Genetic patterns of Black-tailed Godwit populations and their implications for conservation

Trimbos, K.B.

CHAPTER

6

Patterns in nuclear and mitochondrial DNA reveal historical and recent isolation in the Black-tailed Godwit *(Limosa limosa)*

PLOS One (Submitted)

Krijn B. Trimbos, Camiel Doorenweerd, Ken Kraaijeveld, C.J.M. Musters, Niko M. Groen, Peter de Knijff, Theunis Piersma and Geert R. de Snoo

Abstract

On the basis of morphological differences, three subspecies of Black-tailed Godwit (*Limosa limosa*) have been recognized (*L. l. limosa, L. l. islandica* and *L. l. melanuroides*). In previous studies mtDNA sequence data showed minimal genetic divergence between the three subspecies and an absence of substructuring within *L. l. limosa*. Here, population genetic structure and phylogeographic patterns have been analyzed using COI, HVR1 and HVR2 mtDNA sequence data as well as 12 microsatellite loci (nuDNA). The nuDNA data suggest genetic differentiation between *L. l. limosa* from Sweden and The Netherlands, between *L. l. limosa* and *L. l. islandica*, but not between *L. l. limosa* and *L. l. melanuroides*. However, the mtDNA data were not consistent with the nuDNA pattern. mtDNA did support a split between *L. l. melanuroides* and *L. l. limosa*/*L. l. islandica* and also demonstrated two *L. l. limosa* haplotype clusters that were not geographically isolated. This genetic structure can be explained by a scenario of isolation of *L. l. melanuroides* from *L. l. limosa* in Beringia during the Last Glacial Maximum, possibly followed by recent introgression. During the Pleistocene separation of *L. l. islandica* from *L. l. limosa* occurred, followed by colonization of Iceland by the *L. l. islandica* during the Holocene. Within *L. l. limosa* founder events, followed by population expansion, took place during the Holocene also. According to the patterns observed in both markers together and their geographic separation, we propose that the three traditional subspecies indeed represent three separate genetic units.

Keywords DNA barcode · Incomplete lineage sorting · Microsatellites · Population Genetics · Pleistocene · Subspecies

Introduction

Until a few centuries ago, breeding Black-tailed Godwits (*Limosa limosa* Linneaus, 1785) were confined to raised bogs, moorlands, lake margins and damp grassy depressions in steppe (Beintema *et al.* 1995, Haverschmidt 1963). Since the early Middle Ages the bog habitats in north-western Europe became converted into increasingly nutrient-rich meadows for dairy farming. Black-tailed Godwits were probably quick to exploit this new opportunity and as a result the number of breeding pairs in The Netherlands alone increased to approximately 120,000 in 1967 (Mulder 1972). However, over the last few decades further agricultural intensification with ever earlier mowing dates has led to low recruitment (Kleijn *et al*. 2010, Schekkerman *et al.* 2008). In addition, urbanization of rural areas has led to fragmentation of their breeding habitat. As a result, the mainland European breeding population has been in decline over the last 40 years (Bijlsma *et al.* 2001, Birdlife international 2004, Zwarts *et al.* 2010, Schekkerman *et al.* 2008). This has prompted the IUCN to qualify the species as Near-Threatened (www.birdlife.org/datazone/speciesfactsheet.php?id=3003).

Black-tailed Godwits are migratory shorebirds breeding mainly in temperate and boreal lowlands. Their breeding range extends across Eurasia, from Iceland to Kamchatka and Sakhalin (Cramp and Simmons 1982). Currently, three subspecies are recognized within this range (Figure 1): the European Black-tailed Godwit (*L. l. limosa*), Icelandic Black-tailed Godwit (*L. l. islandica*) and Asian Black-tailed Godwit (*L. l. melanuroides*) (Cramp and Simmons 1982). These subspecies have been distinguished on the basis of morphological traits. *L. l. islandica* has a shorter bill and tarsus and has more extensive rufous-cinnamon and barred plumage than *L. l. limosa*, while *L. l. melanuroides* is distinctly smaller compared with *L. l. limosa* (Cramp and Simmons 1982, Roselaar and Gerritsen 1991). Aside from morphological differences, the subspecies also differ in breeding range and migratory routes, although there is some overlap (Gill *et al.* 2007). The breeding range of nominate *L. l. limosa* extends from Britain to West Russia. *L. l. islandica* breeds mainly on Iceland, with some breeding pairs occasionally found in Scotland and Northern Norway. *L. l. melanuroides* breeds at isolated locations in Russia east of the Yenisey river. *L. l. limosa* winters in parts of southern Europe and south-west Asia, but mainly in sub-Saharan Africa. *L. l. islandica* migrates to Britain, western France, The Netherlands and Iberia. The wintering grounds of *L. l. melanuroides* are in south-east Asia, from the Bay of Bengal to Taiwan, the Philippines and Australia (Cramp and Simmons 1982, Gill *et al.* 2007).

Höglund *et al.* (2009) found slight diagnostic differences between the subspecies on the basis of mitochondrial DNA (mtDNA) sequence data, but found no population structure within *L. l. limosa*. Although they had sequenced part of the highly variable control region (CR) of the mtDNA, they used a relatively conserved part in their analyses (Liebers *et al.* 2001, Ruokonen and Kvist 2002). Using microsatellite markers (nuDNA), Trimbos *et al.* (2011) found moderate levels of genetic variation among Black-tailed Godwits breeding in The Netherlands, and also failed to detect any form of population structure. This suggests either that fragmentation of Black-tailed Godwit breeding populations is too recent for lineage sorting to be complete, or that gene flow has not been restricted. However, genetic structure has yet to be studied throughout the entire breeding range of the Black-tailed Godwit.

Owing to its four times smaller effective population size, mtDNA exhibits faster lineage sorting compared with nuDNA (Moore 1998, Rubinoff *et al.* 2006, Zink and Barrowclough 2008). This difference in effective population size is attributed to the different ways in which the two genomes are inherited. Nuclear DNA is diploid, and recombined between both parents in every generation, whereas mtDNA is haploid and only inherited maternally. In theory, then, mtDNA could thus reflect changes in population structure faster. It has been argued, however, that the best measures of population genetic structure derive from the accumulated signals from multiple loci (Edwards and Bensch 2009), while the entire mtDNA is effectively a single locus. With this in mind, we used a combination of both nuDNA and mtDNA data to account for the shortcomings of each DNA type (Rubinoff and Holland 2005, Mantooth and Riddle 2011). More specifically, we first determined the DNA barcode, part of the mitochondrial Cytochrome C Oxidase I (COI) gene, for a subset of samples. Today, there is a large and growing database of COI barcodes (Barcoding of Life datasystems, www.boldsystems.org), including barcodes for many bird species (Schindel *et al.* 2011). COI data allowed for easy comparison of the results from our samples with those of other studies. Secondly, we used next-generation sequencing on the Illumina HiSeq platform to determine primer sites for the amplification of the hypervariable regions HVR1 and HVR2 of the mitochondrial control region (CR). For the nuDNA data we used a set of 12 microsatellite markers (Verkuil *et al.* 2009)used previously in Trimbos *et al.* 2011.

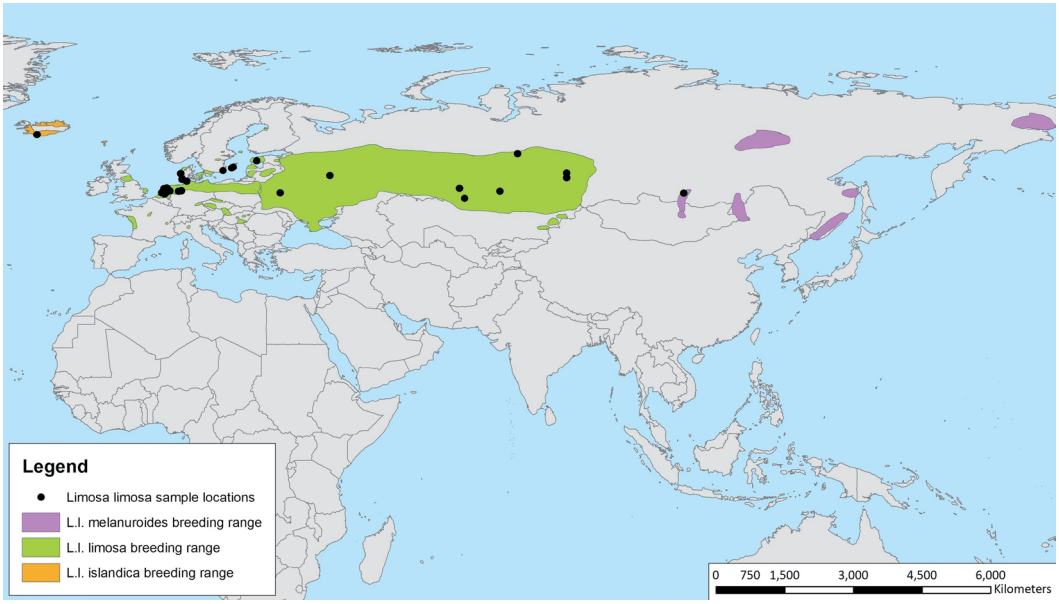

Figure 1. Sample locations of the Black-tailed Godwit *Limosa limosa*. *L. l. limosa*: the Netherlands, Mid-Germany, Northern Germany, Denmark, Sweden, Belarus/Moscow, Kazachstan/SW Russia; *L. l. islandica*: Iceland; and *L. l. melanuroides*: Eastern Russia/Selanga delta.

