetic Analyzer and sizes assigned using
169 GENEMAPPER software (Applied Biosystems, Foster City, CA, USA). Marker polymorphism was
170 calculated using GenAlEx (Peakall and Smouse 2012) and is presented in Table 1. Genotyping of the
171 WW1 and WW2 bi-allelic markers followed previously published methods (Bensch et al. 2002; 2009).
172 Some markers did not amplify for a small number of individuals; see results tables for specific
173 numbers.

174 A possible population genetic structure of willow warblers in Norway was assessed through the
175 software STRUCTURE v 2.3.4 (Pritchard et al. 2000) using the three anticipated neutral microsatellites
176 *Pocc1*, *Pocc6* and *Pocc8* (Bensch et al. 1997) and the polyglutamine repeat in the *Clock* gene (Johnsen

et al. 2007). Although this repeat is located in a functional gene important in circadian rhythm (e.g. Young and Kay 2001), the photoperiodicity at different latitudes might affect resident more than migratory species (Johnsen et al. 2007). Johnsen et al. (2007) found that the migratory bluethroat did not show any latitudinal pattern in allele frequencies while the mainly resident blue tit (*Cyanistes caeruleus*) ~~did~~ showed higher frequency of longer alleles at higher latitudes. As the willow warbler also is a migratory bird, we expected this locus to behave like a neutral genetic marker in this species. For further confirmation of this, we tested the *Clock* gene for geographic signal, and found that there was no correlation between average allele size of the *Clock* gene and latitude (linear model, $n = 330$, $t = 0.65$, $p = 0.52$). Furthermore, a STRUCTURE analysis (see below) including only the *Clock* locus revealed no evidence for geographic structuring (data not shown). We therefore grouped the *Clock* gene together with the three neutral microsatellites for the rest of the analysis. The bi-allelic markers *WW1* and *WW2*, for which there is previously demonstrated genetic structure (e.g. Bensch et al. 2002; 2009), were excluded, to avoid adding noise and masking the results when testing other markers. STRUCTURE was run using default parameters with the admixture model and correlated allele frequencies among populations, with 10 iterations for every K between $K=1$ and $K=5$. The length of burnin was 1 000 000, as was the number of Markov Chain Monte Carlo (MCMC) steps. In order to detect the real number of clusters in the dataset (K), we used the ΔK approach of Evanno et al. (2005), founded on the rate of change in the log probabilities for each K . The results were visualized by Structure Harvester v 0.6.94 (Earl and vonHoldt 2012) and the online version of CLUMPAK (February 2015) (Kopelman et al. 2015).

197 Statistical analyses

198 All statistics were performed with R statistical software v 3.2.2 (R Core Team 2014), using the *car* and
199 *stats* packages.

200 To determine the response of the *WW1* and *WW2* markers to geographical variables, we used
201 generalised linear models (GLM). The frequencies of the northern (*N*) allele at each site were used as
202 response variables in two separate models, one for each bi-allelic marker. The predictor variables
203 were latitude and elevation. The interaction between them was initially included, but was not
204 significant for any model and was subsequently removed. Elevation was included due to the previous
205 finding of *WW1* allele frequency being associated with elevation (Larson et al. 2014). Our *WW1* data
206 are identical to those presented for Norway in Larson et al. (2014). Longitude was not used because
207 the orientation of the country (southwest-northeast) meant that latitude could be used as a proxy for
208 both latitude and longitude. A binomial error structure and a logit link function were used due to the
209 proportional structure of the data. We used multiple linear regression models in the same way for

analysing geographical variation in sperm and body traits. Linear models were appropriate here due to the normal distribution of the data.

Finally, for determining the response of sperm traits to the presence of the northern (N) and southern (S) alleles at the WW1 and WW2 loci in each individual, we used one-way ANOVAs for all traits.

For the distribution maps, we followed the methods provided in Larson et al. (2014), using the Spatial Analyst ordinary kriging function in ArcMap 10 (ESRI 2014) to make a raster surface of the distribution of the relevant variables, overlaid on a map of Norway. The sample sites were overlaid on a digital elevation model (DEM) background.

Results

Geographical structure in genotypic and phenotypic traits

There was no indication of any substructuring in the willow warblers across Norway for the four microsatellites. This is implied from the STRUCTURE results, where there is no evidence for $K>1$, as $K=1$ is the most likely number of clusters (see Fig. 2a). Further, the algorithms did not converge for any K other than $K=1$ (Fig. 2a). By applying the ΔK approach (Evanno et al. 2005), $K=2$ was most likely (Figure 2b). However, as ΔK makes no sense for $K=1$ and each individual is approximately $1/K$ assigned to each cluster (Fig. 2c), there is presumably no population substructuring at the loci tested (Pritchard et al. 2010).

WW1 and WW2 allele frequencies showed a strong north/south structure, with a clear shift in central Norway (Table 2; Fig. 3a, 3b). However, there was variation throughout the sampling area; ~~neither the extreme south nor north had 100% occurrence of the S and N allele respectively for~~ example, at the most southern site (Storaker) there was 90% occurrence of the S-allele for WW1 but only 65% occurrence of the S-allele for WW2, whereas at the most northern site (Olderfjord) there was 85% occurrence of N-allele for WW1 and 95% occurrence of the N-allele for WW2 (see electronic appendix for per-site data). The intergradation zone was at about 65° N for both allele markers. This is slightly further north than has been found in Sweden, where the zone is at about 62° N -63° N (Bensch et al. 2009). WW1 allele frequency was related to elevation as well as latitude (Table 2), with northern alleles being more common at high elevation in southern Norway, which can be seen in Fig. 3a where the lighter areas correspond to mountainous regions of southern Norway. WW2 allele frequency was not related to elevation (Table 2).

Wing length and body mass showed a positive correlation with latitude (Table 3), i.e. individuals were larger and had longer wings in the north (Fig. 3c, 3d). The intergradation zone for both wing length and body mass was similar to the intergradation zone for the genetic markers. Body mass had a more gradual cline than wing length. Body mass was also lower at high elevation, and higher when birds were captured late in the day (Table 3). Tarsus length showed no geographical pattern (Table 3).

Geographical variation in sperm traits

Total sperm length (Fig. 4a), midpiece length and flagellum length did not show any geographical or clinal variation (Table 4). R^2 values in the models approached zero which means that very little of the variation could be explained by latitude or elevation (Table 4).

