High genetic diversity and no detectable founder effects following a rapid range expansion of a long-distance migrant (*Acrocephalus scirpaceus*)

Nora M. Bergman[1,2†], Petteri Lehikoinen[3], Edward Kluen[1,2], Staffan Bensch[4], Camilla Lo Cascio Sætre [5], Fabrice Eroukhmanoff[5], Frode Fossøy[6], Petr Procházka[7], William J. Smith[1,2,8], Bård G. Stokke[6], Craig R. Primmer[2,9], Rose Thorogood[1,2], Katja Rönkä[1,2]

[1]HiLIFE Helsinki Institute of Life Science, University of Helsinki, [2]Research programme in Organismal and Evolutionary Biology, Faculty of Biological & Environmental Sciences, University of Helsinki, [3]Finnish Museum of Natural History, University of Helsinki, Finland, [4]Department of Biology, Lund University, Lund, Sweden, [5]CEES, Department of Biosciences, University of Oslo, [6]Norwegian Institute for Nature Research (NINA), Trondheim, Norway, [7]Institute of Vertebrate Biology, Czech Academy of Sciences, [8]Department of Biology, University of Oxford, UK, [9]Institute of Biotechnology, HiLIFE Helsinki Institute of Life Science, University of Helsinki

[†]Correspondence should be addressed to Nora M. Bergman; E-mail: nora.bergman@helsinki.fi

Abstract

Many species are currently experiencing range shifts in response to climate change. The colonisation of new regions can have serious genetic consequences for the shifting population, given that repeated founder events and strong genetic drift are expected to erode genetic variation along the expansion axis. Such a loss of genetic diversity could reduce adaptive potential, slowing or even halting the expansion. These genetic consequences are known to be particularly acute in species with low dispersal capabilities and therefore low levels of gene flow. Whether highly mobile species experience similar genetic bottlenecks as they undergo environment-driven range shifts is, however, less clear. Here we combined ecological and genomic evidence to (1) infer the origin and (2) quantify the genetic consequences of a recent and rapid northward range expansion of the common reed warbler (*Acrocephalus scirpaceus*), a philopatric long-distance migratory passerine capable of high dispersal. Together, historical observations, ringing data, and genomic RAD-seq data covering the species' European breeding range support an expansion origin from southwest. There were no reductions in levels of nucleotide diversity or allelic richness along the colonisation route, and no detectable allele frequency clines persisting from potential past founder effects. The genomic dataset confirmed the occurrence of low but significant levels of geographically continuous population structure across the European range. These results suggest that high dispersal capabilities can help even strongly philopatric species to escape the genetic costs of rapid range expansions. The lack of genetic patterns also highlights the value of combining multiple lines of evidence to reveal range shifts as genomic data alone was not sufficient to detect the occurrence of this large-scale range expansion.

1. Introduction

The current rate of climate and ecosystem change is causing rapid range shifts across taxonomic groups (Parmesan & Yohe, 2003; Chen et al., 2011). Alongside local adaptation and phenotypic plasticity, range shifts allow species to respond to changes in their habitat by tracking suitable environmental conditions that are shifting in space (Waldvogel et al., 2020). To mitigate the threats that environmental change poses to global and local biodiversity, it is crucial to understand both the prerequisites for, and consequences of, range shifts. On one hand, range shifts are expected to have ecological consequences in the recipient communities (Wallingford et al., 2020). On the other hand, the shifting population itself is often subject to various evolutionary forces, including strong genetic drift at the expanding range edge (Excoffier et al., 2009a). Much of the earlier research into the genetics of contemporary range expansions (as opposed to longer, post-glacial timescales) has focused on systems involving human-mediated invasions or threatened species, and less attention has been paid to environment-driven expansions or geographically widespread species (Garnier & Lewis, 2016; Wallingford et al., 2020). The latter, however, form the bulk of the currently shifting populations (MacLean & Beissinger, 2017; Parmesan & Yohe, 2003) and are therefore likely to have major importance in ecosystem functioning. Furthermore, among the studies investigating such environment-driven expansions, most have been on species that remain resident year-round (e.g. wall lizards; Schulte et al., 2013). Migratory species, by comparison, are highly mobile and have greater potential for dispersal (Weatherhead & Forbes, 1994). Highly mobile species could maintain gene flow between the range edge and core areas, and consequently have reduced genetic signatures compared to those currently predicted for range-front populations (Excoffier et al., 2009a). By understanding the history and genomic consequences of contemporary environment-driven expansions, we can better understand the capacity of species to shift their ranges and maintain adaptive potential under the influence of contemporary environmental change.

Following decades of theoretical and empirical work, we know that genetic drift is one of the most important processes in determining evolutionary outcomes during range expansions (Nei et al., 1975; Welles & Dlugosch, 2019). When new areas are colonized, repeated founder effects and small (effective) population sizes can cause strong drift at the expanding range edge (Excoffier, 2004 Excoffier et al., 2009). In the absence of sufficient gene flow from areas further away from the edge, we expect loss of genetic diversity and increasing genetic differentiation along the expansion axis (Austerlitz et al., 1997), as well as stochastic spatial spread of alleles that are not lost due to genetic drift (i.e., allelic surfing) (Edmonds et al., 2004 Klopfstein et al., 2006). These effects of drift are predicted to be more severe in fast expansions compared to slower ones (Garnier & Lewis, 2016), and may have serious consequences for colonising success: the loss of genetic diversity has been shown to cause reduced responses to selection and negative fitness consequences in newly founded populations in both theoretical work (Peischl et al., 2015) and empirical studies (Pujol & Pannell, 2008; Szűcs et al., 2017). In some cases, however, there can also be positive effects if reduced maladaptive gene flow from the range core allows local adaptation to persist (Haldane 1956 Kirkpatrick & Barton, 1997 but see Kottler et al., 2021), or if the stochastically spread alleles happen to be beneficial (e.g. Gralka et al.,

2016; Moreau et al., 2011). The impacts of genetic drift can determine the future of colonising populations, but how severe and long-standing are these consequences in contemporary range shifts?

Numerous empirical range expansion studies have documented the expected consequences of repeated founder effects, both in terms of reduced genetic diversity (ranging from e.g. microalgae; Grupstra et al., 2017, to trees; Keller et al., 2010, and mammals; Garroway et al., 2011, White et al., 2013, Garcia-Elfring et al., 2017) and increasing genetic differentiation in the direction of the colonisation (from insects; Swaegers et al., 2013, cephalopods; Ramos et al., 2018, mammals; Hagen et al., 2015, to birds; Rózsa et al., 2016). However, there are also cases where these expected genetic patterns have been found to be weak or non-existent, many of these coming from species with high dispersal capabilities (Engler et al., 2016; Haase et al., 2019; Heppenheimer et al., 2018; Song et al., 2013b). This suggests that, at least in some cases, higher dispersal capabilities also lead to sufficient levels of gene flow to avoid severe bottlenecks during range expansion. However, in most of these studies, the use of low-density markers (mtDNA and microsatellites), as well as other factors alleviating founder effects (e.g. relatively small expansion scales or hybridisation with other species) limit the reliability of conclusions drawn about fine-scale genomic patterns (see Adams et al., 2023 for a recent example where high-resolution markers allowed the detection of slightly reduced genetic diversity). In addition to the need for high-resolution genomic data, contradictory expectations complicate the prediction of outcomes of environment-driven range expansions in species with high dispersal capabilities. For example, the retention of genetic diversity is expected to be higher during slow expansions (Excoffier et al., 2009b), but sufficient gene flow from areas behind the colonisation front can prevent the erosion of genetic diversity even when the expansion is rapid (e.g. invasive starlings *Sturnus vulgaris*, Berthouly-Salazar et al., 2013). Long-distance dispersal (LDD), a context-dependent but characteristically rare and extreme dispersal event (Jordano, 2017), is a key process that can increase gene flow from more distant areas and alleviate initial founder effects. However, very rare LDD events can have the opposite effect and cause greater loss of genetic diversity compared to short-range dispersal, as isolated populations (with low diversity) being founded ahead of the main expansion wave can hinder the settling of later-arriving individuals (Bialozyt et al., 2006). Additionally, dispersal ability and propensity do not always go hand in hand, as exemplified in many highly philopatric bird species that are capable of annual migrations across thousands of kilometers. Using genome-wide datasets makes it possible to delineate genetic signatures that persist along different timescales: both the more rapidly decaying patterns (e.g. reduced homozygosity or nucleotide diversity) and those that persist for longer (e.g. allele frequency clines in the direction of the expansion or allelic richness) (Allendorf, 1986; Peter & Slatkin, 2013). Comparing these will help to clarify the strength and persistence of founder effects in contemporary range expansions of highly mobile species.

Before we can investigate the genetic consequences of range expansion, it is important to reconstruct the colonisation history of the species involved (Welles & Dlugosch, 2019). Identifying the geographic origin of colonisers is vital for meaningful comparisons, which facilitate the assessment of changes in genetic diversity

during an expansion. Knowledge of the colonisation history can also help to identify the ecological factors that underpin range expansions and detect local adaptation that has occurred at the range edge. However, reconstructing the expansion pathway is a non-trivial task as it requires (a) direct, often anecdotal observations, (b) careful monitoring schemes, or (c) molecular data with sufficient resolution across an appropriate geographical area. Drawing together multiple lines of evidence benefits range expansion studies by improving the confidence for inferring colonisation history, and alleviating potential biases associated with any single approach (e.g. observational bias from less conspicuous species and hard to access habitats; Hughes et al., 2021, as well as technical imperfections and lacking genomic resolution even with the currently most sensitive methods available; Kemppainen et al., 2024). Furthermore, if the expanding population manages to avoid the genomic patterns characteristic of a range expansion, having different lines of evidence can nevertheless make it possible to elucidate its origins.