Material and methods

Sample collection and DNA extraction

Samples were collected between 1991 and 2010 from sites across the *Limosa limosa* breeding range (Figure 1). In The Netherlands, blood samples were collected as described in Trimbos *et al.* (2011). Other blood samples, previously collected (in Sweden, Russia/Moscow, Kazakhstan, western Russia, Iceland, Eastern Russia/Selenga Delta and Canada) by Höglund *et al.* (2009), were made available by the University of Groningen, where they were stored (Table 1). Blood was stored in 97% alcohol at -70 ˚C.

Table 1. Sample location, exact geographical location, number of samples per geographical location used for microsatellite analysis (nuDNA), number of samples per sample location used for microsatellite analysis (msat pooled), number of samples per geographical location used for mtDNA analysis (mtDNA), number of samples per sample location used for mtDNA analysis (mtDNA pooled) and the *Limosa* species or *Limosa limosa* subspecies per sample location.

Sample location	Exact geographical location	nuDNA	nuDNA pooled	mtDNA	mtDNA pooled	Limosa species/ subspecies
Netherlands (140)	Eemnespolder/Arkemheen	24	140	6	46	*Limosa limosa limosa*
	Grote Zoeterwoudse polder	11		4		*Limosa limosa limosa*
	Vijfheerenlanden	10		4		*Limosa limosa limosa*
	Uitdam	11		4		*Limosa limosa limosa*
	Polder Zeevang	11		4		*Limosa limosa limosa*
	Normerpolder	7		4		*Limosa limosa limosa*
	Overijssel/Zwolle	10		4		*Limosa limosa limosa*
	South-west Frysland	38		12		*Limosa limosa limosa*
	Idzegea	18		4		*Limosa limosa limosa*
Germany (35)	Mid-Germany, Schneckenbruch	3	35	3	13	*Limosa limosa limosa*
	Mid-Germany, Dummer	20		6		*Limosa limosa limosa*
	Northern Germany, Fohr	11		3		*Limosa limosa limosa*
	Northern Germany, Meggerdorf	1		1		*Limosa limosa limosa*
Denmark (11)	Tipperne	11	11	4	4	*Limosa limosa limosa*
Belarus (6)	Belarus	3	6	3	6	*Limosa limosa limosa*
	Moscow	3		3		*Limosa limosa limosa*
Sweden (42)	Kristianstad/Faludden/ Hummelbosholm/Oland	42	42	4	4	*Limosa limosa limosa*
Kazakhstan, SW Russia (23)	Novosibirsk	5	23	2	8	*Limosa limosa limosa*
	Lake Ubinsky	2		0		*Limosa limosa limosa*
	Lake Sharkol	5		2		*Limosa limosa limosa*
	Lake Baituma	2		1		*Limosa limosa limosa*
	Lake Big Aksuhat	1		1		*Limosa limosa limosa*
	Lake Shoskaly	2		1		*Limosa limosa limosa*
	Juganski	6		1		*Limosa limosa limosa*
Iceland (27)	W. Iceland	27	27	5	5	*Limosa limosa islandica*
Eastern Russia (3)	River Selenga Delta	3	3	3	3	*Limosa melanuroides*
Canada (2)	Churchill, Manitoba	2	2	2	2	*Limosa heamastica*

Additionally, eggshells were obtained between 2008 and 2010 (Trimbos *et al.* 2009) in The Netherlands, Germany, Belarus and Denmark, all breeding areas of *L. l. limosa* (Table 1). Eggshell remains were collected in the nest (after hatching) and were individually stored in plastic bags at room temperature. DNA was extracted from 6-10 µl of blood using ammonium acetate (Richardson *et al.* 2001) or the Qiagen DNeasy Blood and Tissue Kit according to the manufacturer's protocol (Qiagen 2003). DNA from eggshell membranes was also extracted using the Qiagen DNeasy Blood and Tissue Kit (Qiagen 2003), with minor modifications as described by Trimbos *et al.* (2009). Publicly available sequences from the Barcoding of Life Database (BOLD) were used to supplement the COI barcodes and to provide an outgroup for the COI tree. The Hudsonian Godwit *Limosa haemastica*, an arctic-breeding godwit of Canada and Alaska, was used as outgroup for the HVR analysis.

Microsatellite analysis

A total of 289 birds from 10 different breeding locations were genotyped at 12 microsatellite loci. These 12 loci (LIM3, LIM5, LIM8, LIM10, LIM11, LIM12a, LIM24, LIM25, LIM26, LIM30, LIM33) were specifically developed for Black-tailed Godwits (Verkuil *et al.* 2009). A Fisher's exact test for linkage disequilibrium was carried out using all 289 samples, with 1,000 dememorization steps, 100 batches and 1,000 iterations per batch (GENEPOP web version 4.0; Raymond and Rousset 1995). Deviations from Hardy-Weinberg equilibrium and heterozygote excess or deficiency were tested for each locus and sampling location separately using 1,000 dememorization steps, 100 batches and 1,000 iterations per batch (GENEPOP; Raymond and Rousset 1995). Bonferroni correction for multiple testing was applied (Rice 1989). To detect scoring and amplification errors, we employed MICRO-CHECKER with a 95% confidence interval over 10,000 runs (Oosterhout *et al.* 2004).

For each location, observed (H_o) and expected (H_e) heterozygosities and inbreeding values (F_{IS}) were estimated using ARLEQUIN 3.11 (Excoffier *et al.* 2005) set at 20,000 permutations. An analyses of molecular variance (AMOVA) was performed, allowing variance among sample locations (V_a), variance within sample locations (V_b) and residual variance to be computed (V_c), using ARLEQUIN with 20,000 permutations, followed by Bonferroni correction. Additionally, D was calculated using SPADE (Cho *et al.* 2008), as recent studies have shown that this statistic provides more accurate estimates of genetic differentiation than Fst (Jost 2008; Meirmans and Hedrick 2011). The number of private alleles was determined using CONVERT 1.31 (Glaubitz 2004). FSTAT 2.9.3.2 (Goudet 1995) was used to calculate allelic range, number of alleles per sample location and allelic richness per sample location. To correct for sample size, this program uses the rarefaction index, as described by Hurlbert (1971).

Historic and recent gene flow were investigated using MIGRATE v 3.2.7 (Beerli and Felsenstein 1999;2001) and BAYESASS (Wilson and Rannala 2003). MIGRATE uses coalescent theory in combination with maximum likelihood or Bayesian calculations and MCMC sampling to estimate the relative effective population size (4Neµ) and asymmetric gene flow *M* (m/µ) over approximately 4Ne generations in the past (Beerli 2008). We ran MIGRATE three times using the Bayesian inference; run 1 sampled 200,000 values (1 long chain, 5000 recorded steps, an increment of 10 and 4

replicate chains), run 2 sampled 400,000 (1 long chain, 5000 recorded steps, an increment of 10 and 8 replicate chains) and run 3 sampled 800,000 values (1 long chain, 8000 recorded steps, an increment of 10 and 10 replicate chains). Four-chain heating at temperatures of 1, 1.5, 3 and 10000 was implemented to increase the efficiency of the MCMC. 4Neμ and *M* were compared between the three runs, as were the Bayesian posterior distribution graphs to ascertain whether chains had converged. Here, the results from the longest run are presented, as posterior distribution graphs showed convergence for most of the pairwise comparisons and 4Neμ and *M* values were similar between runs. BAYESASS uses a Bayesian approach and MCMC sampling to generate *m* values which reflect gene flow over the last 5 generations (Wilson and Rannala 2003). BAYESASS was run as described by Chiucchi and Gibbs (2010), with one small modification. The gene flow results of the short run that fit the data best according to a Bayesian deviance measure (Chiucchi and Gibbs 2010) were not comparable with the results of the longer run with the same parameter settings, indicating that with these settings the runs did not converge. We therefore chose the run that fit the data second best, as this run did show similar results in longer runs with the same parameter settings. It is these data that are presented here.