Sperm head length was shorter in the north (Table 4, Fig. 4b), and was associated with the *WW1* and *WW2* genotypes (Table 5; Fig. 5a, 5b) – i.e. birds that were homozygous for the *S* alleles at *WW1* had longer sperm heads, and the same for *WW2*. Based on the results of these initial analyses, a separate ANCOVA was run with latitude, *WW1* genotype and *WW2* genotype as predictor variables, to determine which of the predictors had the strongest association with sperm head length. *WW1* genotype was significantly associated with sperm head length ($F_{3,317} = 3.56$, $p = 0.03$), whereas *WW2* genotype ($F_{3,317} = 1.03$, $p = 0.36$) and latitude ($F_{3,317} = 0.55$, $p = 0.46$) had no significant additional effects. Latitude and elevation together explained only 2.3% of the variation in head length (Table 4). Midpiece length was shorter in individuals that were homozygous for the southern allele at *WW1* (Table 5, Fig. 5c), although the mean difference was only 0.6 μm (69.2 for *SS* and 69.8 for *NN*). There was no effect of *WW2* on midpiece length (Fig. 5d). None of the other sperm morphology variables were associated with *WW1* or *WW2* genotype (Table 5).

We found no significant heterogeneity in sperm CV_{bm} of total sperm length among the 33 sampling sites (Levene's test, $F_{32,292} = 0.88$, $p = 0.66$). We also analysed the CV_{bm} metric on a more regional basis by merging study sites into larger groups and thereby increasing sample size for the CV_{bm} estimates. Three geographical zones were defined based on the location of the intergradation zone of the other variables in the study (genotype, wing length, mass): south ($< 63^\circ\text{N}$, $n = 166$ males), central ($63^\circ\text{N} - 67^\circ\text{N}$, $n = 79$) and north ($> 67^\circ\text{N}$, $n = 80$). We found no evidence for a difference in CV_{bm} across these three zones ($F_{2,322} = 0.72$, $p = 0.49$). CV_{bm} for the south zone was 2.30%, the central zone 2.07%, and the north zone 1.91%. There was also no difference in CV_{bm} between *WW1* genotypes ($F_{2,322} = 0.45$, $p = 0.64$) nor between *WW2* genotypes ($F_{2,318} = 0.07$, $p = 0.93$). Sperm length CV_{bm} for the entire sampling area was 2.15%.

273

274 **Discussion**

275 We found a clear population structure in certain genetic (*WW1*, *WW2*) and morphological (body
276 mass, wing length) traits. We also show that the *WW2* marker has a similar geographical structure in
277 Norway to that previously reported in Sweden, although the intergradation zone is further north
278 (Bensch et al. 2009). However, despite this population structure, and despite high levels of sperm
279 competition (Bjørnstad and Lifjeld 1997), sperm traits of the willow warbler were largely
280 homogenous across the sampling area in Scandinavia. There was also a lack of structure in
281 microsatellites. However, there were associations between sperm head length and alleles at the
282 *WW1* and *WW2* loci, and between midpiece length and the *WW1* locus.

283 The willow warbler is thought to have had only one panmictic glacial refugium (Bensch et al. 1999),
284 and following the colonisation of Fennoscandia in a ring-like manner after the last ice age, the
285 migratory divide is considered a secondary contact zone (Bensch et al. 1999; 2009). However, it has
286 been shown that the width of the contact zone is substantially smaller than expected assuming
287 random mating and no selection (Bensch et al. 1999; 2009). As there is little evidence of pre-
288 copulatory selection in terms of assortative mating (Liedvogel et al. 2014), the narrow contact zone
289 could be explained by on-going post-zygotic selection against hybrids with intermediate migration
290 behaviour (e.g. Bensch et al. 2002, Liedvogel et al. 2014).)-This could be explained by on-going post-
291 zygotic selection against hybrids with intermediate migration behaviour (e.g. Liedvogel et al. 2014). If
292 there is strong selection against hybrids across the migratory divide, we could expect to see signs
293 of differentiation at other rapidly evolving traits. Due to the high occurrence of extra-pair paternity in
294 the willow warbler (Bjørnstad and Lifjeld 1997; Fridolfsson et al. 1997; Gil et al. 2007), sperm
295 competition is expected to be relatively high in this species, which could lead to rapid evolution in
296 sperm morphology (Rowe et al. 2015a). However, we found no geographic structure in total sperm
297 length for the willow warbler. This means that it is unlikely that sperm morphology is maintaining the
298 narrow contact zone by acting as a post-copulatory, pre-zygotic reproductive barrier. This-The
299 geographic homogeneity in sperm morphology indicates that there is little effect of latitude
300 (including related factors such as daylight hours, climate or migration distance) on total sperm length.
301 Alternatively, there has not been enough time for sperm length to diverge between the two
302 subspecies, which represent a very recent divergence (Bensch et al. 2009), or there is too much
303 hybridisation for sufficient selection to occur. The results are consistent with the hypothesis that
304 sperm traits in Phylloscopidae are in evolutionary stasis (Supriya et al. in review 2016).

Sperm heads were on average slightly longer in the southern part of the country, and birds that were homozygous for the southern (S) alleles at *WW1* and *WW2* had longer sperm heads. Homozygotes for the S-allele of *WW1* also had shorter midpieces, which presumably resulted in the lack of an effect on total length. Head length might therefore be associated with environmental selection, as reflected by the *WW1* marker, or a result of divergence during the evolutionary history of the two subspecies, as reflected by the *WW2* marker. Further studies are required to test whether genes coding for sperm head length are localised close to the *WW2* locus or in the divergent chromosome region that includes *WW1* (Lundberg et al. 2011), or elsewhere in the genome~~Further studies are required to test whether genes coding for sperm head length are localized close to the *WW1* or the *WW2* locus, or elsewhere in the genome.~~ Recent evidence suggests that head length evolves more slowly in response to sperm competition than other sperm traits (Rowe et al. 2015a), and it is known that head length tends to be a conserved trait in birds (Jamieson 2007; Rowe et al. 2015a). However, there have also been several studies where head length has been found to diverge, both intraspecifically (Schmoll and Kleven 2011; Immler et al. 2012) and among closely related species (Omotoriogun et al. 2016). This suggests that head length evolves separately from other sperm traits, although the reason for this separate evolution is unclear. It should be emphasized that head length is a composite trait consisting of the acrosome and the nucleus, and that it is not yet possible to ascertain whether the observed differences in head length is due to the length of the acrosome, the nucleus, or both. Further analyses in this respect would require scanning electron microscopy (cf. Rowe et al. 2015b).