The common reed warbler (*Acrocephalus scirpaceus*) provides an excellent case study of the genetic consequences of environment-driven range expansions in species with high dispersal capabilities. This long-distant migratory species has a very large distribution, with the nominate subspecies (*A. s. scirpaceus*, from here on "reed warbler") wintering in sub-Saharan Africa and breeding in reed beds (stands of *Phragmites*) across Europe, western Turkey, and north-western Africa (Gill et al. 2023). Despite undergoing a rapid expansion at the northern edge of its breeding range, the species' dispersal propensity at its breeding range core is rather low and characterised by strong breeding philopatry, slightly more natal dispersal, and very rare LDD events (Davies, 2019; Paradis et al., 1998), representing a site fidelity pattern typical in songbirds (Leisler & Schulze-Hagen, 2015). In northern Europe, the reed warbler has established large breeding populations in Fennoscandia and the Baltic countries during the past ~150 years, settling in areas as far north as 65°N. This expansion is likely a result of an increase in habitat due to eutrophication and decreased grazing pressure by livestock (Jutila, 2001; Pitkänen et al., 2013; von Numers, 2011), as well as climate change, as the range of reed warblers seems to be temperature-limited at its northern edge (Davies, 2019; Virkkala et al., 2005). There are written records of the species' occurrences and expansion from around the 1850s onwards, but the collation of these observations from parts of the Baltic Sea region and western Russia remains incomplete (e.g. Malchevsky & Pukinsky 1983, Glutz von Blotzheim, 1991). Evidence for the origin of the reed warbler's expansion can also be inferred from ringing records and population genomics. Reed warblers show a prominent migratory divide in Central Europe: individuals roughly southeast of the Czech Republic take a southeastern route to their wintering grounds, while according to current knowledge, individuals from other areas head southwest (Procházka et al., 2018). As migration bearing is known to be heritable in birds (Liedvogel et al., 2011), the migratory direction in newly colonized areas can indicate which side of the migratory divide colonisation originated from. It is known from earlier studies that reed warblers in both Fennoscandia and the Baltic countries take the SW route (Wynn et al., 2022), with local (between-country) differences in the mean migratory bearing (Fransson & Stolt, 2005). However, within-country differences could also reveal a genetic contribution from more eastern areas. Based on earlier work, we know that genetic differentiation within the

breeding range of *A. s. scirpaceus* is low (Procházka et al., 2011), but a recent study using genomic data suggests that there should be sufficient structure to detect genetic patterns of range expansion, if such patterns were to persist (Sætre et al., 2022).

Here we combine a literature search of historical observations, ringing records, and genomic data from restriction-site associated DNA sequencing (RAD-seq, Baird et al., 2008) to address the following two questions: (i) What are the origins of the recent and rapid northward range expansion of the reed warbler, and (ii) can loss of genetic diversity or clines in allele frequencies be detected along the expansion axis? The genomic dataset covers most of the breeding range of *A. s. scirpaceus*, including sequences filling in a previous sampling gap in recently colonized northern areas (Sætre et al., 2022). We expect that the most likely expansion origin is in Central or Western Europe, with very little or no contribution from the eastern side of the migratory divide (Fransson & Stolt, 2005; Wynn et al., 2022). We supplement the genomic data with ringing records from the range edge population in Finland, where recent immigration from southeastern Europe should be detectable from differing individual migratory directions, and a significant historical contribution should be visible in within-country differences in mean direction (eastern and western Finland). We hypothesise that whilst genetic diversity is likely to be mostly maintained along the colonisation route (Procházka et al., 2011; Sætre et al., 2022), the longer-preserving allele frequency clines are likely to persist after such a rapid expansion in a mainly philopatric species (e.g. Garnier & Lewis, 2016). This dataset allows us to inspect the fine-scale genetic signals of past founder effects in a documented, environment-driven range expansion, providing a novel example for understanding the conditions under which species can successfully respond to ongoing environmental change.

2. Methods

2.1 Literature search for historical observations

We conducted a literature search to locate as many historical records of local first occurrences of reed warblers and their distributions in Northern Europe (e.g. Denmark, Estonia, Finland, Latvia, Lithuania, Norway, Russia, Sweden) as possible. We relied mostly on published observations, supplemented by secondary sources such as yearbooks of local birding societies and handbooks. To find records from the areas of the former Soviet Union and East Prussia, we contacted a local expert (N. Chernetsov) for guidance.

2.2 Migratory direction

We collated all available data from reed warblers ringed in Finland and re-encountered outside of Finland (years 1969-2023, n = 326) from the databank of the Ringing Centre at the Finnish Museum of Natural History. As Finland is at the species' range edge and there should be no individuals migrating through the country to other areas, we used ringing data from all months. However, these should be interpreted as ringing locations instead of confirmed natal or breeding sites. To assess within-country variation in mean migratory direction and individual migration routes, possibly indicating contributions from different sides of the species' migratory divide, we split the Finnish data into a western (n = 173) and an eastern group (n = 153) by the mean longitude of the site of ringing. Mean individual migratory directions were calculated as the consistent compass bearing between the ringing and re-encounter events. The mean direction of each group was calculated based on re-encounters south of 49°N (western group n = 26, eastern group: n = 33) to 1) avoid the large amount of birds controlled in Belgium (due to high catching effort) affecting the mean (see Fransson & Stolt, 2005), and 2) to avoid bias from re-encounters made near the breeding area, since directions calculated based on short distances may reflect shorter-scale movements (e.g. post-fledgling and post-breeding movements; Chernetsov, 1998; or avoiding geographic barriers such as the Baltic Sea) and less reliably represent innate migratory direction (e.g., Helbig, 1992). Finally, to statistically test whether ringing longitude affects migratory direction, we fit the following simple linear models in R (v4.3.0; R Core Team 2023) to both the full dataset and to re-encounters made south of 49°N: 1) recovery longitude ~ ringing longitude, 2) recovery direction ~ ringing longitude.

2.3 Population genomics

2.3.1 Sample collection, DNA extraction, and RAD-seq library preparation

We collected or received blood samples from 172 reed warblers representing 20 sites at least 50 km apart from each other across the species' European breeding range (Fig. 1, Table 1). Blood samples were taken from resident birds during the breeding season in years 2002-2019 and stored in either Queen's lysis buffer (+4°C or -20°C), ethanol (+4°C) or on FTA cards™ (Whatman, Maidstone, United Kingdom) (room temperature). For sample details, see Supplementary S1.

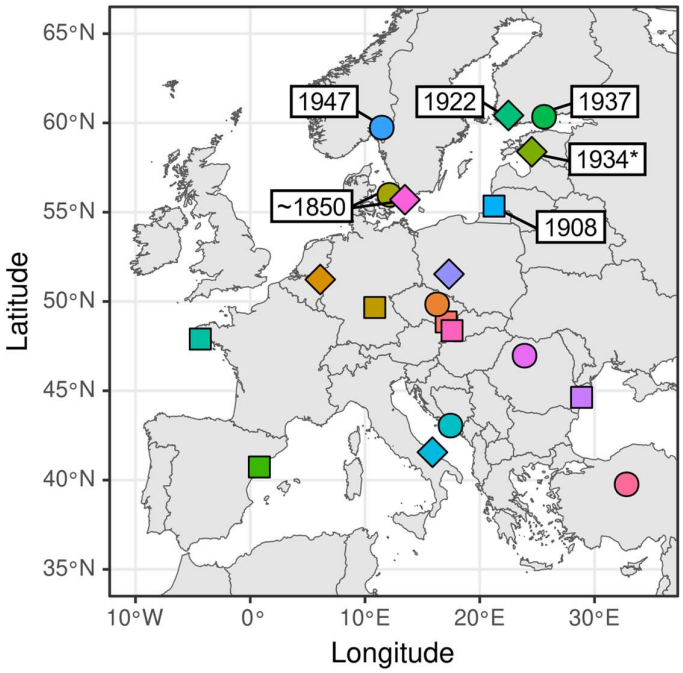

Figure 1. Map of the sampling sites for the molecular data. The labels denote the year of the first recorded occurrence of the reed warbler in each sampling site area within the newly colonised range. The * by the Estonian first occurrence indicates the year of the first record of the species in over 60 years, although a single nest was reported already in 1870.

Genomic DNA was extracted using a QIAGEN DNeasy Blood & Tissue Kit (Qiagen Inc., California, U.S.A) at the University of Oslo or the University of Helsinki. The samples were first handled according to their storage medium, after which the manufacturer's protocol was followed with minor changes (Supplementary S2). After extraction, the purity and DNA concentration of the eluates was assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific) and a Qubit 4 fluorometer (Invitrogen). The DNA concentration of each sample was normalized to 20 ng/µl before sequencing. Gel electrophoresis was used to ensure that the DNA fragments were of desired size (molecular weight > 10 kb).

Single-digest, single end RAD-seq libraries were prepared at Floragenex, Inc. (Oregon, U.S.A) for a total 285 samples (including 12 replicate controls, as well as samples sequenced as a part of another project and not analysed further here), following Baird et al. (2008). DNA was digested with the *Sbf*I enzyme and sequenced on an Illumina HiSeq 2000 platform. Each sample was barcoded and sequenced on two separate lanes. Samples were sequenced on three plates, one in 2017 (read length 91bp) and two in 2021 (read length 150bp). Replicate samples were included between all plates to allow the detection of potential batch effects, and samples sequenced in 2021 were randomized across the two batches to decouple the biological signal from any batch differences (Meirmans, 2015).

Table 1. Information about the sampling sites and samples included in the final analyses.