STRUCTURE 2.3.1 (Pritchard *et al.* 2000) was used to cluster genotypes from all sampling locations. We determined the deltaK (Structure Harvester), a calculation of the second-order rate of change in log likelihood Ln P(X|K), as recommended by Evanno *et al.* (2005). The most likely number of genetic clusters (K) in our sample set was also investigated by determining the maximum average log likelihood Ln P(X|K). Values computed with both methods were plotted using Structure Harvester 0.56.3 (Dent 2009, web version). The Structure model was run using admixture and correlated allele frequencies. Additionally, the LOCPRIOR model, incorporated into STRUCTURE 2.3.1, was used. This model assumes that individuals sampled close together are often from the same population and can assist in the clustering when population structure is weak. The program was run 5 times with a burn-in period of 200,000 iterations and a length of 1,000,000 MCMC iterations for K (1-13). Convergence was assessed by checking whether the graphs provided by the program reached equilibrium before the end of the burn-in phase. CLUMPP was used to estimate the number of identical repeat runs per K. The output of CLUMPP was accordingly used to generate graphs from the STRUCTURE results using Microsoft Excel.

A Mantel test with 9999 permutations was performed using GENALEX 6.2 (Peakall and Smouse 2006) to test for correlation between the genetic and geographic distance matrices (Smouse and Long 1992, Smouse *et al.* 1986).

Mitochondrial DNA sequencing

To identify suitable primer sites around the hypervariable sites (HVR1 and HVR2) in the control region of the mtDNA, we sequenced the entire mtDNA of three *L. l. limosa* samples (from The Netherlands, Sweden and SW Russia) at low coverage. For each sample, 1000 ng of genomic DNA was sheared to 500bp fragments using a Covaris S2. These fragments were end-repaired and fitted with an A-overhang at the 3' end using NEBNext TruSeq. Adapters were ligated to these

fragments, after which they were sequenced on an Illumina HiSeq2000. The resulting reads were aligned against the complete mitochondrial sequence of the Ruddy Turnstone *Arenaria interpres* (Paton *et al.* 2002) using Stampy (http://www.well.ox.ac.uk/project-stampy). Barcoded DNA pools sequenced on part of a single lane of an Illumina HiSeq resulted in 817,335, 6,804,981 and 3,273,078 paired-end reads from *L. limosa* samples from the Netherlands, Sweden and SW Russia, respectively. Alignment of the Illumina reads to the *A,interpres* mtDNA with the substitution rate set to 0.1 resulted in 982, 10,068 and 806 aligned reads, respectively. These covered the mtDNA genome 0.58, 9.37 and 2.6 times, respectively. A consensus sequence was constructed using Samtools pileup (http://samtools.sourceforge.net/). On the basis of this consensus sequence, primers were developed amplifying the first and third domain of the *L. limosa* CR (5'-3'; F-primer: L13F – AGCAGTTCCTGCTTGGCTTT, R-primer: L13R – GCAAGTTGTGCTAGGGGTTT and 5'-3'; F-primer: L23F – TTCAAGTGTCCGGGGAATCA, R-primer: L23R TTTGTCTCTGGGTGCATGGG). As sequencing with L13F and L23R proved to be problematic owing to long T-trains and CAAACAAAA repeats, further sequencing was performed unidirectionally using only primers L13R and L23F. For HVR1 and HVR2, 649 bp were sequenced in 91 samples, including 81 *L. limosa* individuals from 23 different *L. l. limosa* breeding locations, five *L. l. islandica* individuals from Iceland, three individuals *L. l. melanuroides* from Eastern Russia and two *L. haemastica* individuals (Table 1). However, for other HVR1 and HVR2 analysis five sequences of poor quality, including the two samples from *L. haemastica* were excluded, adding up to a sample set of 78 samples from *L. l. limosa* breeding locations, five *L. l. islandica* from Iceland and three *L. l. melanuroides* from Eastern Russia.

The universal COI mitochondrial barcode region was amplified using primers BirdF1, BirdR1 and BirdR2 with the addition of M13 tails (Hebert *et al.* 2004). A cocktail of all three primers was used to increase PCR success rate. A section of 658 bp of the COI gene was sequenced for a subset of 56 samples, which included 52 individuals from several *L. l. limosa* breeding locations, three *L. l. islandica* from Iceland and one *L. l. melanuroides* from Eastern Russia.

PCR amplification reactions for L13 and L23 primer pairs were carried out in a total volume of 25µl consisting of 10 ng genomic DNA, 2.5µl PCR Buffer 10x including 15 mM $MgCl_2$, 2.5mM dNTP, 110 pmol of each primer, 1.25 U Taq DNA polymerase (Qiagen) and 18.8 µl DNA mQ water. For COI the same volume and PCR mix reagents were used with the exception of the amount of primer, which was now 250 pmol of each primer (M13F-BirdF1, M13R-BirdR1 and M13R-BirdR2). PCR was conducted on a BIORAD S1000 thermal cycler using the following PCR program: 94°C for 3 min; 40 cycles of 94°C for 15 s, locus-specific Ta 30 s, 72°C 40 s; 72°C for 5 min. Ta was 50°C for COI and 58°C for L13 and L23. With each PCR a negative control was included and sequenced to check for contamination issues. Sequencing was outsourced to Macrogen Europe. Forward and Reverse chromatograms were combined in Sequencer v4.10.1 (Gene Codes Coorporation), checked manually for ambiguities, exported as FASTA files and aligned using BioEdit v7.0.9 (Hall 1999). All novel sequences generated for this study are deposited at GenBank (accession numbers JQ657268-JQ657500).

Mitochondrial DNA analysis

For HVR1 and HVR2 the number of haplotypes, haplotype diversity and nucleotide diversity were calculated using dnaSP v5.0 (Librado and Rozas 2009), with gaps excluded as potential sequence variability. To detect past population expansions we calculated Fu's F_S statistic and Tajima's D-test (Fu 1997, Tajima 1989). To test for background selection Fu and Li's D* and F* statistics were used (Fu and Li 1993). To obtain pairwise Φst between sampling sites, pairwise Juke and Cantor distances and haplotype frequencies were calculated in ARLEQUIN 3.11 (Excoffier *et al.* 2005) with 20,000 permutations. A median-joining haplotype network was constructed using NETWORK v. 4600 (Fluxus-engineering).

DNA barcodes are available for 91% of all bird species (Schindel *et al.* 2011), allowing for a comparison of the genetic variation of the mtDNA within *Limosa limosa* with other bird species (Kerr *et al.* 2007, Schindel *et al.* 2011). As DNA barcoding aims to identify species, the BOLD data structure does not recognize subspecies. However, subspecies clusters were recognized nonetheless through our own added subspecies COI sequence data and comments in the 'notes' field in some BOLD records. Phylogenetic analysis of HVR1 and HVR2 was performed using two different approaches, Bayesian and maximum likelihood, using *L. haemastica* (CAN) as outgroup. For this analysis, samples with missing sequence data were included with this data part encoded as 'missing data'. The Akaike Information Criterium in MrModeltest v2.3 (Nylander 2004) concluded that the HKY + G model was most suitable for the combined HVR1 and 2 dataset. MrBayes (Huelsenback and Ronquist 2001) was subsequently used for the Bayesian analysis, with the HKY + G model, a melting temperature of 0.01, two runs of four chains each, a burn-in of 1 million and a total of 10 million generations. PhyML (Guindon and Gascuel 2003) was used for the maximum likelihood analysis, with the HKY85 substitution model and 10,000 bootstrap generations. For COI a neighbor-joining tree was created using PAUP* 4.10b with uncorrected P distance (Srivathsan and Meijer 2011).

Results

Microsatellite analysis (nuDNA)

A total of 132 different alleles were amplified. The number of alleles per locus ranged from 4 to 15, with no more than 2 alleles per individual. After sequential Bonferroni correction the breeding populations in The Netherlands showed a significant global heterozygote deficit at 6 loci, indicating low heterozygosity in this population. No significant linkage disequilibrium was found between any of the loci after sequential Bonferroni correction. MICROCHECKER detected no null alleles at any of the loci in the complete dataset.

For each sampling location, Table 2 reports the absolute number of alleles, allelic richness, F_{IS}, and private alleles. Neither *L. l. islandica* nor *L. l. melanuroides* showed the presence of private alleles.

Table 2. Sample location and *Limosa limosa* subspecies; number of HVR1 and HVR2 sequence alignments (n), nucleotide diversity (pi), haplotype diversity (h), number of haplotypes (nh) for mtDNA; and number of individuals (n), absolute number of alleles (A), allelic richness (A_R), number of private alleles (P_a) and inbreeding coefficient (F_{IS}) for microsatellite fragment analysis (Msats).