The genetic structure of willow warblers in Fennoscandia has received interest previously due to the presence of the migratory divide (Bensch et al. 1999; 2002; 2009; Lundberg et al. 2013; Larson et al. 2014), between the southern migratype *trochilus*, which migrates southwest, and the northern migratype *acredula*, migrating south-southeast (Hedenström and Pettersson 1987; Chamberlain et al. 2000). Our results from Norway show a latitudinal geographic pattern in *WW2* allele frequency, corresponding to this migratory divide, which has previously been found in Sweden (Bensch et al. 2009), although the intergradation zone is slightly further north in Norway than was reported from Sweden (~ 65° N versus 62° N - 63° N). A similar geographic pattern has already been shown in Norway and Sweden for *WW1* (Larson et al. 2014). The *N*-allele (northern allele) at *WW1* also tends to be associated with an adaptation to subalpine birch forest (Bensch et al. 2002; Larson et al. 2014). The *WW2*-locus, on the other hand, shows a latitudinal cline with no association with elevation. Bensch *et al.* (2009) suggest that the *WW2*-locus may be linked to genes that are important in migratory behaviour, and our results support this due to the corresponding intergradation zones of wing length and *WW2* allele frequency, at about 65° N. Using the *Clock* gene and the microsatellites

Pocc1, *Pocc6* and *Pocc8*, the STRUCTURE results imply little or no geographic variation at these markers. This is in concordance with the lack of differentiation previously shown at neutral loci (Bensch et al. 1999; 2009) and at certain coding genes (Bensch et al. 2006). The recent split between the two migratypes after the last glaciation (Bensch et al. 2009) might restrict the amount of differentiation observed at neutral loci.

Body mass and wing length both showed an increase with latitude. This is as expected considering Bergmann's Rule, stating that body size increases with increasing latitude, which has robust support from several studies on birds (Blackburn et al. 1999; Ashton 2002; Ramirez et al. 2008; Olson et al. 2009). Wing length has been shown to be positively correlated with migration distance (Marchetti et al. 1995; Copete et al. 1999), so it is expected that the birds at high latitudes which migrate longer distances also have longer wings. Intriguingly, body mass was lower at higher elevation. The literature does not provide any obvious explanation for this pattern – the studies documenting elevational variation in body mass tend to show positive correlations (e.g. Traylor 1950; Diamond 1973; Blackburn and Ruggiero 2001), and we are therefore undecided as to whether this is a real and interesting pattern, or reflect some confounding effect we have not been able to take into account (e.g. weather).

We found no evidence of geographical or clinal structure in sperm competition in the willow warbler in Norway. The intensity of sperm competition appears to be fairly uniform across the study area, which is consistent with current literature (Bjørnstad and Lifjeld 1997; Fridolfsson et al. 1997; Gil et al. 2007). Our estimate of sperm competition using CV_{bm} would correspond to an EPY rate of 23%, which is also in the range of the existing EPY studies on the willow warbler in northern Europe (Bjørnstad and Lifjeld 1997; Fridolfsson et al. 1997; Gil et al. 2007).

Conclusion

Our results show that sperm morphology in the willow warbler is rather undifferentiated between the two subspecies in Scandinavia, which are otherwise differentiated in certain genetic (the *WW1* and *WW2* loci) and phenotypic traits (wing length and body mass). The only exception is a weak geographical structure in sperm head length, which appears to be associated with the *WW1* and *WW2* markers, and a similarly weak association between midpiece length and *WW1*. We suggest that the lack of differentiation in total sperm length is-could be due to a very shallow genetic divergence between the two subspecies. Our data also indicate that the level of sperm competition in this species is consistently high across Scandinavia with no geographical trend.

371

372 **Acknowledgments**

373 We thank Lars Erik Johannessen, Joachim Tørum Johansen and Trond Øigarden for assistance in the
 374 field. We also thank the anonymous reviewers for their useful comments. Our research was
 375 conducted in adherence to the Norwegian guidelines for the use of animals in research and
 376 supported by a grant from the Research Council of Norway.

377

378

379 **References**

- 380 Ashton K. G. 2002. Patterns of within-species body size variation of birds: strong evidence for
 381 Bergmann's rule – Global. Ecol. Biogeogr. 11:505-523. doi: 10.1046/j.1466-
 382 822X.2002.00313.x
- 383 Bensch S., Åkesson S., Irwin D. E. 2002. The use of AFLP to find an informative SNP: genetic
 384 differences across a migratory divide in willow warblers – Mol. Ecol. 11:2359-2366
- 385 Bensch S., Andersson T., Åkesson S. 1999. Morphological and molecular variation across a migratory
 386 divide in willow warblers, *Phylloscopus trochilus* - Evolution 53:1925-1935
- 387 Bensch S., Grahm M., Müller N., Gay L., Åkesson S. 2009. Genetic, morphological, and feather isotope
 388 variation of migratory willow warblers show gradual divergence in a ring – Mol. Ecol.
 389 18:3087-3096
- 390 Bensch S., Price T., Kohn J. 1997. Isolation and characterization of microsatellite loci in a *Phylloscopus*
 391 warbler – Mol. Ecol. 6:91-92. doi: 10.1046/j.1365-294X.1997.00150.x
- 392 BirdLife International 2013. *Phylloscopus trochilus*. The IUCN Red List of Threatened Species. Version
 393 2014.3. www.iucnredlist.org. Accessed 09 April 2015
- 394 Birkhead T. R., Hosken D. J., Pitnick S. S. 2009. Sperm biology: an evolutionary perspective. Academic
 395 press
- 396 Birkhead T. R., Immler S., Pellatt E. J., Freckleton R. 2006. Unusual sperm morphology in the Eurasian
 397 bullfinch *Pyrrhula pyrrhula* - Auk 123:383-392. doi: 10.1642/0004-
 398 8038(2006)123[383:Usmite]2.0.Co;2
- 399 Bjørnstad G., Lifjeld J. T. 1997. High frequency of extra-pair paternity in a dense and synchronous
 400 population of willow warblers *Phylloscopus trochilus* – J. Avian. Biol. 28:319-324. doi:
 401 10.2307/3676945