Site	*Locality*	*Site code*	*Longitude*	*Latitude*	*N (individuals)*
Czech Republic (1)	Lužice	CZLU	17.07	48.85	5
Czech Republic (2)	Osík	CZOS	16.28	49.84	5
Germany (1)	Diergarten	DEDG	6.10	51.23	10
Germany (2)	Mohrhof	DEMH	10.85	49.66	9
Denmark	Arresø	DK	12.12	55.97	10
Estonia	Pärnu	EE	24.53	58.38	10
Spain	Ebro Delta	ES	0.79	40.73	10
Finland (1)	Porvoo	FIE	25.60	60.34	10
Finland (2)	Kaarina	FIW	22.50	60.42	6
France	Trunvel	FR	-4.35	47.90	9
Croatia	Neretva	HR	17.45	43.05	8
Italy	Lago Salso	IT	15.87	41.56	10
Lithuania	Ventės ragas	LT	21.24	55.35	10
Norway	Hemnessjøen	NO	11.46	59.74	10
Poland	Milicz	PL	17.31	51.53	10
Romania (1)	Lacul Ştiucilor	ROTR	23.90	46.97	9
Romania (2)	Lacul Sinoe	RODD	28.87	44.63	7
Sweden	Krankesjön	SE	13.48	55.70	6
Slovakia	Trnava	SK	17.58	48.37	9
Turkey	Mogan Gölü	TR	32.79	39.76	8

2.3.2 Sequence processing and reference alignment

The raw sequence quality was assessed using FastQC (Andrews, 2010), after which the data was cleaned and processed with STACKS (v2.62; Catchen et al., 2013; Rochette et al., 2019). The raw reads from each sequencing lane were cleaned and demultiplexed separately with the STACKS *process_radtags* program, using default settings: allowing for maximum one mismatch in the barcode or restriction site, and discarding the read if the average Phred score in a sliding window drops below 10. This resulted in a total of 2.14 billion reads across all three sequencing batches, out of which 90.6% were retained. The non-retained reads were dropped due to ambiguity in barcodes (8.6%) or restriction sites (0.2%), or a low quality score (0.6%). All reads were truncated to equal length (91bp), and sequences from the two lanes of each plate were combined to increase total coverage. The cleaned reads were aligned to the *A. scirpaceus* reference genome (Sætre et al., 2021) using BWA-MEM (v0.7.17-r1188, Li, 2013) with default settings. The resulting BAM files were sorted using SAMtools (Li et al., 2009), and the STACKS *ref_map.pl* program was run to assemble the loci according to alignment positions, and to perform SNP calling in each sample. The program retained 69.6%

of the primary alignments, the rest were dropped due to insufficient mapping qualities (25.7%), excessive soft-clipping (1.4%), or as unmapped (3.3%). From these, 452 673 RAD loci (containing 997 942 SNPs) were built with an effective per-sample mean coverage of 47.1x (SD=16.8x, min=15.1x, max=111.7x).

To check for any batch effects between our three sequencing plates (2017, 2021_plate1, and 2021_plate2), we ran STACKS *populations* without filters for the 12 control samples that were re-sequenced across plates. We checked the relatedness of each replicate pair, which corresponded to the same individual in all samples except one, where a similar relatedness value to another sample indicated likely mislabelling. These samples were excluded from further analyses. We then checked that the control samples cluster by sample identity and not by batch in a principal component analysis (O'Leary et al., 2018), which showed no signs of batch-based clustering across any of the three batches (Supplementary S3). With VCFtools (v0.1.16; Danecek et al., 2011), we outputted the observed and expected heterozygosity, mean depth, and missing data per individual and compared sample values across batches using paired, two-tailed t-tests. The two plates sequenced in 2021 did not differ significantly from each other in any of the tested values (Supplementary S3), and therefore the 2021 controls were treated as one group in the comparisons with the 2017 plate. Between the 2021 and 2017 batches, we detected small but statistically significant differences ($p < 0.05$) in all tested statistics (Supplementary S3), observed heterozygosity being on average 3.7% lower and expected heterozygosity 0.1% lower in the 2017 batch. As even slight differences in heterozygosity due to technical artefacts might produce a false signal when biological differences are small, we only used samples free from batch effects (i.e. Datasets 3 & 4 in Table 2) for the most sensitive analyses regarding allele frequency clines and genetic diversity.

2.3.3 Creating the filtered datasets for population genomic analyses

We first ran an exploratory PCA using the glPca function in the R package *adegenet* (Jombart & Ahmed, 2011), and excluded two extreme outlier individuals from further analyses (likely misidentified species or subspecies, see Arbabi et al., 2014). We calculated the mean sequencing depth and missingness per site and individual using VCFtools; no individuals were discarded based on these. With VCFtools, we extracted a list of loci containing sites with minimum mean depth 10 and maximum mean depth corresponding to 2x mean depth of all sites. These loci, together with all loci on sex chromosomes, were converted into a blacklist for STACKS *populations*. After re-running *populations* with this blacklist and basic quality filters, we calculated the relatedness between samples using --relatedness2 in VCFtools, based on the method by Manichaikul et al. (2010). Only one individual from each pair or group of close relatives (1st-3rd-degree) was kept (e.g., Wang, 2018). We also removed two pairs of samples with pairwise relatedness values corresponding to a duplicate or monozygotic twin, likely indicating contamination or mislabelling. STACKS *populations* was run again without control samples and mean depth was re-calculated for a new blacklist, again containing the sex chromosomes and sites with mean depth <10 or >2x the mean depth of all sites. To determine a thinning

threshold for analyses that assume unlinked markers, we used the PopLDdecay software (v3.42; Zhang et al., 2019) to calculate the decay of linkage disequilibrium over physical distance (Supplementary S4).

Finally, we downsampled the data to a maximum of 10 individuals (n = 5–10) per sampling site to avoid the effects of uneven sampling on the genomic analyses (e.g., Lawson et al., 2018). STACKS *populations* was run to create the final datasets shown in Table 2. In all four datasets, we required a locus to be present in at least 80% of individuals (-r) in 75% of the sampling sites (-p) to reduce missing data, applying these filters haplotype-wise (-H) in the datasets with full RAD loci. We set a minimum minor allele count threshold to exclude singletons and private doubletons (--min-mac 3; Linck & Battey, 2019), a maximum heterozygosity threshold of 70% across all samples (--max-obs-het 0.70), and applied the blacklist described above (excluding sex chromosomes to prevent individuals clustering by sex instead of population structure, as well as low- and high-depth loci to prune out potential false homozygotes and paralogs). To exclude any remaining genotyping errors, we scanned for markers out of Hardy-Weinberg equilibrium in all populations with n = 10 (to increase statistical power), but there were none left to be removed at this stage. For the analyses assuming independent SNPs or RAD haplotypes (Table 2), we used VCFtools to thin the data according to the calculated LD decay distance and ran STACKS *populations* a final time using just a whitelist of the kept markers: SNPs min 3000 bp apart (**Dataset 1)**, or RAD loci min 3000 bp apart (**Dataset 2**). To be certain that the slight batch effect does not drive the results even in the most sensitive analyses, we additionally created two datasets with only the batches sequenced in 2021: a linkage-pruned SNP dataset (**Dataset 3**) and a non-linkage-pruned SNP dataset (**Dataset 4**).

Table 2. The filtered datasets that were used in the final analyses.

Dataset	Description	Analyses	N (ind.)	Sampling sites	#Loci; #SNPs
1) Full, linkage-pruned SNP dataset	All sampling sites included; SNPs pruned by physical distance	PCA, conStruct, F_{ST}	172	20	35564; 35564
2) Full, linkage-pruned RAD locus dataset	All sampling sites included; full RAD loci pruned by physical distance	fineRADstructure, D_{XY}	172	20	35456; 246760
3) Batch-effect-free, linkage-pruned SNP dataset	Only sampling sites from 2021 batches; SNPs pruned by physical distance	RangeExpansion	103	20	34533; 34533
4) Batch-effect-free, full RAD locus dataset (not linkage-pruned)	Only sampling sites from 2021 batches; not linkage-pruned	π, (private) allelic richness	103	12	78011; 246916

2.3.4 Population structure and differentiation analyses

While population structure analyses are not explicit inferences of colonisation history, they provide useful information about patterns of genetic variation and connectivity between populations. As recommended for population structure inference (Linck & Battey, 2019), we used complementary non-parametric and model-based methods to depict the structure and differentiation in our data. The specific dataset used for each of the analyses are listed in Table 2.

We performed a principal component analysis (PCA) for the full, linkage-pruned SNP dataset (Dataset 1) using the glPca function in the R package *adegenet* (Jombart & Ahmed, 2011) and plotted the results with *ggplot2* (Wickham 2011). To use the resolution provided by full RAD haplotypes for analysing population structure, we ran the population inference package fineRADstructure (v0.3.2; Malinsky et al., 2018) for the full, linkage-pruned RAD locus dataset (Dataset 2), allowing maximum 10 SNPs per locus (-n 10). FineRADstructure's MCMC clustering algorithm was run with 100,000 burn-in iterations, followed by 100,000 iterations sampled every 1000 iteration steps (-x 100000, -y 100000, -z 1000), and tree-building algorithm with 10,000 iterations (-m T, -x 10000). To disentangle any patterns of discrete structuring that cannot be explained by isolation by distance only, we used the spatial, model-based clustering package conStruct (v1.0.5; Bradburd et al., 2018) for the full, linkage-pruned SNP dataset (Dataset 1). We first ran a cross-validation procedure implemented in conStruct to select between spatial and non-spatial models and the number of discrete layers (values of K). Cross-validation was run for K=1–5, with four repetitions with each value of K, 5000 iterations per repetition, and a training proportion of 0.9. To choose the most informative value of K for our data, we required that each new layer should contribute at least 5% to the overall covariance of the model. The diagnostic plots for MCMC performance were checked for all repetitions of the chosen value of K to confirm that the chains were well-mixed. Finally, we calculated pairwise F_{ST} values and their statistical significance based on 5000 bootstrap iterations using the R package StAMPP (v1.6.3; Pembleton et al., 2013) for the full, linkage-pruned SNP dataset (Dataset 1).

2.3.5 Origins of the range expansion

To test whether a genetic pattern of past founder effects persists, we used an approach by Peter & Slatkin (2013, 2015), implemented in the RangeExpansion R package. The directionality index (ψ) detects allele frequency asymmetries, caused by past founder effects, between pairs of populations. Clinal variation in ψ is still the most sensitive statistic for detecting signals of range expansions in genomic data, its power decaying slower than that of F_{ST} or heterozygosity (Kemppainen et al., 2024; Peter & Slatkin, 2013). We calculated the

pairwise ψ values for all sampling sites in the batch-effect-free, linkage-pruned SNP dataset (Dataset 3) with the modified function (get.all.psi.mc.bin) from Kemppainen et al. (2024), which fixes a major bug in the original package by inverting the polarity of the result matrix, and applies a binomial test to calculate the significance of each value.

2.3.6 Genetic diversity

To compare levels of genetic diversity across the sampled range of the reed warbler, we used STACKS *populations* to calculate nucleotide diversity (π) for each sampling site using the batch-effect-free, non-linkage-pruned full RAD locus dataset (Dataset 4), including variant and invariant sites to avoid biases stemming from only using variant sites (e.g. Korunes & Samuk, 2021). Allelic richness was calculated using ADZE (v1.0; Szpiech et al., 2008) for the same dataset (Dataset 4). The number of samples per site was trimmed down to the smallest sample size across sites (n = 6) as uneven sampling impacts the calculations (Kalinowski, 2004; Szpiech et al., 2008), and a maximum of 10% missing data per locus was allowed. For both π and allelic richness, we tested for correlation with latitude using a simple linear regression and compared the mean values between the recently colonised range (Fennoscandia and Baltic countries) and the other sites using a two-tailed t-test.