Sample location / *L. limosa* subspecies	mtDNA (n)	pi	h	nh	Msats (n)	A	A_R	P_a	F_{IS}
Netherlands *Limosa limosa limosa*	46	0.006	0.896	16	140	123	2.689	11	0.041*
Mid-Germany *Limosa limosa limosa*	9	0.007	0.972	8	23	84	2.673	0	-0.023
Northern Germany *Limosa limosa limosa*	4	0.007	1.000	4	12	70	2.582	0	0.072
Denmark *Limosa limosa limosa*	4	0.009	0.833	3	11	68	2.579	1	0.002
Belarus/Moscow *Limosa limosa limosa*	6	0.009	1.000	4	6	55	2.581	0	0.189*
Sweden *Limosa limosa limosa*	4	0.002	0.500	2	42	100	2.656	3	0.034
Kazachstan/SW Russia *Limosa limosa limosa*	8	0.005	0.929	6	23	97	2.695	4	-0.002
Iceland *Limosa limosa islandica*	5	0.008	0.900	4	27	62	2.355	0	0.054
Eastern Russia *Limosa limosa melanuroides*	3	0.010	0.667	2	3	41	2.667	0	0.143

F_{IS} values were significantly different from zero in The Netherlands and Belarus. AMOVA calculations showed significance for all the calculated variances (3% Va = 0.03 P < 0.0001, 3% Vb = 0.04 P = 0.0001 and 94% Vc = 0.06 P < 0.0001). The molecular variance present in the sample set was explained for 3% by differences between sample locations. An additional 3% of the variance was explained by differences between individuals within locations. The remaining 94% was randomly distributed over populations, indicating the existence of genetic differentiation, although small, between populations. D supported differentiation between samples from Iceland and the other sampling locations (Table 3). Also, D indicated weak but significant differentiation between Dutch and Swedish samples (Table 3).

Table 3. Below the diagonal: D values for the microsatellite loci; above the diagonal: pairwise Φst for mtDNA HVR1 and HVR2 sequences. CIs not overlapping with zero for D values and significant P values after sequential bonferroni correction for Φst are indicated by *.

	Netherlands	M Germany	N Germany	Denmark	Belarus	Sweden	Kaz/W Rus	Iceland	E Russia
Netherlands	-	-0.03529	-0.06135	-0.04698	-0.05796	0.23901	-0.00394	0.53332*	0.91115*
M Germany	0.005	-	-0.13251	-0.11034	-0.09773	0.32468	-0.05243	0.51159	0.91407
N Germany	0.026	0.022	-	-0.21049	-0.11098	0.33619	-0.13143	0.47302	0.92038
Denmark	0.009	0.029	0.018	-	-0.06555	0.30287	-0.09829	0.44205	0.90956
Belarus	0.000	0.039	0.037	0.030	-	0.39894	-0.09067	0.51269	0.92515
Sweden	0.022*	-0.010	0.036	0.027	0.018	-	0.32384	0.59999	0.97015
Kaz/SW Rus	0.011	0.002	0.019	0.019	-0.000	0.017	-	0.54314	0.93078
Iceland	0.106*	0.088*	0.111*	0.094*	0.175*	0.134*	0.129*	-	0.90610
E Russia	-0.071	-0.061	-0.004	-0.040	-0.127	-0.042	-0.093	0.081	-

Results from STRUCTURE strongly supported a scenario with four genetic clusters, three within *Limosa limosa* and one comprising *Limosa haemastica*. The maximum average log likelihood Ln P(X|K) showed a maximum at K= 4 (Figure 2). Birds from Iceland (*L. l. islandica*) were assigned to a

separate cluster. Birds from the breeding range of *L. l. limosa* were assigned to two different clusters, hereafter clusters 1 and 2 (Figure 2). Birds from The Netherlands were assigned to cluster 1 only. Assignment of the other *L. l. limosa* populations was more ambiguous, with individuals and populations being assigned to either cluster 1 or both clusters 1 and 2. Only in the Swedish populations did assignment values for cluster 2 exceed 60%. Eastern Russian birds (*L. l. melanuroides*) were not recognized as a distinct genetic entity, but clustered together with birds from *L. l. limosa* breeding locations. Mantel tests detected significant correlation between genetic distance and geographic distance (P = 0.006), but not when Icelandic birds were excluded (P = 0.313).

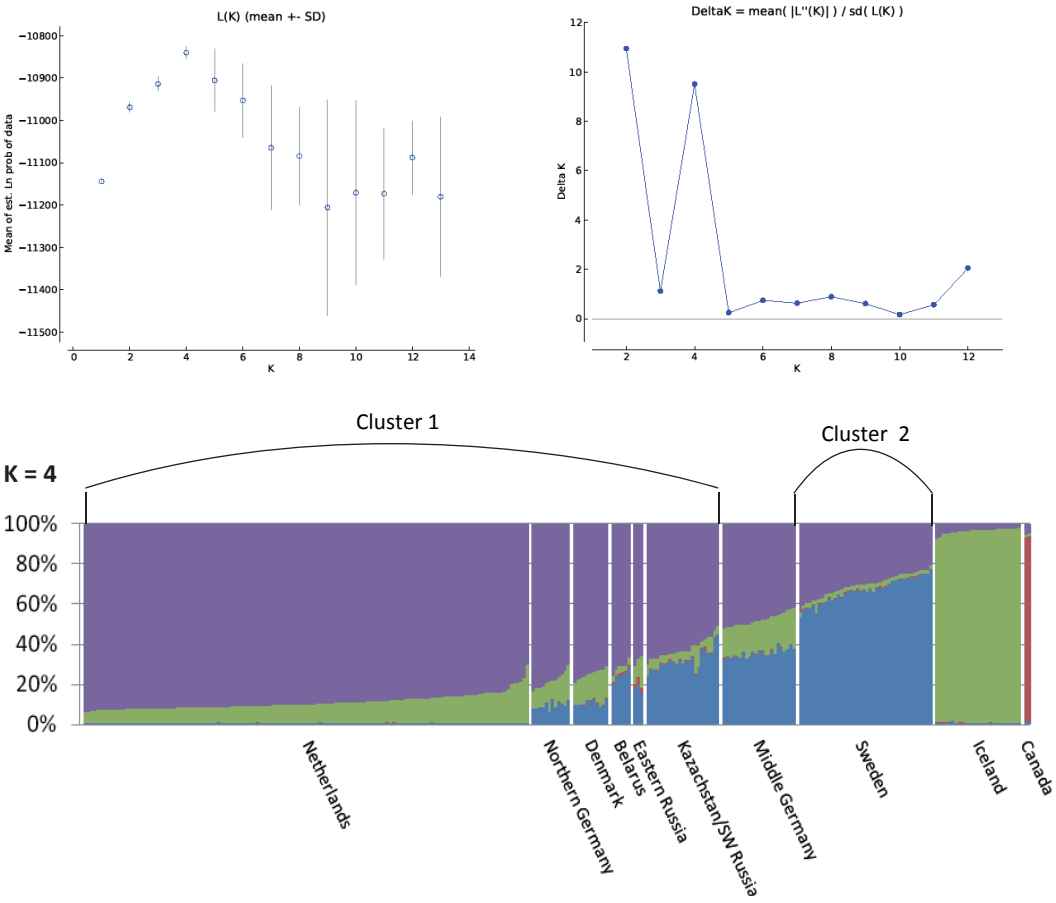

Figure 2. Above: mean log likelihood Ln P(X|K) and DeltaK as a function of the number of genetic clusters (K) averaged over 5 consecutive STRUCTURE runs for each K (error bars indicate one standard deviation). Below: representation of the assignment values, estimated relative contribution of each member of the population to that individual's microsatellite-based genome, per individual at the different sample locations for K=4. Within *L. l. limosa* 2 clusters according to assignment values are depicted.

Recent gene flow estimates demonstrated migration from The Netherlands to all other sampling areas except Iceland (Table 4). Historic gene flow also showed high emigration rates from The Netherlands towards all other areas, including Iceland (Table 5).

Table 4. Bayesian estimates of recent gene flow among sample locations according to Bayesass. Migration is represented as the percentage of the total amount of migration towards a sample location with 95% confidence intervals. Rows represent emigration rates, columns immigration rates.

	Netherlands	M Germany	N Germany	Denmark	Belarus	Sweden	Kaz/W Rus	Iceland	E Russia
Netherlands	99.26 (97.80-99.92)	27.65 (21.01-32.22)	21.03 (10.51-30.39)	18.20 (8.04-29.44)	13.21 (3.43-27.24)	27.53 (20.84-31.70)	27.40 (20.10-32.39)	0.18 (0.00-1.55)	5.31 (0.32-16.21)
M Germany	0.04 (0.00-0.30)	68.05 (66.70-71.35)	1.17 (0.00-5.61)	1.28 (0.00-5.66)	1.77 (0.00-7.13)	0.27 (0.00-1.74)	0.49 (0.00-3.04)	0.12 (0.00-1.15)	2.43 (0.00-9.81)
N Germany	0.16 (0.00-0.73)	0.54 (0.00-3.00)	69.02 (66.73-75.09)	1.42 (0.00-6.90)	1.81 (0.00-7.90)	0.29 (0.00-1.77)	0.46 (0.00-2.59)	0.12 (0.00-0.98)	2.48 (0.00-10.32)
Denmark	0.05 (0.00-0.40)	0.49 (0.00-3.19)	1.26 (0.00-5.81)	69.28 (66.75-75.36)	1.71 (0.00-7.45)	0.28 (0.00-1.67)	0.49 (0.00-2.92)	0.14 (0.00-1.16)	2.39 (0.00-9.19)
Belarus	0.04 (0.00-0.34)	0.56 (0.00-3.31)	1.16 (0.00-5.65)	1.36 (0.00-6.33)	70.44 (66.80-78.95)	0.27 (0.00-1.62)	0.43 (0.00-2.57)	0.12 (0.00-1.04)	2.50 (0.00-10.63)
Sweden	0.09 (0.00-0.63)	0.56 (0.00-3.59)	1.54 (0.00-6.52)	3.29 (0.02-9.71)	3.42 (0.02-11.62)	69.97 (67.08-75.28)	1.08 (0.00-5.02)	0.16 (0.00-1.53)	2.44 (0.00-10.28)
Kaz/SW Rus	0.05 (0.00-0.39)	0.50 (0.00-2.85)	1.20 (0.00-5.95)	1.33 (0.00-6.19)	1.94 (0.00-8.01)	0.28 (0.00-1.77)	67.99 (66.71-71.26)	0.14 (0.00-1.33)	2.41 (0.00-9.49)
Iceland	0.19 (0.00-1.32)	0.61 (0.00-3.36)	1.35 (0.00-6.24)	1.40 (0.00-6.77)	1.93 (0.00-7.98)	0.57 (0.00-2.91)	0.63 (0.00-4.01)	98.74 (95.53-99.96)	2.59 (0.00-10.25)
E Russia	0.04 (0.00-0.34)	0.53 (0.00-3.17)	1.11 (0.00-5.88)	1.27 (0.00-5.98)	1.86 (0.00-7.82)	0.26 (0.00-1.72)	0.55 (0.00-3.23)	0.13 (0.00-1.09)	74.97 (67.00-90.68)