Formatted: Font: Calibri, Do not check spelling or grammar

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Blackburn T. M., Gaston K. J., Loder N. 1999. Geographic gradients in body size: a clarification of
Bergmann's rule – Divers. Distrib. 5:165-174. doi:10.1046/j.1472-4642.1999.00046.x

Blackburn, T. M., Ruggiero, A. 2001. Latitude, elevation and body mass variation in Andean passerine
birds. Global Ecol. Biogeogr. 10:245–259. doi: 10.1046/j.1466-822X.2001.00237.x

Briskie J. V., Montgomerie R. 1992. Sperm Size and Sperm Competition in Birds – P. Roy. Soc. B.-Biol.
Sci. 247:89-95. doi: 10.1098/rspb.1992.0013

Chamberlain C. P., Bensch S., Feng X., Åkesson S., Andersson T. 2000. Stable isotopes examined
across a migratory divide in Scandinavian willow warblers (*Phylloscopus trochilus trochilus*
and *Phylloscopus trochilus acredula*) reflect their African winter quarters – P. Roy. Soc. B.-Biol.
Sci. 267:43-48

Copete J. L., Marine R., Bigas D., Martinez-Vilalta A. 1999. Differences in wing shape between
sedentary and migratory reed buntings *Emberiza schoeniclus* - Bird Study 46:100-103

Diamond, A.W. 1973. Altitudinal variation in a resident and migrant passerine on Jamaica – Auk
90:610-618

Earl D. A., vonHoldt B. M. 2012. STRUCTURE HARVESTER: a website and program for visualizing
STRUCTURE output and implementing the Evanno method – Conserv. Genet. Resour. 4:359-
361. doi:10.1007/s12686-011-9548-7

ESRI 2014. ArcGIS Desktop: Release 10.3. Environmental Systems Research Institute, Redlands, CA

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the
software STRUCTURE: a simulation study – Mol. Ecol. 14:2611-2620. doi:10.1111/j.1365-
294X.2005.02553.x

Fonstad T., Hogstad O. 1981. Geographical variation and sexual dimorphism in Scandinavian willow
warblers *Phylloscopus trochilus* - Fauna Norv. Ser. C., Cinclus 4:82-88

Fridolfsson A. K., Gyllensten U. B., Jakobsson S. 1997. Microsatellite markers for paternity testing in
the willow warbler *Phylloscopus trochilus*: high frequency of extra-pair young in an island
population - Hereditas 126:127-132. doi: 10.1111/j.1601-5223.1997.00127.x

Gage M. J. G. 1994. Associations between body-size, mating pattern, testis size and sperm lengths
across butterflies – P. Roy. Soc. B.-Biol. Sci. 258:247-254. doi: 10.1098/rspb.1994.0169

Gibbs H. L., Weatherhead P. J., Boag P. T., White B. N., Tabak L. M., Hoysak D. J. 1990. Realized
reproductive success of polygynous red-winged blackbirds revealed by DNA markers -
Science 250:1394-1397. doi: 10.1126/science.250.4986.1394

Gil D., Slater P. J. B., Graves J. A. 2007. Extra-pair paternity and song characteristics in the willow
warbler *Phylloscopus trochilus* – J. Avian. Biol. 38:291-297. doi:10.1111/j.2007.0908-
8857.03868.x

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

- Gomendio M., Roldan E. R. S. 1993. Coevolution between male ejaculates and female reproductive-biology in eutherian mammals – P. Roy. Soc. B.-Biol. Sci. 252:7-12. doi: 10.1098/rspb.1993.0039
- Gray E. M. 1996. Female control of offspring paternity in a western population of red-winged blackbirds *Agelaius phoeniceus* – Behav. Ecol. Sociobiol. 38:267-278. doi: 10.1007/s002650050242
- Griffith S. C., Owens I. P. F., Thuman K. A. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function – Mol. Ecol. 11:2195-2212. doi: 10.1046/j.1365-294X.2002.01613.x
- Hedenström A., Pettersson J. 1987. Migration routes and wintering areas of willow warblers *Phylloscopus Trochilus* (L) ringed in Fennoscandia - Ornis Fennica 64:137-143
- Higginson D. M., Miller K. B., Segraves K. A., Pitnick S. 2012. Female reproductive tract form drives the evolution of complex sperm morphology – P. Natl. Acad. Sci. USA 109:4538-4543. doi:10.1073/pnas.1111474109
- Hogner S., Laskemoen T., Lifjeld J. T., Pavel V., Chutny B., Garcia J., Eybert M. C., Matsyna E., Johnsen A. 2013. Rapid sperm evolution in the bluethroat *Luscinia svecica* subspecies complex – Behav. Ecol. Sociobiol. 67:1205-1217. doi: 10.1007/s00265-013-1548-z
- Immler S., Birkhead T. R. 2007. Sperm competition and sperm midpiece size: no consistent pattern in passerine birds – P. Roy. Soc. B.-Biol. Sci. 274:561-568. doi: 10.1098/rspb.2006.3752
- Immler S., Griffith S. C., Zann R., Birkhead T. R. 2012. Intra-specific variance in sperm morphometry: a comparison between wild and domesticated zebra finches *Taeniopygia guttata* - Ibis 154:480-487. doi:10.1111/j.1474-919X.2012.01232.x
- Immler S., Pitnick S., Parker G. A., Durrant K. L., Lüpold S., Calhim S., Birkhead T. R. 2011. Resolving variation in the reproductive tradeoff between sperm size and number – P. Natl. Acad. Sci. USA 108:8065-8065. doi: 10.1073/pnas.1105524108
- Jamieson B. G. M. 2007. Avian spermatozoa: structure and phylogeny. In: Jamieson B. G. M. (ed) Reproductive biology and phylogeny of birds. Science Publisher
- Johnsen A., Fidler A. E., Kuhn S., Carter K. L., Hoffmann A., Barr I. R., Biard C., Charmantier A., Eens M., Korsten P., Siitari H., Tomiuk J., Kempnaers B. 2007. Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies – Mol. Ecol. 16:4867-4880. doi:10.1111/j.1365-294X.2007.03552.x
- Johnsen A., Lifjeld J. T. 2003. Ecological constraints on extra-pair paternity in the bluethroat - Oecologia 136:476-483. doi:10.1007/s00442-003-1286-4