3. Results

3.1 Historical observations suggest a southwestern expansion origin

During the 19th and early 20th century, reed warblers were documented expanding northwards along both sides of the Baltic Sea. In terms of latitude, the colonisation wave along the western coast through Scandinavia seems to have been a few decades ahead of the wave advancing through countries along the eastern coast of the Baltic Sea. In the mid-1800s (the beginning of available observation data for the species), reed warblers were likely breeding in most of mainland Denmark (Dybbro, 1976; Løppenthin, 1967), the regions south of Kaliningrad, Russia (at the time Königsberg, East Prussia; Tischler, 1914), and a few locations in southern Sweden (Risberg, 1990). The species expanded rapidly from the late 1800s onwards, in the west reaching the Stockholm area around the 1910s (Risberg, 1990; Svensson et al., 1999), and in the east settling in northern parts of the Curonian Lagoon in the early 1900s (Tischler, 1941). The expansion reached southwestern Finland in the 1920s (Turku 1922, Åland islands 1926) (Leivo, 1937; Wikström, 1945). The colonisation wave through the Baltic countries followed closely: reed warblers were first found breeding in both Latvia and Estonia during the early 1930s (Riga 1931, Pärnu and Matsalu Bay 1934, although a single, possible nesting has been reported from Estonia already in 1870) (Mang, 1936; Sits, 1936; Tischler, 1941; Russow, 1874; cited in Sits, 1936). However, it is possible that the first colonisers went

undetected, as multiple individuals and nests were found in all these locations already in the year of the first observation. The first finding from the Helsinki region in Finland is from 1933, but based on how common and widespread reed warblers were found to be in 1937, Leivo (1937) suggests that the species might have already occurred in the area some years before. It therefore seems possible that the two expansion waves, from the southwest (Sweden) and south (Estonia), both reached Finland in the 1920s-30s. In Norway, the first confirmed breeding was in 1947 at Lake Borrevatn near Oslo, although the first sighting was made a decade earlier on the island of Utsira at the SW coast of the country (Haftorn, 1971; Løvenskiold, 1947). To the easternmost tip of the Baltic Sea in the Saint Petersburg region, Russia (at the time Leningrad, USSR), reed warblers only arrived in the 1960s, 30-40 years after the first records from neighbouring areas in Finland and Estonia (Malchevsky & Pukinsky, 1983). A distribution map from the 1950s shows that the species was absent directly east of the Leningrad area 10 years before the first observation near Saint Petersburg, and ~500 km away in the south (Dement'ev et al., 1968). Observations support that the expansion proceeded from west to east and from coastal areas inland (Malchevsky & Pukinsky, 1983), making east an unlikely origin for the Fennoscandian and Baltic reed warbler range expansion.

3.2 Migratory direction supports expansion from the SW side of the migratory divide

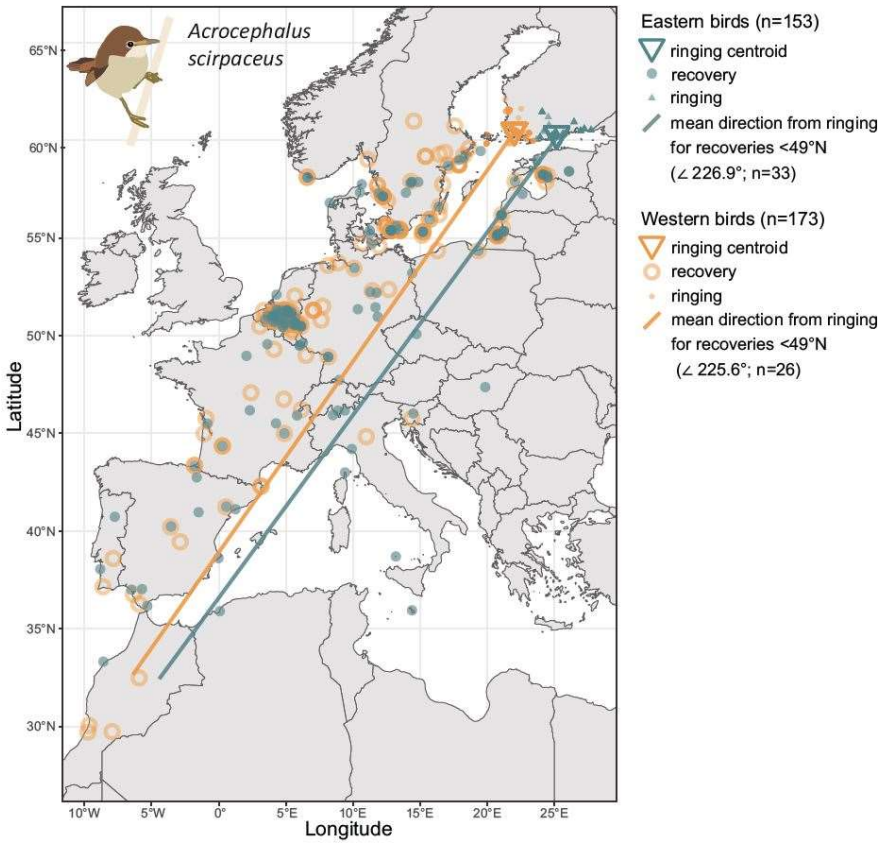

Figure 2. Map of all re-encounters from reed warblers ringed in Finland. The re-encounters are divided into a western (orange) and an eastern (teal) group based on the mean longitude of the ringing site. Opacity indicates multiple, overlapping points. Mean migratory direction for each group is calculated from re-encounters made south of 49°N. For individual trajectories during spring and autumn migrations separately, see Supplementary S5.

As expected under the hypothesis that the reed warbler's range expansion originated from the southwestern side of the species' migratory divide, none of the individuals ringed in Finland (n = 326) were re-encountered along the southeastern migration route. However, there was some variation in where birds were re-encountered along the SW migration route. For recoveries south of 49°N (n = 59), the birds ringed in eastern parts of Finland had a marginally significant tendency to also be recovered at more eastern longitudes, and vice versa ($\beta = 0.98$, SE = 0.51, t = 1.92, p = 0.059), but the correlation between ringing longitude and recovery direction was not significant ($\beta = 0.50$, SE = 0.72, t = 0.695, p = 0.49). Neither relationship was statistically significant when all recoveries (n = 326) were included (recovery longitude: $\beta = 0.13$, SE = 0.23, t = 0.57, p = 0.57; recovery direction: $\beta = 0.81$, SE = 0.57, t = 1.42, p = 0.16). Interestingly, birds ringed in the western (n = 173) and eastern (n = 153) parts of Finland were both found to be migrating to and from their breeding grounds along either side of the Baltic Sea. While most individuals seem to cross the Mediterranean Sea through the Iberian Peninsula, the re-encounter data suggest that Finnish reed warblers may rarely use the Italian peninsula for crossing (Fig. 2), during both the spring and autumn migrations (Supplementary S5).

3.3 Weak but significant, geographically continuous population structure

The overall population structure of the reed warbler across the European range largely reflects geography. In the principal component analysis, PC1 (1.13% of variation) appears to align samples along a continuous south-north cline, PC2 (0.81%) reflects a west-east cline, and PC3 (0.76%) separates most Spanish individuals from the rest of the samples (Fig. 3A-B). However, PC1 also reflects the northern expansion as Norway is positioned at the end of the genetic continuum, despite samples being collected from 1 degree south of Finland. The results from fineRADstructure are in line with the results of the PCA: individuals generally cluster by sampling site and geography in the coancestry matrix, but most of the main splits are not well-supported as separate clusters (posterior probability < 95%), indicating continuous rather than discrete structuring (Fig. 3C). The same patterns are reflected in the F_{ST}, where in general, differentiation increases with increasing geographic distance (Fig. 4). Pairwise F_{ST} values are low but mostly statistically significant ($p < 0.05$), significant comparisons ranging from 0.021 between eastern Finland and Turkey to 0.001 between eastern Finland and Estonia.