Table 5. Bayesian estimates of long-term gene flow among sample locations according to Migrate-n. All pairwise estimates of M (m/u) are shown with 95% confidence intervals. Rows represent emigration rates, columns immigration rates.

	Netherlands	M Germany	N Germany	Denmark	Belarus	Sweden	Kaz/W Russia	Iceland	E Russia
Netherlands	0	10.122 (7.400-12.800)	5.000 (2.267-7.667)	4.778 (2.533-7.000)	3.694 (0.667-6.667)	15.531 (13.267-17.733)	9.193 (4.733-12.867)	10.566 (7.867-13.133)	2.411 (0.133-4.667)
M Germany	2.212 (0.333-4.067)	0	1.325 (0.000-3.267)	1.150 (0.000-3.000)	0.541 (0.000-2.133)	1.764 (0.000-3.533)	1.561 (0.000-3.333)	2.212 (0.200-4.133)	1.121 (0.000-2.933)
N Germany	1.448 (0.000-3.067)	1.257 (0.000-3.067)	0	0.602 (0.000-2.267)	0.951 (0.000-2.733)	1.220 (0.000-2.933)	1.215 (0.000-3.000)	0.921 (0.000-2.667)	0.718 (0.000-2.400)
Denmark	1.367 (0.000-3.000)	0.672 (0.000-2.400)	0.937 (0.000-2.733)	0	0.556 (0.000-2.133)	0.572 (0.000-2.200)	1.783 (0.000-3.533)	0.976 (0.000-2.733)	1.076 (0.000-2.867)
Belarus	0.807 (0.000-2.467)	0.576 (0.000-2.200)	0.905 (0.000-2.667)	0.499 (0.000-2.067)	0	1.057 (0.000-2.800)	1.094 (0.000-2.867)	0.471 (0.000-2.000)	0.571 (0.000-2.200)
Sweden	4.023 (1.933-6.067)	2.811 (0.400-5.200)	1.450 (0.000-3.600)	1.858 (0.000-3.800)	1.017 (0.000-2.867)	0	3.469 (1.000-5.867)	2.190 (0.067-4.133)	0.712 (0.000-2.400)
Kaz/SW Russia	2.133 (0.267-3.933)	1.547 (0.000-3.467)	0.874 (0.000-2.600)	1.320 (0.000-3.067)	0.578 (0.000-2.133)	2.166 (0.200-4.067)	0	0.756 (0.000-2.467)	1.140 (0.000-2.933)
Iceland	2.350 (0.467-4.200)	1.971 (0.000-3.867)	1.209 (0.000-3.000)	1.507 (0.000-3.267)	1.484 (0.000-3.467)	2.475 (0.467-4.467)	2.314 (0.133-4.400)	0	0.301 (0.000-1.733)
E Russia	0.635 (0.000-2.267)	0.792 (0.000-2.533)	0.591 (0.000-2.200)	0.659 (0.000-2.333)	0.447 (0.000-1.933)	0.883 (0.000-2.600)	0.966 (0.000-2.733)	0.497 (0.000-2.067)	0

Mitochondrial analysis (mtDNA)

The subset of COI barcode sequences from our dataset was combined with the public *Limosa* sequences on BOLD as well as several *Limnodromus scolopaceus* and *Limnodromus griseus* sequences as outgroup (Figure 3). According to previous phylogenetic studies (Thomas *et al.* 2004), *Limnodromus* is the closest sister genus of *Limosa*. Genetic distances between COI barcodes have been shown to be a good indicator of phylogenetic relationships (Wilson *et al.* 2011). In the COI Neighbor-Joining tree, the clade containing *L. haemastica* and *L. fedoa* was the nearest sister to *L. limosa*, with 8.3% and 8.5% pairwise distance to each species, respectively.

This makes them both appropriate as outgroup for the HVR phylogenetic analysis. *L. lapponica* was placed as sister to the above, with 10.4% pairwise distance to *L. limosa*. Within *L. l. limosa*, COI sequences were 100% identical for 57 individuals from samples throughout the breeding distribution of *L. l. limosa*. COI sequences were derived from different PCR batches, with samples from diverse sources including blood, eggshell and muscle tissue, from which DNA was extracted by different people and in different laboratory rooms. Moreover, all the public BOLD sequences also consisted of this most common haplotype. Lack of variation due to large-scale contamination issues can thus be ruled out. *L. l. islandica* sequences were placed within the *L. l. limosa* cluster, distinguished by a single diagnostic character. Our *L. l. melanuroides* sequences as well as several BOLD sequences formed a monophyletic sister cluster to *L. l. limosa* and *L. l. islandica*, with minimally 2.0% pairwise distance. However, five BOLD sequences of specimens from the distribution range of *L. l. melanuroides* contained COI haplotypes that differed at a single position from the most common *L. l. limosa* haplotype and fell within that cluster. Some of these specimens were collected in Vietnam and could therefore not be linked to a specific breeding location. However, two were collected at the Selenga river delta area, the same location as our *L. l. melanuroides* samples and a known *L. l. melanuroides* breeding area.

Nucleotide diversity (pi), haplotype diversity (h) and number of haplotypes (nh) are summarized in Table 2. A total of 37 different haplotypes are found within the HVR 1 and 2 dataset. Phylogenetic trees of mitochondrial HVR 1 and 2 derived from Bayesian and Maximum Likelihood analysis are shown in a Bayesian tree (Figure 4). Where branch topology agree between the Bayesian and Maximum likelihood analysis, both support values are displayed on the respective Bayesian tree branches. As with the COI barcode data, both analyses support two monophyletic clades: one containing the individuals from Eastern Russia (PP > 0.95; bootstrap value > 80), the other containing all other individuals (PP > 95%; bootstrap value > 80%). The resolution of the HVR data was much greater than that of COI barcode, however. All Icelandic samples but one were recovered on a monophyletic sister clade to the *L. l. limosa* clade, while a single sample from Iceland (H072) fell within the *L. l. limosa* clade, making *L. l. limosa* and *L. l. islandica* paraphyletic.

Next page:

Figure 3. Neighbor-joining tree based on COI barcode mitochondrial sequences of *Limosa* with *Limnodromus* as outgroup (Thomas *et al.* 2004). Aside from the barcode sequences generated for this study, public sequences for *Limosa haemastica*, *Limosa fedoa*, *Limosa lapponica*, *Limnodromus scolopaceus* and *Limnodromus griseus* available through BOLD were included as well, indicated by their BOLD ID.