469 Kleven O., Laskemoen T., Fossøy F., Robertson R. J., Lifjeld J. T. 2008. Intraspecific variation in sperm
470 length is negatively related to sperm competition in passerine birds - Evolution 62:494-499.
471 doi: 10.1111/j.1558-5646.2007.00287.x

472 Kopelman N. M., Mayzel J., Jakobsson M., Rosenberg N. A., Mayrose I. 2015. Clumpak: a program for
473 identifying clustering modes and packaging population structure inferences across K – Mol.
474 Ecol. Resour. 15:1179-1191. doi:10.1111/1755-0998.12387

475 Larson K. W., Liedvogel M., Addison B., Kleven O., Laskemoen T., Lifjeld J. T., Lundberg M., Åkesson S.,
476 Bensch S. 2014. Allelic variation in a willow warbler genomic region is associated with climate
477 clines - PLoS One 9. doi: 10.1371/journal.pone.0095252

478 Laskemoen T., Albrecht T., Bonisoli-Alquati A., Cepak J., de Lope F., Hermosell I. G., Johannessen L. E.,
479 Kleven O., Marzal A., Mousseau T. A., Moller A. P., Robertson R. J., Rudolfsen G., Saino N.,
480 Vortman Y., Lifjeld J. T. 2013a. Variation in sperm morphometry and sperm competition
481 among barn swallow *Hirundo rustica* populations – Behav. Ecol. Sociobiol. 67:301-309. doi:
482 10.1007/s00265-012-1450-0

483 Laskemoen T., Kleven O., Fossøy F., Lifjeld J. T. 2007. Intraspecific variation in sperm length in two
484 passerine species, the bluethroat *Luscinia svecica* and the willow warbler *Phylloscopus*
485 *trochilus* - Ornis Fennica 84:131-139

486 Laskemoen T., Kleven O., Johannessen L. E., Fossøy F., Robertson R. J., Lifjeld J. T. 2013b.
487 Repeatability of sperm size and motility within and between seasons in the barn swallow
488 *Hirundo rustica* – J. Ornithol. 154:955-963. doi:10.1007/s10336-013-0961-4

489 Liedvogel M., Larson K. W., Lundberg M., Gursøy A., Wassenaar L. I., Hobson K. A., Bensch S., Åkesson
490 S. 2014. No evidence for assortative mating within a willow warbler migratory divide – Front.
491 Zool. 11. doi: 10.1186/s12983-014-0052-2

492 Lifjeld J. T., Hoenen A., Johannessen L. E., Laskemoen T., Lopes R. J., Rodrigues P., Rowe M. 2013. The
493 Azores bullfinch *Pyrrhula murina* has the same unusual and size-variable sperm morphology
494 as the Eurasian bullfinch *Pyrrhula pyrrhula* – Biol. J. Linn. Soc. 108:677-687. doi:
495 10.1111/j.1095-8312.2012.02040.x

496 Lifjeld J. T., Laskemoen T., Kleven O., Albrecht T., Robertson R. J. 2010. Sperm length variation as a
497 predictor of extrapair paternity in passerine birds - PLoS One 5.
498 doi:10.1371/journal.pone.0013456

499 Lundberg, M., Åkesson S., Bensch, S. 2011. Characterisation of a divergent chromosome region in the
500 willow warbler *Phylloscopus trochilus* using avian genomic resources. J Evol Biol 24:1241-
501 1253

502 Lundberg M., Boss J., Canback B., Liedvogel M., Larson K. W., Grahm M., Åkesson S., Bensch S., Wright
503 A. 2013. Characterisation of a transcriptome to find sequence differences between two

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

Formatted: English (U.S.)

Formatted: Font: Calibri, English (U.S.), Do not check spelling or grammar

- 504 differentially migrating subspecies of the willow warbler *Phylloscopus trochilus* - BMC
505 Genomics 14. doi:10.1186/1471-2164-14-330
- 506 Lüpold S., Westneat D. F., Birkhead T. R. 2011. Geographical variation in sperm morphology in the
507 red-winged blackbird *Agelaius phoeniceus* – Evol. Ecol. 25:373-390. doi: 10.1007/s10682-010-
508 9410-5
- 509 Marchetti K., Price T., Richman A. 1995. Correlates of Wing Morphology with Foraging Behavior and
510 Migration Distance in the Genus *Phylloscopus* – J. Avian Biol. 26:177-181. doi:
511 10.2307/3677316
- 512 Morrow E. H., Gage M. J. G. 2000. The evolution of sperm length in moths – P. Roy. Soc. B.-Biol. Sci.
513 267:307-313
- 514 Olson V. A., Davies R. G., Orme C. D. L., Thomas G. H., Meiri S., Blackburn T. M., Gaston K. J., Owens I.
515 P. F., Bennett P. M. 2009. Global biogeography and ecology of body size in birds – Ecol. Lett.
516 12:249-259. doi:10.1111/j.1461-0248.2009.01281.x
- 517 Omotoriogun T. C., Laskemoen T., Rowe M., Albrecht T., Sedláček O., Horák D., Ottosson U., Lifjeld J.
518 T. 2016. ~~S~~Variation in sperm morphology ~~variation~~ among Afrotropical sunbirds - Ibis 155-166
519 ~~In press~~. doi:10.1111/ibi.12334
- 520 Peakall R., Smouse P. E. 2012. GenAEx 6.5: genetic analysis in Excel. Population genetic software for
521 teaching and research-an update - Bioinformatics 28:2537-2539.
522 doi:10.1093/bioinformatics/bts460
- 523 Pentzold S., Tritsch C., Martens J., Tietze D. T., Giacalone G., Lo Valvo M., Nazarenko A. A., Kvist L.,
524 Packert M. 2013. Where is the line? Phylogeography and secondary contact of western
525 Palearctic coal tits (*Periparus ater*: Aves, Passeriformes, Paridae) – Zool. Anz. 252:367-382.
526 doi:10.1016/j.jcz.2012.10.003
- 527 Pitnick S., Hosken D. J., Birkhead T. R. 2009. Sperm morphological diversity. In: Pitnick S., Hosken D. J.,
528 Birkhead T. R. (eds) Sperm biology: an evolutionary perspective. Academic Press
- 529 Pitnick S., Miller G. T., Schneider K., Markow T. A. 2003. Ejaculate-female coevolution in *Drosophila*
530 *mojavensis* – P. Roy. Soc. B.-Biol. Sci. 270:1507-1512. doi: 10.1098/rspb.2003.2382
- 531 Pritchard J., Wen X., Falush D. 2010. Documentation for STRUCTURE software: version 2.3.
- 532 Pritchard J. K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus
533 genotype data - Genetics 155:945-959
- 534 R Core Team 2014. R: A language and environment for statistical computing. R Foundation for
535 Statistical Computing, Vienna, Austria
- 536 Ramirez L., Diniz J. A. F., Hawkins B. A. 2008. Partitioning phylogenetic and adaptive components of
537 the geographical body-size pattern of New World birds – Global. Ecol. Biogeogr. 17:100-110.
538 doi:10.1111/j.1466-8238.2007.00346.x