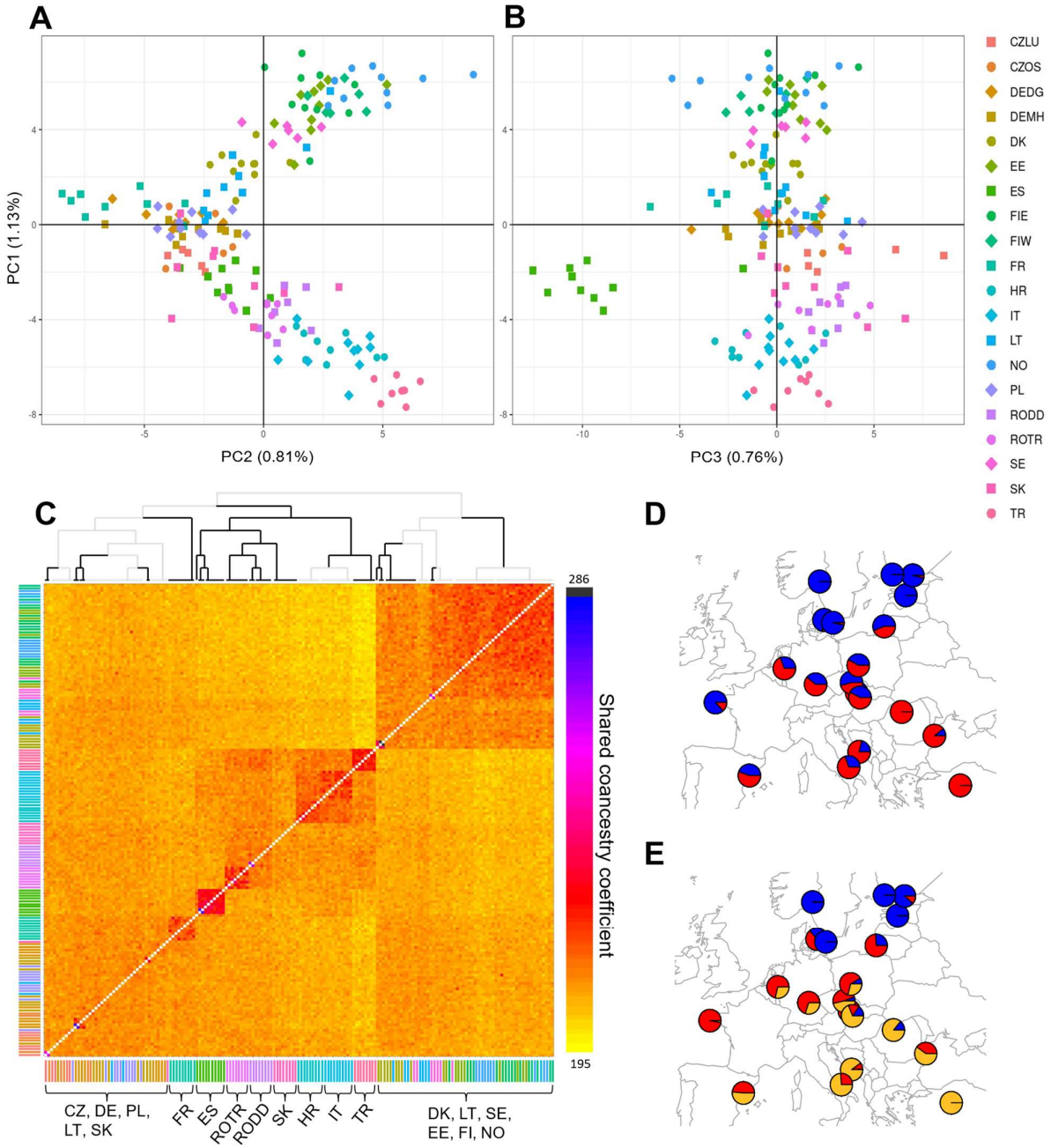

Figure 3. Population structure across the reed warbler's European breeding range. (A) The two first principal components of the PCA show geographically continuous structure (PC1 and PC2 axes rotated to reflect latitude and longitude, respectively). Each point represents an individual, the colour and shape correspond to its sampling site. (B) PC1 against PC3, which separates most Spanish samples from the other sites. (C) The fineRADstructure coancestry matrix similarly shows clustering by geographical proximity. The branches of the tree with high support (>95% posterior probability) are shown in black. Each row and column correspond to an individual and its genetic relatedness every other individual in the data, heatmap colour corresponding

to the relative genetic similarity. Sampling site is indicated with the same colour as in the PCA. (D) The overall best model in the conStruct analysis, a spatial model with K=2, suggests the presence of two genetic clusters after accounting for isolation by distance. However, there is very little genetic variation separating the two layers and all replicates did not converge on the same split. (E) The best non-spatial conStruct model (K=3) shows geographical clustering in line with the other non-spatial analyses.

The spatial population analysis conStruct shows that the reed warbler population structure can be largely explained by isolation by distance, and any additional, discrete patterns are weak and receive mixed support. In the package's cross-validation procedure, a spatial model had a higher predictive accuracy than the corresponding non-spatial model with all tested values of discrete layers (K). Based on the cross-validation and layer contributions (Supplementary S6), a spatial model with K=2 could be considered to be the most relevant description of the data. However, the amount of variation separating the two layers in this model is so small that not all replicates converged on the same split of the layers, despite good MCMC diagnostics (Supplementary S6). In Figure 3D, the most common split for the K=2 spatial model is shown (3 out of 4 replicates converging on it). This split indicates that the northern and southern sites would be more differentiated from each other than expected under IBD alone. However, the most biologically relevant model might in fact be a spatial model with K=1, considering the negligible amount of variation separating the two layers in the K=2 model (Supplementary S6), meaning that nearly all spatial variation could be explained by geographic separation (IBD). The non-spatial model (comparable to an ADMIXTURE model; Alexander et al., 2009; Bradburd et al., 2018) that best describes the data is K=3, which again shows geographical structuring in line with the results from all other non-spatial structure analyses (Fig. 3E).

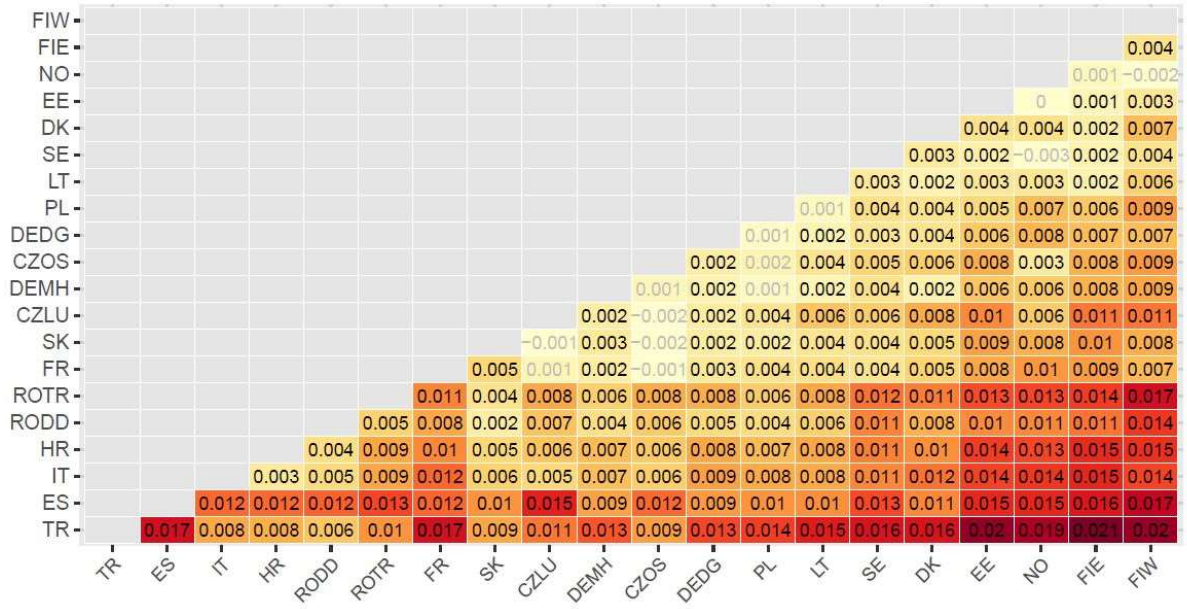

Figure 4. Pairwise F_{ST} between the sampling sites, ordered by latitude. Darker colours indicate higher differentiation. Significant F_{ST} values ($p < 0.05$) are reported in black text, and non-significant values in light grey.

3.4 No detectable signal of range expansion

No allele frequency clines were detectable with high enough confidence to distinguish from an equilibrium scenario. Directionality index (ψ), quantifying directional change in allele frequencies across the sampled range, was very low in all pairwise comparisons between sampling sites (absolute value $|\psi| < 0.03$), and the proportion of significant comparisons (2/66, 3%) was lower than the proportion of false positives expected under the null hypothesis of no expansion (5%, Kemppainen et al., 2024). The values of ψ in our data fall clearly below the cutoff $|\psi| > 3$ used by Peter & Slatkin (2013, 2015). The lack of a persisting signal of founder effects, even in the most sensitive statistic for recent range expansions, ψ, likely indicates significant gene flow after the initial founder events.

3.5 Genetic diversity

Levels of nucleotide diversity (π) and allelic richness were highly consistent across the range and showed no reductions in the newly colonised sites (Fig. 5). The mean value of π ranged from 0.00585 (Sweden) to 0.00607 (Estonia, Poland) per sampling site, with no linear correlation with latitude or significant difference in mean π between recently colonised sites (Fennoscandia and Baltic countries) and the other sites (Supplementary S7). Values of allelic richness ranged from 1.359 (Sweden and Spain) to 1.366 (Poland), also with no latitudinal correlation or mean difference between the recently colonised and other sites, even though allelic richness should retain the signal of founder effects longer than frequency-based measures of genetic diversity, such as π (e.g. Greenbaum et al., 2014).

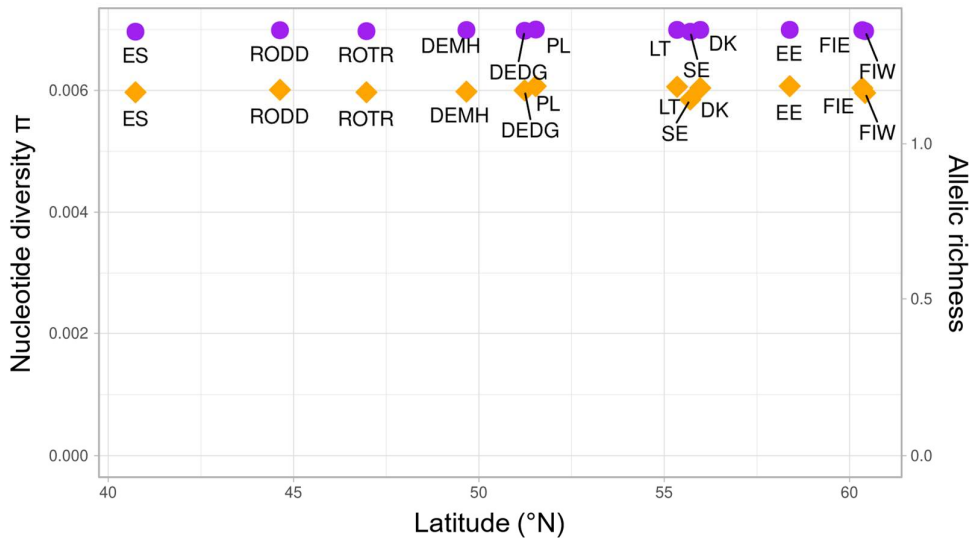

Figure 5. Nucleotide diversity π (orange diamonds) and allelic richness (purple circles) for each sampling site, arranged by latitude.