Neighbor-joining tree of the genus
Limosa COI sequence
625bp

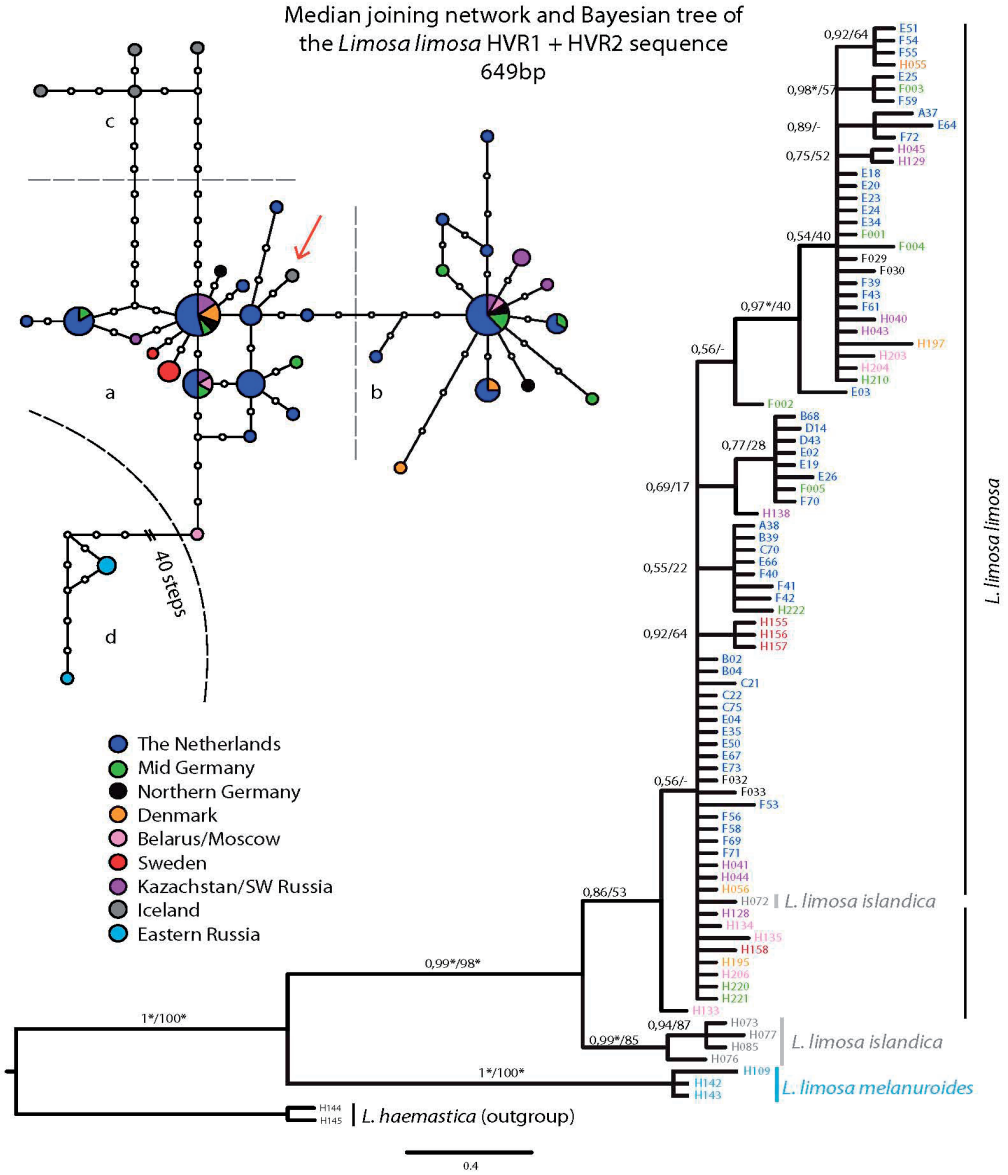

Figure 4. Analysis of the mitochondrial HVR1 and HVR2 for the three *Limosa limosa* subspecies. The colors indicate the sample locations. The support values of the Bayesian (left-hand value) and maximum likelihood (right-hand value) analysis are plotted on the Bayesian tree. Significantly supported (0.95 or higher) branches are indicated with a *. Additionally, a median-joining network of 89 HVR1 and HVR2 mtDNA sequences is depicted. Different clusters are indicated with a/b/c/d. The red arrow indicates one *L. l. islandica* individual (H072) that sorted close to *L. l. limosa* haplotypes.

A split in the *L. l. limosa* clade was significantly supported by the Bayesian analysis. However, these two genetic clusters did not correspond to geographically separated populations.

A median joining network based on the mitochondrial HVR1 and HVR2 sequences is shown in Figure 4. The basic structure of the network strongly resembles the phylogenetic trees. Haplotypes of the individuals from Eastern Russia (d) are separated from all others (abc) by at least 45 steps (Figure 4). Four Icelandic samples (c) are grouped together but separated from sample locations within the *L. l. limosa* breeding range (ab) by at least 11 steps, while one Icelandic sample is found within the *L. l. limosa* cluster (red arrow in Figure 4). The individuals from the *L. l. limosa* breeding locations group into two star-shaped clusters (a and b), with the most common haplotypes separated by eight steps. The two star-shaped clusters do not correspond to geographically separated populations (Figure 4). Swedish *L. l. limosa* individuals belong to cluster a, but display two unique haplotypes. These results are supported by Φst calculations, which showed higher values for pairwise differences between Eastern Russian and all other individuals (Φst values between 0.91-0.97) as compared with pairwise differences between Icelandic and other individuals (Φst values between 0.44-0.54). Φst values between Sweden and other sample locations are moderate (Φst values between 0.24-0.40). Neither Fu's Fs (ranging from -9.47 to 1.61, $P > 0.50$) nor Tajima's D (ranging from -1.32 to 1.32, $P > 0.10$) nor Fu and Li's D* (ranging from -1.01 to 1.29, $P > 0.10$) and F* (ranging from -0.95 to 1.32, $P > 0.10$) are significant for the total population or any of the sampling locations.

Discussion

Three subspecies have been recognized morphologically within *Limosa limosa* (*L. l. limosa*, *L. l. islandica* and *L. l. melanuroides*) and have been confirmed to be genetically identifiable as well in a previous study using the 'conserved domain' of the mitochondrial CR (Höglund *et al.* 2009). Here we confirm this distinction. Nevertheless, the signals found in the nuDNA did not support the split between *L. l. melanuroides* and *L. l. limosa* demonstrated by the mtDNA.

Nuclear DNA showed significant heterozygote deficiency in the Netherlands. MICROCHECKER analysis showed no signs of null alleles within this population, indicating that heterozygote deficiency was not an effect of null alleles. It is also unlikely that it was caused by a Wahlund effect (Wahlund 1928). As previous population genetic research could not detect any genetic population structure among Black-tailed Godwits breeding in different areas in The Netherlands (Trimbos *et al.* 2011), a possible explanation could be that there are few migration events from other locations towards The Netherlands (note recent gene flow analysis and significant F_{IS} value, Table 2). The nuDNA data demonstrate genetic differentiation between *L. l. islandica* on the one hand and *L. l. limosa* and *L. l. melanuroides* on the other. However, no genetic divergence was detected between *L. l. melanuroides* and *L. l. limosa*. Furthermore, within *L. l. limosa*, individuals from The Netherlands and Sweden appear to be divided into two distinct clusters. Individuals from other *L. l. limosa* breeding areas show an admixture of genotypes between the two clusters.

Only three COI barcode haplotypes were found within *L. l. limosa*, 92% of all samples of which contained the same haplotype. The other two subspecies contained only a single haplotype each. The lack of subspecific variation in COI barcode has been noted for other bird species, too, with various explanations being proferred, including selective sweeps or genetic drift through population bottlenecks (Kerr *et al.* 2007). This is probably an artefact of the lower substitution rate in COI compared to the HVR regions of the mtDNA (Buehler and Baker 2005, Wenink and Baker 1996). Even though the resolution exhibited by the COI barcode is less than the resolution of the HVR1 and HVR2 data, the subspecies are distinguishable by both parts of the mtDNA. Interestingly, all three mtDNA regions (COI, HVR1 and HVR2) show genetic differentiation between *L. l. melanuroides* and *L. l. limosa* individuals to be much higher than that between *L. l. limosa* and *L. l. islandica* individuals. Again interestingly, a single individual from Iceland (H072) contains a combined HVR1/HVR2 haplotype that closely resembles that of *L. l. limosa* individuals. To confirm that this was not due to contamination, we re-examined the microsatellite results from this extract. The microsatellite pattern of H072 was unique and contamination of the extract was thus ruled out; the lowest genetic distance found in all pairwise comparisons with H072 was 8 differences. Furthermore, we repeated the HVR PCR and sequencing for this sample twice, with no change in the results. Our results therefore indicate that there might have been recent hybridization events between *L. l. limosa* and *L. l. islandica*. *L. l. limosa* is divided into two large star-like haplotype clusters in the HVR data. These clusters are not supported geographically, as both haplotype clusters are present at nearly all the *L. l. limosa* sample locations. The two *L. l. limosa* haplotype clusters in the HVR mtDNA (Figure 4; cluster a and b) do not completely correspond with the *L. l. limosa* clusters found in the nuDNA (Figure 2; cluster 1 and 2). The two nuDNA clusters in *L. l. limosa* are comprised of the Netherlands, Northern Germany, Denmark, Belarus and Kazachstan/Russia, which make up cluster 1, and Sweden (over 60% assigned), which makes up cluster 2. Mid-Germany shows an almost even admixture between both clusters. So, while the differentiation between *L. l. limosa* and *L. l. islandica* shows similar patterns in mtDNA and nuDNA, differentiation between *L. l. limosa* and *L. l. melanuroides* seems to show opposite patterns in the mtDNA and nuDNA.