539 Rowe M., Albrecht T., Cramer E. R. A., Johnsen A., Laskemoen T., Weir J. T., Lifjeld J. T. 2015a.
540 Postcopulatory sexual selection is associated with accelerated evolution of sperm
541 morphology – *Evolution* 69:1044-1052. doi:10.1111/evo.12620

542 Rowe M., Griffith S. C., Hofgaard A., Lifjeld J. T. 2015b. Subspecific variation in sperm morphology and
543 performance in the long-tailed finch *Poephila acuticauda* - *Avian Res.* 6:1-10.
544 doi:10.1186/s40657-015-0032-z

545 Schmoll T., Dietrich V., Winkel W., Epplen J. T., Lubjuhn T. 2003. Long-term fitness consequences of
546 female extra-pair matings in a socially monogamous passerine – *P. Roy. Soc. B.-Biol. Sci.*
547 270:259-264. doi:10.1098/rspb.2002.2216

548 Schmoll T., Kleven O. 2011. Sperm dimensions differ between two coal tit *Periparus ater* populations
549 – *J. Ornithol.* 152:515-520. doi: 10.1007/s10336-010-0603-z

550 Supriya K., Rowe M., Laskemoen T., Mohan D., Price T., Lifjeld J. T. ~~In review~~2016. Early diversification
551 of sperm size in the evolutionary history of the Old World leaf warblers (Phylloscopidae). – *J*
552 ~~Evol Biol.~~ doi: 10.1111/jeb.12826~~Journal of Evolutionary Biology.~~

553 ~~Traylor, M.A. 1950. Altitudinal variation in Bolivian birds – Condor 52:123-126~~

554 Young M. W., Kay S. A. 2001. Time zones: A comparative genetics of circadian clocks – *Nat. Rev.*
555 *Genet.* 2:702-715. doi: 10.1038/35088576

556

557 **Figure legends**

558 **Figure 1.** The sampling locations of willow warbler males overlaid on a digital elevation model (DEM)
559 of Norway.

560 **Figure 2.** Visualisations of the STRUCTURE results for willow warblers in Norway, based on the
561 polyglutamin repeat of the *Clock* gene and three microsatellite loci, showing a) the mean likelihood
562 of the observed data given each number of clusters (K) and the corresponding variance, b) the rate of
563 change in the log probabilities (ΔK) for every $K > 1$, based on the "Evanno method", and c) the
564 assignment of the individual males grouped by the 33 sample sites to the two clusters at $K = 2$ (each
565 cluster is represented by one colour).

566 **Figure 3.** Spatially interpolated surface maps from a willow warbler population in Norway, showing
567 the distribution of a) frequency of the *S* allele and *N* allele at the *WW1* locus, b) frequency of the *S*
568 and *N* alleles at the *WW2* locus, c) wing length (mm), and d) body mass (g).

569 **Figure 4.** Sperm a) total length and b) head length for 325 willow warbler males in Norway, according
570 to latitude of the sampling sites ($n = 33$). Lines are linear models with latitude as the predictor
571 variable and total length and head length as response variables, respectively.

572 **Figure 5.** Sperm head length for willow warbler males from Norway, for each genotype at a) the
573 *WW1* locus ($n = 325$) and b) the *WW2* locus ($n = 321$), and sperm midpiece length for each genotype
574 at the c) *WW1* locus ($n = 325$) and d) *WW2* locus ($n = 321$). Stars indicate significant differences
575 according to Tukey's post-hoc test. Error bars are 95% confidence intervals. Note that the y-axes do
576 not start at 0.

577

578 **Tables**

579 **Table 1.** Polymorphism among four microsatellite loci in the willow warbler.

Locus	n	A	Allele size range (bp)	H_o	H_e	Reference
<i>Clock</i>	330	6	262-277	0.40	0.39	Johnsen et al. (2007)
<i>Pocc1</i>	328	8	196-209	0.62	0.58	Bensch et al. (1997)
<i>Pocc6</i>	328	11	169-188	0.70	0.70	Bensch et al. (1997)
<i>Pocc8</i>	327	16	202-231	0.68	0.69	Bensch et al. (1997)
Combined				0.60	0.59	

n, number of adult individuals genotyped; A, number of alleles; H_o , observed heterozygosity; H_e , expected heterozygosity. The data are based on genotypes from adult males analysed with GenAlEx v6.501 (Peakall and Smouse 2012).

580

581

Table 2. Geographic variation in single-nucleotide polymorphism (SNP) markers. Generalised linear models (GLMs) with a binomial error structure and a logit link function, modelling the allele frequency per site of the *WW1* and *WW2* SNP markers of willow warblers in Norway as a function of latitude and elevation. The dispersion parameter is calculated from the residuals of each model and used to correct for underdispersion in the dataset (less variability than predicted from the model).

Response variable	Predictor variables	Model statistic	Dispersion parameter	P value
<i>WW1-N</i> allele frequency ($n=33$)	Latitude	$Z_{30} = 11.22$	0.068	< 0.001
	Elevation	$Z_{30} = 4.82$		< 0.001
<i>WW2-N</i> allele frequency ($n = 29$)	Latitude	$Z_{26} = 8.16$	0.075	< 0.001
	Elevation	$Z_{26} = -0.67$		0.50

582

583

584

Table 3. Body morphology measurements of willow warbler males in Norway as a function of latitude and elevation, tested using multiple linear regression models.