4. Discussion

In a world characterised by rapid environmental changes, it is important to understand the origins and consequences of recent environment-driven range shifts, as the loss of genomic variation following an expansion can impact the population's resilience in the face of further environmental challenges (Thorogood et al., 2023). Here we provide an example of the lack of genomic consequences in a remarkably rapid and extensive environment-driven range expansion of a migratory passerine, the common reed warbler. Despite the expansion spanning more than 1000 km in the past 100-150 years, we did not detect even minor losses in nucleotide diversity or allelic richness in the newly colonised areas at the northern range edge. The European breeding range of the reed warbler was characterised by weak, continuous population structure mainly reflecting geographic distance, although we also discuss the possibility that accelerated differentiation during repeated founder effects could lead to a similar pattern. While there was still some uncertainty regarding the origins of the range expansion from the genomic data, the overall genomic patterns were consistent with the historical observations and migratory direction from ringing data, which all support a southwestern origin for the species' northward expansion through Fennoscandia and Baltic countries. Even though colonisation histories can be complex and resolving all details difficult if not impossible, combining independent lines of ecological and genomic evidence allowed us to reconstruct a consistent and well-supported picture of the expansion routes and its associated genomic consequences.

4.1 Complementary evidence for southwestern origin of expansion

Although all three types of evidence for the expansion origin aligned, using them together ensured that potential deficits in any single data type did not lead to an erroneous conclusion. While the historical observations provided valuable reference points and suggested an expansion proceeding from the southwest, observational data of species' range expansions is often incomplete or inaccurate due to e.g. spatial bias in monitoring effort (e.g. Hughes et al., 2021) and does not alone provide means to directly assess the origin of the colonising individuals. However, ringing data also supported an expansion origin from the SW side of the species' migratory divide in Central Europe, as migratory bearing is known to be heritable in birds (Liedvogel et al., 2011) and none of the reed warblers ringed in Finland were recaptured along the SE route (whilst a few individuals were found to cross areas where local breeding birds are SE-migrating, e.g. Hungary or Italy, their migratory direction was still southwestern). Furthermore, there was a continuous pattern of genetic differentiation connecting sampling sites in Central and Western Europe (e.g. Germany and Poland) to the recently colonised range, pointing towards shared ancestry (Estoup & Guillemaud, 2010). Southeastern Europe (e.g. Romania and Croatia) is clearly more differentiated from the range edge in Fennoscandia and Baltic countries. Again, it needs to be noted that relying only on the genomic data could have introduced a possible source of bias due to sampling limitations (Estoup & Guillemaud, 2010) as we were not able to include samples from more eastern sites (e.g. Russia). However, this is unlikely to be a problem in the case of our study: the historical observations were detailed from the time when the expansion proceeded into the St Petersburg region from the 1960s onwards, documenting a chronologically advancing expansion from southwest along the southern coast of the Gulf of Finland (Malchevsky & Pukinsky, 1983).

Why the reed warblers in Central and Western Europe were so successful in colonising Northern Europe might be a consequence of newly available habitat being suitably positioned along their inherited migratory bearing (Bensch 2003, Wynn et al. 2022). While reverse migration driven by bi-axial variation in orientation has been suggested as a potential mechanism for birds to expand their range in directions that align with inverse migratory bearings (Wynn et al., 2024), this process is unlikely to explain the colonisation by reed warblers of previously unoccupied sites. Instead, migratory overshooting – where individuals extend their migratory journey beyond their typical breeding areas (Newton, 2023) – provides a more plausible mechanism for this range expansion. This is particularly likely for juvenile birds, whose navigational inexperience may predispose them to explore new habitats. Such movements may be further reinforced by post-fledging dispersal, a period thought to serve as a mechanism for searching for future breeding areas in juvenile migratory birds, including reed warblers (Nielsen & Bensch, 1995; Mukhin et al. 2005). Once northern habitats are colonised, subsequent generations may integrate magnetic inclination cues from the new range into their navigational toolkit (Wynn et al. 2022), stabilising the population. Another plausible mechanism driving the SW-to-NE range expansion of the reed warblers' northern breeding range could simply be the geographical proximity of SW-migrating populations to the northern range limit, with SE-migrating populations being possibly restricted to areas southeast of Czechia due to greater geographic distances and barriers to move northwards (e.g. the Carpathian Mountains). The SW-migrating populations would therefore have had both the advantage of shorter distances

to newly available habitats and the alignment of their inherited migratory direction with the Baltic coastline, likely providing a natural corridor for dispersal.

4.2 Why was there no detectable signal of range expansion?

The remarkable maintenance of genetic diversity despite the colonisation distance and pace raises the question: is this an exception, or perhaps more common in nature than was previously thought? Despite the theoretical expectation of consecutive founder effects and strong drift eroding genetic diversity during range expansions (Excoffier, 2004; Excoffier et al., 2009a), it is also well-known that a few migrants per generation can have a large impact on the maintenance of genetic diversity (Mills & Allendorf, 1996). While consequences of founder effects have been reported from expansions of species with high dispersal capabilities (Adams et al., 2023; Pierce et al., 2014), the results of this study are in line with a handful of earlier studies that have documented conserved genetic diversity. Using mtDNA and microsatellite markers, maintained diversity was observed in a shorter-scale expansion of melodious warblers (*Hippolais polyglotta*; Engler et al., 2016) and in a rapid invasion of non-native starlings (*Sturnus vulgaris*; Berthouly-Salazar et al., 2013). High levels of genetic diversity were also documented following the expansion of light-vented bulbuls (*Pycnonotus sinensis*) with the exception of the most recently colonised site (~10ya), although with possible human-mediated dispersal through wildlife trade (Song et al., . A recent whole genome sequencing study of a comparably extensive range expansion of Anna's hummingbird (*Calypte anna*; Adams et al., 2023) reported slight but significant reductions in nucleotide diversity at the range edge but no range-wide population structure, seemingly opposite to the reed warblers with some spatial structuring but fully-maintained diversity. The authors discuss the possibility that during their timeframe (30-60 years since colonisation of the range edge), allele frequency differences might not have had time to build up. It is possible that the slightly longer time in our case (85-100 years) has made a difference in pronouncing differentiation and replenishing diversity. However, in terms of generation time (4.3 years for the reed warbler and 2.3 years for Anna's hummingbird; IUCN 2024), the timescales of the two expansions are very similar and the differences might be explained by other factors. Apart from some potentially expansion-limiting factors (e.g., low dispersal propensity and habitat specialisation), the reed warbler has had good prerequisites for a successful range expansion, including a large distribution and census population size, high dispersal capabilities, and a possibility for niche-tracking due to newly available habitat. These all have been found to be positively associated with range expansion success in previous studies (MacLean & Beissinger, 2017; Mair et al., 2014; Pacifici et al., 2020).

Most of the abovementioned studies identify long-distance dispersal (LDD) as one of the most important processes maintaining genetic variation at the range front (Berthouly-Salazar et al., 2013; Engler et al., 2016; Song et al., 2013), can this be the case with the reed warbler? In its range core, the species' dispersal kernel is characterized by strong philopatry (especially between breeding events), some short-distance dispersal (especially between natal site and first breeding), and very rare LDD events (Paradis et al., 1998). When

accounting for detection probabilities and only using birds with confirmed breeding status, site fidelity estimates are even stronger. In Great Britain, 94.8% of adults returned to the same site between breeding events, 2.8% moved 1-3 km and one individual moved 33 km (n=1871). Natal dispersal (n=171) was less than 12 km in 96.7% of the cases, but there were some LDD events of up to 77 km (Davies, 2019). The longest confirmed breeding dispersal event in the reed warbler is 160 km (Hałupka et al., 2022). Despite the rarity of LDD events, it has been estimated that short-distance dispersal alone cannot explain the recent rate of reed warblers' range expansion in Britain (Davies, 2019). Furthermore, especially when considering the maintained genetic diversity in Northern Europe, it does not seem plausible that the colonisation of Fennoscandia and Baltic countries would have been driven by only a handful of long-distance-dispersing individuals. Instead, the expansion may have been by changes in expansion-facilitating traits, such as increased dispersal propensity, a phenomenon documented in other expanding species (e.g. Duckworth & Badyaev, 2007; Phillips et al., 2006). Such traits, combined with large population size, high dispersal capacity and the ability to track suitable habitats, likely contributed to the reed warbler's success.

4.3 Could founder effects still underlie part of the geographically continuous population structure?

An interesting pattern in the data is that levels of differentiation are so similar across the European range, even though the northern populations have had only <150 years to drift apart genetically compared to the thousands of years after postglacial expansion in the Central European range core (Arbabi et al., 2014). This could reflect different processes (or their combinations) behind the observed pattern of isolation by distance, such as isolation by colonisation (IBC), which can create very similar patterns to IBD in the case of serial colonisation events (Orsini et al., 2013). Founder effects can lead to patterns of divergence persisting even for thousands of generations (Boileau et al., 1992), and the genetic differentiation caused by founder effects can further be preserved if initial colonisers have ecological or evolutionary advantage over later immigrants (Orsini et al., 2013). Different processes coincidentally creating roughly equal amounts of differentiation at the range core and edge could also explain the ambiguous results of conStruct, where a very small amount of variation was supported as a significant departure from pure IBD in the spatial model, despite all replicates not agreeing on the same split of groups. Also in the PCA, the newly colonised sites do not fully reflect geography and instead form a straight cline of differentiation roughly in the order of colonisation, possibly more fitting with IBC. Isolation by environment (IBE) is also a possibility, especially if selective pressures covary with geography. It has been suggested that IBE plays a role in reed warbler differentiation (Sætre et al., 2022, preprint), but further research is still needed to incorporate more ecological drivers of variation. Disentangling the processes behind the observed pattern would require assessment of neutral variation and loci under selection separately, and even then may not be straightforward (Orsini et al., 2013), making it a promising direction for further research.

4.4 Conclusions and broader implications

As anthropogenic environmental change is affecting ecosystems across the globe, it is valuable to understand the consequences of range shifts not only in rare or invasive species, but also for the native and abundant species that make significant contributions to ecosystem functioning. Founder effects and drift are generally expected to erode genetic diversity during range expansions, reducing adaptive potential and even affecting expansion success. While it is known that gene flow can alleviate these impacts, here we present a rare case where no genomic signatures of founder effects could be detected despite a remarkably rapid and geographically extensive expansion of a philopatric long-distance migrant, the common reed warbler. Even with a fine-resolution genomic dataset and the use of statistics which are highly sensitive to historical founder effects, we found no detectable reductions in genetic diversity or allele frequency clines in the direction of the expansion. Instead, levels of diversity were highly consistent across the species' European breeding range, and the low levels of population structure did not deviate from that expected under isolation by distance. Consequently, this study also illustrates the value of combining genomic and ecological data – by looking solely at the genomic patterns, it would not have been evident that the population went through a recent range expansion (see Campbell et al. 2018 for a similar example in detecting the origins of an invasive species). Reconstructing the expansion history enables biologically relevant comparisons of genetic and phenotypic changes between the source and range edge populations. This also paves the road for further studies, which are needed to disentangle the factors underlying expansion success, such as possible changes in expansion-facilitating behaviours. Another key question is whether local adaptation can persist at range edges under such considerable gene flow, even though genomic diversity is high. This study shows that even philopatric species with sufficient dispersal capabilities can avoid the genomic consequences of range expansions, providing insights into the conditions under which species can successfully keep up with current rates of environmental change.