Within the mtDNA private haplotypes in *L. l. islandica* do not support a scenario of mitochondrial gene flow between *L. l. limosa* and *L. l. islandica*, but the similarity between *L. l. islandica* and *L. l. limosa* haplotypes might indicate past mtDNA introgression. Indeed, the nuDNA data support historical rather than recent gene flow between *L. l. limosa* and *L. l. islandica*. Furthermore, *L. l. islandica* does not possess private alleles but differs from *L. l. limosa* only by its allele frequencies. Together, the nuDNA and mtDNA thus suggest relatively recent separation of *L. l. islandica* and *L. l. limosa*. The difference between the mtDNA and the nuDNA in regards to the differentiation between *L. l. melanuroides* and *L. l. limosa* was unexpected. The mtDNA exhibited a sharp divergence between *L. l. melanuroides* and the remaining Black-tailed Godwits, while in the nuDNA there was a lack of divergence between *L. l. melanuroides* and the *L. l. limosa*. It is known that programs like STRUCTURE are very conservative in assigning samples from a certain group to a cluster when the sample sizes of such a group is small or sampling scheme is biased (Schwartz and McKelvey 2009). Therefore, these results can most likely be explained by the low sample size of the *L. l. melanuroides* group which have probably biased microsatellite analysis conservatively in regards to the split between *L. l. melanuroides* en *L. l. limosa*. Alternatively, the combined nuDNA and mtDNA

results might suggest a distant split between *L. l. melanuroides* and the other Black-tailed Godwits, with recent gene flow between *L. l. melanuroides* and *L. l. limosa*. Other studies have explained discordant patterns between nuDNA and mtDNA through introgression events (Daly-Engel *et al.* 2012, Hefti-Gautschi *et al.* 2009, Jones *et al.* 2005, Zarza *et al.* 2011). It should be noted, though, that this scenario with respect to *L. l. melanuroides* and *L. l. limosa* would only hold for the *L. l. melanuroides* breeding population at the Selanga River Delta; other *L. l. melanuroides* breeding populations might still be isolated from *L. l. limosa*. Interestingly, this study and another recent study both show the presence of two COI *L. l. melanuroides* haplotype groups, one sorting close to *L. l. limosa* and the other showing more distinct divergence from *L. l. limosa*, at the Selanga River Delta area (Elbourne 2011). These groups might constitute two disjunct *L. l. melanuroides* breeding colonies. An alternate hypothesis would be the presence of a slightly diverged *L. l. limosa* breeding population and a *L. l. melanuroides* breeding population. This would suggest current overlap in *L. l. limosa* and *L. l. melanuroides* breeding sites, making a gene flow scenario between *L. l. limosa* and *L. l. melanuroides*, as proposed here, more likely.

Wenink and Baker (1996) and Buehler and Baker (2005) estimated the mutation rates for HVR1 and HVR2 at around 10% per Myr. For a sequence length of 649 bp this would translate to 6.4×10^{-5} mutations per year, with a range of 3.2×10^{-5} to 9.6×10^{-5}. This results in split estimates of approximately 347 (±174) Ky for *L. l. limosa* vs. *L. l. melanuroides*, 85 (± 43) Ky for *L. l. limosa* vs. *L. l. islandica* and 62 (± 31) Ky for the two mtDNA *L. l. limosa* clusters. This would indicate that the mtDNA population structure, according to HVR1 and HVR2, arose during the Pleistocene. Other studies have also reported the origin of lineage diversity of several bird species to lie within the Pleistocene (Jones *et al*. 2005, Otvall *et al*. 2005, Ronka *et al*. 2008). Iceland was covered in ice during the Weichselien (occuring between 116Ky – 11,5Ky), making it unlikely that *L. l. islandica* (85Ky ago) colonized the island during that period (Adams 1997, Schmitt 2007, www.ngdc.noaa.gov. We hypothesize that the most recent common ancestor of *L. l. islandica* colonized Iceland after the Pleistocene (i.e. in the last 12Ky) and that since then genetic isolation and drift have resulted in the genetic differentiation observed between these subspecies today. Lineage diversification between *L. l. limosa* and *L. l. melanuroides* lineages could have occurred via separate southward or northward founder events. During the Pleistocene the ice sheets that dominated the landscape in Northern Europe and America were absent in large parts of far eastern Russia and there is strong evidence from Beringia and north-eastern Asia that several species of plant and animal survived the last glaciation at high altitudes (Adams 1997, Schmitt 2007, www.ngdc.noaa.gov. We suggest that the ancestral *L. l. melanuroides* became isolated from the remaining Black-tailed Godwit population in the Beringian refugium during periods of glacial cooling in the Pleistocene, resulting in the split in the mtDNA. After glacial conditions alleviated, *L. l. melanuroides* and *L. l. limosa* gene flow might have been regained, with introgression as a result. Northward founder events by two separate *L. l. limosa* lineages subsequently expanding throughout the current *L. l. limosa* breeding range followed by recent isolation and genetic drift could explain the two distinct star-shaped HVR mtDNA haplotype clusters for *L. l. limosa* (a and b). Similar patterns have been found in the Herring Gull *Larus argentatus* complex (Liebers *et al.* 2004). While the mtDNA demonstrated that haplotypes belonging to both cluster a and b were present in all sample locations, the nuDNA shows that *L. l. limosa* individuals from the Netherlands are assigned to cluster 1 only. As the HVR mtDNA shows that *L. l. limosa*

structure is more recent than the divergence between *L. l. islandica* and *L. l. limosa*, one explanation for the different *L. l. limosa* patterns in mtDNA and microsatellites might be incomplete lineage sorting in the microsatellites. Alternatively, the Dutch *L. l. limosa* population expansion that took place during the first half of the 20th century (Beintema *et al.* 1995, Haverschmidt 1963) may have caused introgression among Dutch *L. l. limosa* breeding locations. This could have resulted in the genetic homogenization of the Dutch *L. l. limosa* breeding population in the nuDNA. Additionally, this recent population expansion of Dutch *L. l. limosa* may have resulted in introgression between Dutch *L. l. limosa* individuals and individuals from other *L. l. limosa* breeding locations as well. This is supported by the recent gene flow estimates as well as by most *L. l. limosa* individuals being assigned to the Dutch *L. l. limosa* cluster (cluster 1). Some divergence between Sweden and other *L. l. limosa* sampling locations is shown by the microsatellites. While the Swedish *L. l. limosa* individuals do not share any mtDNA haplotypes with other *L. l. limosa* individuals, they are closely related to other *L. l. limosa* individuals, which might indicate recently restricted gene flow between Swedish *L. l. limosa* and other *L. l. limosa* individuals.

Our data confirm divergence between the three *Limosa limosa* subspecies (cf. Höglund *et al.* 2009). According to the patterns observed and their geographic separation, we propose that the three traditional subspecies should be managed as three separate units. We believe the most likely explanation for the genetic structure found in this study is post-Pleistocene geographical separation of *L. l. islandica* and a distant Pleistocene split of *L. l. melanuroides*. Possibly, *L. l. limosa* and *L. l. melanuroides* have regained secondary contact at Selanga River Delta recently. The two star-shaped haplotype clusters visible in the mtDNA of *L. l. limosa* are most likely the result of one or more successful *L. l. limosa* populations carrying two ancestral haplotypes expanding post-Pleistocene. *L. l. limosa* individuals from The Netherlands may have spread throughout the *L. l. limosa* breeding range, resulting in the two still slightly admixed clusters in the microsatellites. However, incomplete lineage sorting and homoplasy affecting the microsatellite analysis could not be ruled out in this case. Our data highlight the importance of using both nuDNA and mtDNA simultaneously when studying range-wide population genetic structure in birds.

Acknowledgements

Lida Kanters, David Kleijn, Rene Faber, Astrid Kant, Gerrit Gerritsen, Wim Tijsen, Dirk Tanger, Angela Helmecke, Heinrich Belting, Ole Thorup, Jacob Höglund helped collect blood and eggshell membrane samples. We thank Jos Hooijmeijer, Petra de Goeij, Pedro Lourenço and the rest of the University of Groningen Black-tailed Godwit team for their help in collecting eggshells and blood samples in southwest Fryslân and for sharing their blood samples collected at other Dutch Black-tailed Godwit breeding sites. Marco van der Velde arranged laboratory space for DNA extractions in Groningen. Dick Groenenberg, Frank Stokvis, Menno Schilthuizen, Pim Arntzen, Kees Roselaar and Ben Wielstra helped with the interpretation of the data. Finally, we are grateful to Nigel Harle of Gronsveld for brushing up our English.

References

Adams JM (1997) Global land environments since the last interglacial. Oak Ridge National Laboratory, TN, USA. http://www.esd.ornl.gov/projects/qen/nerc.html.

Ballard JW, Chernoff B, James AC (2002) Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology or behavior in *Drosophila simulans*. Evolution, 56, 527-545.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics, 152, 763-773.

Beerli P, Felsenstein J (2001) Maximum liklihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences of the United States of America, 98, 4563-4568.

Beintema A, Moedt O, Ellinger D (1995) Ecologische Atlas van de Nederlandse Weidevogels. Schuyt & Co, Haarlem

BirdLife International (2009) Species factsheet: *Limosa limosa*. Downloaded from http://www.birdlife.org on 29/1/2010

Birdlife International (2004) Birds in Europe: population estimates, trends and conservation status. Cambridge, UK.