Response variable	Predictor variables	Model statistic	Adj. R ²	P value
Tarsus length	Latitude	t ₂₉₇ = -0.29	0.006	0.77
(n = 300)	Elevation	t ₂₉₇ = -1.90		0.06
Wing length	Latitude	t ₃₂₇ = 10.22	0.289	< 0.001
(n = 330)	Elevation	t ₃₂₇ = -1.86		0.06
Mass <u>Body mass</u>	Latitude	t ₃₁₄₂₅ = 7. 9586	0. 234326	< 0.001
(n = 328318)	Elevation	t ₃₁₄₂₅ = -3. 1995		< 0.002001
	<u>Time of day</u>	<u>t₃₁₄ = 6.70</u>		<u>< 0.001</u>

Formatted Table

585

586

587

Table 4. Sperm morphology measurements from willow warblers in Norway as a function of latitude and elevation, tested using [multiple linear regression models](#).

Response variable	Predictor variables	Model statistics	Adj. R ²	P values
Total sperm length	Latitude	$t_{2,322} = 0.44$	0	0.66
(<i>n</i> = 325)	Elevation	$t_{2,322} = 0.13$		0.90
Head length	Latitude	$t_{2,322} = -3.01$	0.023	0.003
(<i>n</i> = 325)	Elevation	$t_{2,322} = -1.69$		0.09
Midpiece length	Latitude	$t_{2,322} = 1.64$	0.001	0.10
(<i>n</i> = 325)	Elevation	$t_{2,322} = 0.12$		0.91
Flagellum length	Latitude	$t_{2,322} = 1.01$	0	0.32
(<i>n</i> = 325)	Elevation	$t_{2,322} = 0.44$		0.66

588

589

590

Table 5. Results from ANOVA tests comparing the difference in sperm morphology measurements between heterozygosity (*NS*) and homozygosity (*NN/SS*) of the bi-allelic *WW1* and *WW2* loci for each individual willow warbler. Each row represents a separate ANOVA test.

Response variable	Predictor variable	Model statistics	P values
Total sperm length	<i>WW1</i> (<i>n</i> = 325)	$F_{2,322} = 0.83$	0.44
	<i>WW2</i> (<i>n</i> = 321)	$F_{2,318} = 0.72$	0.49
Head length	<i>WW1</i> (<i>n</i> = 325)	$F_{2,322} = 3.57$	0.029
	<i>WW2</i> (<i>n</i> = 321)	$F_{2,318} = 3.77$	0.024
Midpiece length	<i>WW1</i> (<i>n</i> = 325)	$F_{2,322} = 3.27$	0.039
	<i>WW2</i> (<i>n</i> = 321)	$F_{2,318} = 0.60$	0.44
Flagellum length	<i>WW1</i> (<i>n</i> = 325)	$F_{2,322} = 1.65$	0.19
	<i>WW2</i> (<i>n</i> = 321)	$F_{2,318} = 0.74$	0.48

591

Figure 1. The sampling locations of willow warbler males overlaid on a digital elevation model (DEM) of Norway.

190x200mm (299 x 299 DPI)

Figure 2. Visualisations of the STRUCTURE results for willow warblers in Norway, based on the polyglutamin repeat of the Clock gene and three microsatellite loci, showing a) the mean likelihood of the observed data given each number of clusters (K) and the corresponding variance, b) the rate of change in the log probabilities (ΔK) for every $K > 1$, based on the "Evanno method", and c) the assignment of the individual males grouped by the 33 sample sites to the two clusters at $K = 2$ (each cluster is represented by one colour).

50x87mm (300 x 300 DPI)

Figure 3. Spatially interpolated surface maps from a willow warbler population in Norway, showing the distribution of a) frequency of the S allele and N allele at the WW1 locus, b) frequency of the S and N alleles at the WW2 locus, c) wing length (mm), and d) body mass (g).
199x210mm (300 x 300 DPI)

Figure 4. Sperm a) total length and b) head length for 325 willow warbler males in Norway, according to latitude of the sampling sites (n = 33). Lines are linear models with latitude as the predictor variable and total length and head length as response variables, respectively.
179x231mm (300 x 300 DPI)

Figure 5. Sperm head length for willow warbler males from Norway, for each genotype at a) the WW1 locus ($n = 325$) and b) the WW2 locus ($n = 321$), and sperm midpiece length for each genotype at the c) WW1 locus ($n = 325$) and d) WW2 locus ($n = 321$). Stars indicate significant differences according to Tukey's post-hoc test. Error bars are 95% confidence intervals. Note that the y-axes do not start at 0.

194x233mm (300 x 300 DPI)

Site	Municipality	Latitude (°)	Longitude (°)	Elevation (m)	Tarsus (m)
					Mean
Storaker	Mandal	58.037	7.541	51.5	22.66
Ålgård	Gjesdal	58.724	5.776	123.0	22
Treungen	Nissedal	58.995	8.529	250.0	22.45
Femsjøen	Halden	59.135	11.508	160.7	NA
Løfallstrand	Kvinnherad	60.055	6.026	40.0	21.21
Rjukan	Tinn	60.028	8.725	489.5	NA
Stangnes	Eidskog	59.960	12.037	167.5	NA
Sperillen	Ådal	60.498	10.089	261.1	22.20
Lavik	Høyanger	61.120	5.569	44.6	21.98
Fagernes	Nord-Aurdal	61.019	9.190	400.0	22.44
Bergeberget	Elverum	61.003	11.770	326.2	22.27
Øvre Heimdalen	Øystre Slidre	61.421	8.893	1120.0	22.04
Stai	Stor-Elvdal	61.543	11.052	261.8	22.39
Røberg	Stryn	61.863	6.515	400.4	21.78
Tofte	Dovre	61.998	9.216	474.6	22.11
Sømå	Engerdal	62.047	11.696	652.0	22.18
Nørdalen	Os	62.463	11.362	673.0	21.97
Torvikdalen	Gjemnes	62.923	7.871	67.0	22.09
Budal	Midtre Gauldal	62.965	10.312	485.0	22.12
Møsjødalen	Tydal	62.878	11.893	683.0	22.18
Selnes	Åfjord	63.906	9.962	12.8	22.30
Gjelsås	Steinkjer	63.955	11.604	96.5	22.57
Finnvollan	Namsskogan	64.863	13.103	265.0	22.22
Tjønndalstjønnna	Hemnes	66.022	13.844	114.5	22.35
Almo	Saltdal	66.967	15.339	47.5	22.22
Botelvatn	Hamarøy	67.978	15.956	24.5	22.37
Bjørnstad	Bardu	68.981	18.526	72.0	22.04
Rafsbotn	Alta	70.021	23.507	18.0	22.29
Olderfjord	Porsanger	70.481	25.056	36.0	22.22
Rustefjelbma	Tana	70.401	28.147	36.0	22.34
Svanvik	Sør-Varanger	69.456	30.017	30.0	22.51
Bahkilskaidi	Karasjok	69.435	25.259	137.0	22.08
Niittojavnt	Kautokeino	69.100	23.216	328.5	22.29