Acknowledgements

We thank the many people involved in ringing reed warblers and contributing to EURING databases across Europe, and those who assisted in collection of blood samples in the field (including those previously mentioned in Procházka et al. 2011 and Sætre et al. 2022, preprint). New samples collected in Finland and Estonia were possible due to assistance from Julius Mäkinen and Hanna Huitu, and advice from Aleksi Lehikoinen, Margus Ellermaa and Vallo Tilgar. In addition, we thank Kirsi Kähkönen and Airi Lamminmäki for technical assistance in the Molecular Ecology & Systematics lab at the University of Helsinki, Nikita Chernetsov for helpful suggestions on the historical literature, and the EvoConGen research group, Petri Kemppainen, and Jochen Wolf for stimulating discussions about the methods and results. NB was supported

by the Doctoral Programme in Wildlife Biology (LUOVA), University of Helsinki, and the work was funded by Academy of Finland project grant no. 333801 and HiLIFE Fellows grant to RT.

Ethics statement

All blood samples were collected under permission from animal ethics experimentation boards and with permission to derogate from wildlife protection laws in respective countries. In Finland these were granted by the Project Authorization Board of the Regional State Administrative Agency (ESAVI/7857/2018) and the Centre for Economic Development, Transport and the Environment (VARELY/758/2018), in Estonia permission was granted by Keskkonnaamet (permit nr 02-2019). For samples from other locations, see Procházka et al. 2011 and Sætre et al. 2022, preprint.

Data availability statement

All data and analysis code will be made available on publication. For supplementary materials, please contact corresponding author.

Author contributions

Conceptualization: Nora M. Bergman, Rose Thorogood, and Katja Rönkä. **Data curation**: Nora M. Bergman, Petteri Lehikoinen, Camilla Lo Cascio Sætre, Frode Fossøy, and Katja Rönkä. **Formal analysis**: Nora M. Bergman, Petteri Lehikoinen, and Katja Rönkä. **Funding acquisition:** Nora M. Bergman, Petteri Lehikoinen, Staffan Bensch, Camilla Lo Cascio Sætre, Fabrice Eroukhmanoff, Petr Procházka, Bård G. Stokke, Rose Thorogood, and Katja Rönkä. **Investigation:** Nora M. Bergman, Petteri Lehikoinen, Edward Kluen, Camilla Lo Cascio Sætre, Fabrice Eroukhmanoff, Rose Thorogood, and Katja Rönkä. **Methodology:** Nora M. Bergman, Petteri Lehikoinen, Edward Kluen, Craig R. Primmer, and Katja Rönkä. **Project administration:** Nora M. Bergman, Fabrice Eroukhmanoff, Rose Thorogood, and Katja Rönkä. **Resources:** Nora M. Bergman, Petteri Lehikoinen, Staffan Bensch, Camilla Lo Cascio Sætre, Fabrice Eroukhmanoff, Frode Fossøy, Petr Procházka, William J. Smith, Bård G. Stokke, Rose Thorogood, and Katja Rönkä. **Software:** Nora M. Bergman and Katja Rönkä. **Supervision:** Fabrice Eroukhmanoff, Craig R. Primmer, Rose Thorogood, and Katja Rönkä. **Validation:** Nora M. Bergman and William J. Smith. **Visualization:** Nora M. Bergman, Petteri Lehikoinen, and Rose Thorogood. **Writing - original draft:** Nora M. Bergman. **Writing - review & editing:** Nora M. Bergman, Petteri Lehikoinen, Edward Kluen, Staffan Bensch, Camilla Lo Cascio Sætre, Fabrice Eroukhmanoff, Frode Fossøy, Petr Procházka, William J. Smith, Bård G. Stokke, Craig R. Primmer, Rose Thorogood, and Katja Rönkä.

References

Adams, N. E., Bandivadekar, R. R., Battey, C. J., Clark, M. W., Epperly, K., Ruegg, K., Tell, L. A., & Bay, R. A. (2023). Widespread gene flow following range expansion in Anna's Hummingbird. *Molecular Ecology*, 32(12), 3089–3101. https://doi.org/10.1111/mec.16928

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, 5(2), 181–190. https://doi.org/10.1002/zoo.1430050212

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arbabi, T., Gonzalez, J., Witt, H.-H., Klein, R., & Wink, M. (2014). Mitochondrial phylogeography of the Eurasian Reed Warbler *Acrocephalus scirpaceus* and the first genetic record of *A. s. fuscus* in Central Europe. Ibis, 156(4), 799–811. https://doi.org/10.1111/ibi.12174

Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P.-H. (1997). Evolution of Coalescence Times, Genetic Diversity and Structure during Colonization. *Theoretical Population Biology*, 51(2), 148–164. https://doi.org/10.1006/tpbi.1997.1302

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, 3(10), e3376.

Bensch, S. (1999). Is the range size of migratory birds constrained by their migratory program? *Journal of Biogeography*, *26*(6), 1225-1235.

Berthouly-Salazar, C., Hui, C., Blackburn, T. M., Gaboriaud, C., van Rensburg, B. J., Van Vuuren, B. J., & Le Roux, J. J. (2013). Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology*, 22(23), 5793–5804. https://doi.org/10.1111/mec.12538

Bialozyt, R., Ziegenhagen, B., & Petit, R. J. (2006). Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology,* 19(1), 12–20. https://doi.org/10.1111/j.1420-9101.2005.00995.x

Boileau, M. G., Hebert, P. D. N., & Schwartz, S. S. (1992). Non-equilibrium gene frequency divergence: Persistent founder effects in natural populations. *Journal of Evolutionary Biology*, 5(1), 25–39. https://doi.org/10.1046/j.1420-9101.1992.5010025.x

Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring Continuous and Discrete Population Genetic Structure Across Space. *Genetics*, 210(1), 33–52. https://doi.org/10.1534/genetics.118.301333

Campbell, C. D., Sarre, S. D., Stojanovic, D., Gruber, B., Medlock, K., Harris, S., ... & Holleley, C. E. (2018). When is a native species invasive? Incursion of a novel predatory marsupial detected using molecular and historical data. *Diversity and Distributions*, 24(6), 831-840.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, 333(6045), 1024–1026.

Chernetsov, N. (1998). Habitat distribution during the post-breeding and post-fledging period in the Reed Warbler *Acrocephalus scirpaceus* and Sedge Warbler *A. schoenobaenus* depends on food abundance. *Ornis Svecica*, 8(3), Article 3. https://doi.org/10.34080/os.v8.22946

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., & Sherry, S. T. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.

Davies, J. (2019). *The role of dispersal in range change in birds*. Doctoral dissertation, University of York.

Dement'ev, G. P., Birron, A., & Cole, Z. S. (1968). *Birds of the Soviet Union*. Israel Program for Scientific Translations.

Duckworth, R. A., & Badyaev, A. V. (2007). Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. *Proceedings of the National Academy of Sciences*, 104(38), 15017–15022. https://doi.org/10.1073/pnas.0706174104

Dybbro, T. (1976). *De danske ynglefugles udbredelse: Resultaterne af Atlas-projektet, kortlægningen af Danmarks ynglefugle 1971-74*. Dansk Ornithologisk Forening.

Edmonds, C. A., Lillie, A. S., & Cavalli-Sforza, L. L. (2004). Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences*, 101(4), 975–979. https://doi.org/10.1073/pnas.0308064100

Engler, J. O., Secondi, J., Dawson, D. A., Elle, O., & Hochkirch, A. (2016). Range expansion and retraction along a moving contact zone has no effect on the genetic diversity of two passerine birds. *Ecography*, 39(9), 884–893.

Estoup, A., & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology*, 19(19), 4113–4130. https://doi.org/10.1111/j.1365-294X.2010.04773.x

Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: Lessons from the infinite-island model. *Molecular Ecology*, 13(4), 853–864. https://doi.org/10.1046/j.1365-294X.2003.02004.x

Excoffier, L., Foll, M., & Petit, R. J. (2009b). Genetic Consequences of Range Expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Fransson, T., & Stolt, B.-O. (2005). Migration routes of North European Reed Warblers *Acrocephalus scirpaceus*. *Ornis Svecica*, 15(3), 153–160. https://doi.org/10.34080/os.v15.22735

Garcia-Elfring, A., Barrett, R., Combs, M., Davies, T., Munshi-South, J., & Millien, V. (2017). Admixture on the northern front: Population genomics of range expansion in the white-footed mouse (*Peromyscus leucopus)* and secondary contact with the deer mouse (*Peromyscus maniculatus*). *Heredity,* 119(6), 447–458.

Garnier, J., & Lewis, M. A. (2016). Expansion Under Climate Change: The Genetic Consequences. *Bulletin of Mathematical Biology*, 78(11), 2165–2185. https://doi.org/10.1007/s11538-016-0213-x

Garroway, C. J., Bowman, J., Holloway, G. L., Malcolm, J. R., & Wilson, P. J. (2011). The genetic signature of rapid range expansion by flying squirrels in response to contemporary climate warming. *Global Change Biology*, 17(5), 1760–1769. https://doi.org/10.1111/j.1365-2486.2010.02384.x

Gill F, D Donsker & P Rasmussen (Eds). (2023). *IOC World Bird List (v13.1)*. doi: 10.14344/IOC.ML.13.1.

Glutz von Blotzheim, U. N. (1991). *Handbuch der Vögel Mitteleuropas*. *Bd 12, Passeriformes (3. Teil), 2, Sylviidae*. AULA-Vlg.