Buehler DM, Baker AJ (2005) Population divergence time and historical demography in Red Knots and Dunlins. Condor, 107, 497-513.

Chao A, Jost L, Chiang SC, Jiang Y-H, Chazdon RL (2008) A two-stage probabilistic approach to multiple-community similarity indices. Biometrics, 64, 1178-1186.

Chiucchi JE, Gibbs HL (2010) Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. Molecular Ecology, 19, 5345-5358.

Cramp S, Simmons KEL (1982) *The Birds of the Western Palearctic*. Vol III. Oxford: Oxford University press.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW (2012) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered Scalloped Hammerhead Shark (*Sphyrna lewini*). PLOS One, 7, 1-11.

Dent EA (2009) STRUCTURE HARVESTER version 0.56.3 http://taylor0.biology.ucla.edu/struct_harvest/

Edwards S, Bensch S (2009) Looking forward or looking backwards in avian phylogeography? Molecular Ecology, 18, 2930-2933.

Elbourne R (2011) Masters of Science thesis: COI Barcoding of the Shorebirds: Rates of Evolution and the Identification of Species. Department of Ecology and Evolutionary Biology, Department of Toronto.

Estoup A, Jarne P, Cournuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Molecular Ecology, 11, 1691-1604.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1, 47-50.

Fu Y (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. Genetics, 147, 915-925.

Fu Y, Li WH (1993) Statistical tsts of neutrality of mutations. Genetics, 133, 693-709.

Gill JA, Langston RHW, Alves JA, Atkinson PW, Bocher P, Cidraes Vieira N, Chrockford NJ, Gelinaud G, Groen N, Gunnarsson TG, Hayhow B, Hooijmaijer J, Kentie R, Kleijn D, Lourenco P, Masero, JA, Meunier F, Potts P, Roodbergen M, Schekkerman H, Schroeder J, Wymenga E, Piersma T (2007) Contrasting trends in two black-tailed godwit populations: A review of causes and recommendations. Wader Study Group Bulletin, 114, 43-50.

Glabutz JC (2004) CONVERT: a user friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Molecular Ecology Notes, 4, 309-310.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity, 86, 485-486.

Gunnarsson TG, Potts PM, Gill JA, Croger RE, Gelinaud G, Atkinsson PW, Gardarsson A, Sutherland WJ (2005) Estimating population size in Icelandic Black-tailed Godwits Limosa limosa islandica by colour-marking. Bird Study, 52, 153-158.

Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95–98.

Haverschmidt F (1963) The black-tailed godwit. E.J. Brill. Leiden

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLOS Biology, 2, e312.

Hefti-Gautschi B, Pfunder M, Jenni L, Keller V, Ellegren H (2009) Identification of conservation units in the European Mergus merganser based on nuclear and mitochondrial DNA markers. Conservation Genetics, 10, 87-99.

Höglund J, Johansson T, Beintema A, Schekkerman H (2009) Phylogeography of the Black-tailed Godwit *Limosa limosa*: substructuring revealed by mtDNA control region sequences. Journal of Ornithology, 150,45-53.

Huelsenback JP, Ronquist F (2001) MR-BAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17, 754–755.

Jones KL, Krapu GL, Brrandt DA, Ashley MV (2005) Population genetic structure in migratory sandhill cranes and the role of Pleistocene glaciations. Molecular Ecology, 14, 2645-2657.

Jost L (2008) G_{ST} and its relative do not measure differentiation. Molecular Ecology, 17, 4015-4026.

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, *et al.* (2007) Comprehensive DNA barcode coverage of North American birds. Molecular Ecology Notes, 7, 535-543.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451-1452

Liebers D, Helbig AJ, de Knijff P (2001) Genetic differentiation and phylogeography of gulls in the *Larus cachinnans-fuscus* group (Aves: Charadriiformes). Molecular Ecology, 10, 2447-2462.

Liebers D, de Knijff P, Helbig AJ (2004) The herring gull complex is not a ring species. Proceedings of the Royal Society London B, 271, 893-901.

Mantooth, S. J. and Riddle, B. R. (2011) Molecular Biogeography: The Intersection between Geographic and Molecular Variation. Geography Compass, 5, 1–20.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Molecular Ecology Resources, 11, 5-18.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear-gene trees. Evolution, 49, 718-726.

Moritz C (1994) Defining 'Evolutionarily Significant Units'for conservation. *TREE*, 9, 373- 375.

Paton T, Haddrath O, Baker AJ (2002) Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proceedings of the Royal Society London B, 269, 839-846.

Peakal R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288-295.

Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. Bioinformatics, 14, 817–818.

Pritchard JK, Stphens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Ottvall R, Höglund J, Bensch S, Larsson K (2005) Population differentiation in the redshank (Tringa tetanus) as revealed by mitochondrial DNA and amplified fragment length polymorphism markers. Conservation Genetics, 6, 321-331.

Qiagen (2003) DNeasy Tissue Handbook. Protocol for isolation of total DNA from animal tissues. pp.18-20, QIAGEN. Valencia, California, USA.

Rach J, DeSalle R, Sarkar IN, Schierwater B, Hadrys H (2008) Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proceedings of the Royal Society B-Biological Sciences, 275, 237-247.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity, 86, 248-249.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus sechellensis*). Molecular Ecology, 10, 2263-2273.

Roselaar CS, Gerritsen GJ (1991) Recognition of Icelandic Black-tailed Godwit and its occurrence in The Netherlands. Dutch Birding, 13, 128-135.

Rönkä A, Kvist L, Karvonen J, Koivula K, Pakanen VM, Schamel D, Tracy DM (2008) Population genetic structure in the Themminck's stint *Calidris temminckii*, with an emphasis on Fennoscandian populations. Conservation Genetics, 9, 29-37.

Rubinoff D, Holland BS (2005) Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. Systematic Biology, 54, 952-961.

Rubinoff D, Cameron S, Will K (2006) A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. *Journal of Heredity,* 97, 581-594.

Ruokonen M, Kvist L (2002) Structure and evolution of the avian mitochondrial control region. Molecular Phylogenetics and Evolution, 23, 422-432.

Schekkerman H, Teunissen WA, Oosterveld E (2008) The effect of 'mosaic management' on the demography of black-tailed godwit *Limosa limosa* on farmland. Journal of Applied Ecology, 45, 1067-1075.

Schindel D, Stoeckle M, Milensky C, Trizna M, Schmidt B, Gebhard C, Graves G (2011) Project Description: DNA Barcodes of Bird Species in the National Museum of Natural History, Smithsonian Institution, USA. ZooKeys 152: 87-91. doi: 10.3897/zookeys.152.2473

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Frontiers in Zoology, 4, 1-13.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conservation Genetics, 10, 441-452.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters, 9, 615-629.

Srivathsan A, Meijer R (2011) On the inappropriate use of the Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics, 28, 190-194.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123, 585-595.

Thomas G H, Wills MA, Sze´kely T (2004) A supertree approach to shorebird phylogeny. BMC Evolution and Biology, 28, 1471-2148.

Trimbos KB, Broekman J, Kentie R, Musters CJM, de Snoo GR (2009) Using eggshell membranes as a DNA source for population genetic research. Journal of Ornithology, 150, 915-920.

Trimbos KB, Musters CJM, Verkuil YI, Kentie R, Piersma T, de Snoo GR (2011) No evident spatial genetic structuring in the rapidly declining Black-tailed Godwit *Limosa limosa limosa* in The Netherlands. Conservation Genetics, 12, 629-636.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538.

Verkuil YI, Trimbos K, Haddrath O, Baker AJ (2009) Characterization of polymorphic microsatellite DNA markers in the black-tailed godwit (*Limosa limosa: Aves*). Molecular Ecology Resources, 9, 1415-1418.

Wenink PW, Baker AJ (1996) Mitochondrial DNA lineages in composite flocks of migratory and wintering Dunlins (*Calidris alpina*). Auk, 113, 744-756.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics, 163, 1177-1191.

Wilson JJ, Rougerie R, Schonfeld J, Janzen DH, Hallwachs W, Hajibabaei M, Kitching IJ, Haxaire J, Hebert PDN (2011) When species matches are unavailable are DNA barcodes correctly assigned to higher taxa? An assessment using sphingid moths. *BMC ecology*, 11, 1-14.

Zarza E, Reynoso VH, Emerson BC (2011) Discordant patterns of geographic variation between mitochondrial and microsatellite markers in the Mexican black iguana (Ctensosaura pectinata) in a contact zone. Journal of Biogeography, 38, 1394-1405.

Zink RM (2004) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proceedings of the Royal Society B-Biological Sciences, 271, 561-564.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology, 17, 2107-2121.

Zwarts L, Bijlsma RG, van der Kamp J, Wymenga E (2009) Living on the edge; wetlands and birds in a changing Sahel. KNNV Publishing Zeist.