m) (SD)	Wing (mm)		Body mass (g)		Sperm head (μm)		Sperm mid
	SD	Mean	SD	Mean	SD	Mean	SD
NA	0.54	69.90	1.29	9.34	0.73	14.60	0.32
	0.58	70.25	1.16	9.68	0.38	14.47	0.39
	0.45	70.60	1.31	9.39	0.47	14.70	0.26
		69.50	0.88	8.99	0.27	14.66	0.37
	0.75	69.85	1.49	9.42	0.38	14.86	0.30
NA		70.05	1.19	9.15	0.52	14.49	0.28
NA		69.90	1.35	8.80	0.30	14.85	0.42
	0.54	69.90	1.47	9.42	0.33	14.72	0.39
	0.49	70.00	1.05	9.65	0.42	14.71	0.36
	0.54	69.87	1.20	9.26	0.52	14.72	0.37
	0.61	69.00	1.41	9.28	0.36	14.61	0.19
	0.48	70.05	1.44	9.18	0.51	14.43	0.37
	0.66	70.03	0.82	9.53	0.39	14.64	0.49
	0.31	70.35	1.16	9.53	0.49	14.83	0.38
	0.36	70.25	1.06	9.30	0.33	14.55	0.37
	0.88	70.40	1.56	8.87	0.46	14.70	0.39
	0.60	69.65	1.60	9.34	0.32	14.42	0.40
	0.55	70.60	0.97	9.41	0.31	14.43	0.36
	0.72	70.50	0.47	9.61	0.65	14.72	0.26
	0.72	70.70	1.01	9.28	0.30	14.73	0.18
	0.58	71.35	1.65	9.63	0.63	14.88	0.40
	0.65	71.10	1.35	9.50	0.39	14.61	0.41
	0.39	70.75	1.34	9.48	0.39	14.49	0.29
	0.55	72.20	0.95	10.02	0.62	14.59	0.31
	0.60	71.70	1.72	9.60	0.32	14.33	0.29
	0.47	71.65	1.13	9.74	0.33	14.51	0.21
	0.77	71.50	1.35	9.50	0.37	14.48	0.41
	0.56	71.90	1.02	9.87	0.38	14.78	0.41
	0.76	72.50	1.13	10.08	0.58	14.60	0.18
	0.42	71.60	1.15	10.45	0.37	14.60	0.22
	0.61	72.25	1.44	9.99	0.37	14.59	0.44
	0.70	71.80	1.60	9.49	0.49	14.21	0.36
	0.57	72.25	1.64	10.18	0.48	14.42	0.56

piece (μm)	Sperm tail (μm)		Sperm total length (μm)		CVbm	WV
SD	Mean	SD	Mean	SD		S alleles
1.42	10.09	1.12	93.47	1.76	0.0188	18
1.22	8.82	1.57	93.55	1.64	0.0174	17
1.99	10.07	1.22	93.55	1.79	0.0191	17
1.65	9.44	1.42	92.48	1.98	0.0214	20
1.37	8.57	1.03	91.78	1.47	0.0160	18
1.79	8.66	1.58	92.95	2.80	0.0301	17
1.92	9.37	1.24	93.68	2.44	0.0261	18
1.93	8.64	1.83	92.91	3.01	0.0324	20
1.18	9.53	1.69	93.67	1.67	0.0179	18
1.59	9.05	1.29	91.89	1.72	0.0187	16
2.24	9.46	1.74	93.84	2.40	0.0256	18
1.76	9.07	1.52	93.21	1.88	0.0202	9
1.37	7.93	1.04	93.25	2.02	0.0217	17
1.21	8.66	1.45	92.61	1.97	0.0213	17
1.84	9.36	1.35	93.64	1.96	0.0209	11
1.83	9.47	1.42	92.61	3.01	0.0324	8
2.35	9.05	1.29	92.57	2.15	0.0233	11
1.99	9.12	1.35	93.45	1.62	0.0174	18
1.77	10.16	1.49	93.77	1.74	0.0186	14
1.72	9.86	0.85	94.54	1.58	0.0167	8
1.60	8.95	1.43	92.81	2.33	0.0251	15
1.42	8.55	1.00	92.72	1.20	0.0129	17
1.43	9.46	1.07	93.96	1.54	0.0164	11
2.08	9.33	1.12	93.85	2.67	0.0284	7
1.71	8.72	0.91	92.68	2.25	0.0242	4
1.11	9.67	1.42	93.86	2.06	0.0220	4
1.40	9.14	1.43	93.16	1.71	0.0183	5
1.67	9.64	2.16	94.27	2.20	0.0234	2
1.30	9.57	1.21	93.15	1.61	0.0173	3
1.54	8.96	1.31	92.71	1.59	0.0171	1
1.45	8.24	1.23	93.56	1.02	0.0109	0
1.68	8.97	1.35	93.15	2.07	0.0223	2
1.33	8.71	1.15	92.12	1.35	0.0146	2

V1	WW2		Clock	
N alleles	S alleles	N alleles		
2	13	7	270.10	
3	16	4	271.00	
3	14	6	271.60	
0	13	7	271.30	
2	14	6	270.40	
3 NA	NA		271.00	
2	14	6	270.10	
0	16	4	271.15	
2	16	4	271.30	
4 NA	NA		270.70	
2	15	5	271.30	
11	14	6	270.55	
3	14	6	271.45	
3	15	5	271.15	
9	15	5	270.85	
12	12	8	270.25	
9	14	6	271.00	
2	18	2	270.55	
6	16	4	271.00	
12	10	10	271.60	
5	12	8	271.45	
3 NA	NA		270.40	
9	8	12	270.70	
13	4	16	270.40	
16	4	16	271.00	
16	6	14	271.60	
15	1	19	270.85	
18 NA	NA		270.55	
17	1	19	271.00	
19	4	16	271.00	
20	4	16	271.75	
18	2	18	270.70	
18	2	18	271.00	