Gralka, M., Stiewe, F., Farrell, F., Möbius, W., Waclaw, B., & Hallatschek, O. (2016). Allele surfing promotes microbial adaptation from standing variation. *Ecology Letters,* 19(8), 889–898. https://doi.org/10.1111/ele.12625

Greenbaum, G., Templeton, A. R., Zarmi, Y., & Bar-David, S. (2014). Allelic Richness following Population Founding Events – A Stochastic Modeling Framework Incorporating Gene Flow and Genetic Drift. *PLOS ONE*, 9(12), e115203. https://doi.org/10.1371/journal.pone.0115203

Grupstra, C. G. B., Coma, R., Ribes, M., Leydet, K. P., Parkinson, J. E., McDonald, K., Catllà, M., Voolstra, C. R., Hellberg, M. E., & Coffroth, M. A. (2017). Evidence for coral range expansion accompanied by reduced diversity of Symbiodinium genotypes. *Coral Reefs*, 36(3), 981–985. https://doi.org/10.1007/s00338-017-1589-2

Haase, M., Höltje, H., Blahy, B., Bridge, D., Henne, E., Johansson, U. S., Kaldma, K., Khudyakova, E. A., King, A., Leito, A., Mewes, W., Mudrik, E. A., Ojaste, I., Politov, D. V., Popken, R., Rinne, J., Stanbury, A., Tofft, J., Väli, Ü., & Schmitz Ornés, A. (2019). Shallow genetic population structure in an expanding migratory bird with high breeding site fidelity, the Western Eurasian Crane *Grus grus grus*. *Journal of Ornithology*, 160(4), 965–972. https://doi.org/10.1007/s10336-019-01688-1

Haftorn, S. (1971). *Norges fugler.* Universitetsforlaget.

Hagen, S. B., Kopatz, A., Aspi, J., Kojola, I., & Eiken, H. G. (2015). Evidence of rapid change in genetic structure and diversity during range expansion in a recovering large terrestrial carnivore. *Proceedings of the Royal Society B: Biological Sciences,* 282(1807), 20150092.

Hałupka, L., Liersz-Żelasko, K., & Jelínek, V. (2022). Long-distance breeding dispersal in a male Eurasian Reed Warbler: An unusual case in a highly philopatric species. *Ringing & Migration*, 37(1–2), 51–54. https://doi.org/10.1080/03078698.2023.2207034

Helbig, A. J. (1992). Population differentiation of migratory directions in birds: Comparison between ringing results and orientation behaviour of hand-raised migrants. *Oecologia*, 90(4), 483–488. https://doi.org/10.1007/BF01875441

Heppenheimer, E., Brzeski, K. E., Hinton, J. W., Patterson, B. R., Rutledge, L. Y., DeCandia, A. L., Wheeldon, T., Fain, S. R., Hohenlohe, P. A., Kays, R., White, B. N., Chamberlain, M. J., & vonHoldt, B. M. (2018). High genomic diversity and candidate genes under selection associated with range expansion in eastern coyote (*Canis latrans*) populations. *Ecology and Evolution*, 8(24), 12641–12655. https://doi.org/10.1002/ece3.4688

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, 44(9), 1259–1269. https://doi.org/10.1111/ecog.05926

IUCN 2024. *The IUCN Red List of Threatened Species*. Version 2024-1. Available online at: https://www.iucnredlist.org. (Accessed: 03 September 2024).

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070–3071.

Jordano, P. (2017). What is long-distance dispersal? And a taxonomy of dispersal events. *Journal of Ecology*, 105(1), 75–84. https://doi.org/10.1111/1365-2745.12690

Jutila, H. (2001). How does grazing by cattle modify the vegetation of coastal grasslands along the Baltic Sea? *Annales Botanici Fennici*, 38(3), 181–200.

Kalinowski, S. T. (2004). Counting Alleles with Rarefaction: Private Alleles and Hierarchical Sampling Designs. *Conservation Genetics*, 5(4), 539–543. https://doi.org/10.1023/B:COGE.0000041021.91777.1a

Keller, S. R., Olson, M. S., Silim, S., Schroeder, W., & Tiffin, P. (2010). Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*. *Molecular Ecology*, 19(6), 1212–1226. https://doi.org/10.1111/j.1365-294X.2010.04546.x

Kemppainen, P., Schembri, R., & Momigliano, P. (2024). Boundary Effects Cause False Signals of Range Expansions in Population Genomic Data. *Molecular Biology and Evolution*, 41(5), msae091. https://doi.org/10.1093/molbev/msae091

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *The American Naturalist*, 150(1), 1–23. https://doi.org/10.1086/286054

Klopfstein, S., Currat, M., & Excoffier, L. (2006). The Fate of Mutations Surfing on the Wave of a Range Expansion. *Molecular Biology and Evolution*, 23(3), 482–490. https://doi.org/10.1093/molbev/msj057

Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4), 1359–1368.

Kottler, E. J., Dickman, E. E., Sexton, J. P., Emery, N. C., & Franks, S. J. (2021). Draining the swamping hypothesis: Little evidence that gene flow reduces fitness at range edges. *Trends in Ecology & Evolution*, 36(6), 533–544.

Lawson, D. J., van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, 9(1), 3258. https://doi.org/10.1038/s41467-018-05257-7

Leisler, B., & Schulze-Hagen, K. (2015). *The Reed Warblers: Diversity in a Uniform Bird Family*. BRILL.

Leivo, O. (1937). Lampikertun, *Acrocephalus s. scirpaceus* (Herm.), esiintymisestä Suomessa. *Ornis Fennica*, 14(3–4), 81-90-81–90.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv Preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.

Liedvogel, M., Åkesson, S., & Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology & Evolution*, 26(11), 561–569. https://doi.org/10.1016/j.tree.2011.07.009

Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*, 19(3), 639–647. https://doi.org/10.1111/1755-0998.12995

Løppenthin, B. H. O. F. (1967). *Danske ynglefugle i fortid og nutid*. Odense Universitetsforlag.

Løvenskiold, H. L. (1947). *Håndbok over Norges fugl*er. Gyldendal norsk forlag, Oslo.

MacLean, S. A., & Beissinger, S. R. (2017). Species' traits as predictors of range shifts under contemporary climate change: A review and meta-analysis. *Global Change Biology*, 23(10), 4094–4105. https://doi.org/10.1111/gcb.13736

Mair, L., Hill, J. K., Fox, R., Botham, M., Brereton, T., & Thomas, C. D. (2014). Abundance changes and habitat availability drive species' responses to climate change. *Nature Climate Change*, 4(2), 127–131. https://doi.org/10.1038/nclimate2086

Malchevsky, A. S., & Pukinsky, Y. B. (1983). *The birds of Leningrad region and adjacent areas. Vol. 2. Passerines*. Leningrad Univ. Press, Leningrad (in Russian).

https://libarch.nmu.org.ua/bitstream/handle/GenofondUA/65169/d41bcccf80d3f2e7049869ca016a9b2e.pdf?sequence=1

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, 35(5), 1284–1290.

Mang, A. (1936). Tiigiroolinnu pesitsemisest Pärnu rannas. Eesti Loodus, 4, 140–142.

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867–2873.

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, 24(13), 3223–3231. https://doi.org/10.1111/mec.13243

Mills, L. S., & Allendorf, F. W. (1996). The One-Migrant-per-Generation Rule in Conservation and Management. *Conservation Biology*, 10(6), 1509–1518. https://doi.org/10.1046/j.1523-1739.1996.10061509.x

Moreau, C., Bhérer, C., Vézina, H., Jomphe, M., Labuda, D., & Excoffier, L. (2011). Deep Human Genealogies Reveal a Selective Advantage to Be on an Expanding Wave Front. *Science*, 334(6059), 1148–1150. https://doi.org/10.1126/science.1212880

Mukhin, A., Kosarev, V., & Ktitorov, P. (2005). Nocturnal life of young songbirds well before migration. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1572), 1535-1539.

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, 29(1), 1–10. https://doi.org/10.2307/2407137

Newton, I. (2023). *The migration ecology of birds*. Elsevier.

Nielsen, B., & Bensch, S. (1995). Post-fledging movements of juvenile Reed Warblers *Acrocephalus scirpaceus* and Sedge Warblers *Acrocephalus schoenobaenus*. *Ornis Svecica*, *5*(3–4), 125-131.

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, 27(16), 3193–3206. https://doi.org/10.1111/mec.14792

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, 22(24), 5983–5999. https://doi.org/10.1111/mec.12561

Pacifici, M., Rondinini, C., Rhodes, J. R., Burbidge, A. A., Cristiano, A., Watson, J. E. M., Woinarski, J. C. Z., & Di Marco, M. (2020). Global correlates of range contractions and expansions in terrestrial mammals. *Nature Communications*, 11(1), 2840. https://doi.org/10.1038/s41467-020-16684-w

Paradis, E., Baillie, S. R., Sutherland, W. J., & Gregory, R. D. (1998). Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology,* 67(4), 518–536. https://doi.org/10.1046/j.1365-2656.1998.00215.x

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), Article 6918. https://doi.org/10.1038/nature01286

Peischl, S., Kirkpatrick, M., & Excoffier, L. (2015). Expansion Load and the Evolutionary Dynamics of a Species Range. *The American Naturalist*, 185(4), E81–E93. https://doi.org/10.1086/680220

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13(5), 946–952. https://doi.org/10.1111/1755-0998.12129

Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution*, 67(11), 3274–3289. https://doi.org/10.1111/evo.12202

Peter, B. M., & Slatkin, M. (2015). The effective founder effect in a spatially expanding population. *Evolution*, 69(3), 721–734. https://doi.org/10.1111/evo.12609

Phillips, B. L., Brown, G. P., Webb, J. K., & Shine, R. (2006). Invasion and the evolution of speed in toads. *Nature*, 439(7078), 803–803. https://doi.org/10.1038/439803a

Pierce, A. A., Zalucki, M. P., Bangura, M., Udawatta, M., Kronforst, M. R., Altizer, S., Haeger, J. F., & de Roode, J. C. (2014). Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 281(1797), 20142230. https://doi.org/10.1098/rspb.2014.2230

Pitkänen, H., Peuraniemi, M., Westerbom, M., Kilpi, M., & von Numers, M. (2013). Long-term changes in distribution and frequency of aquatic vascular plants and charophytes in an estuary in the Baltic Sea. *Annales Botanici Fennici,* 50, 1–54.

Procházka, P., Brlík, V., Yohannes, E., Meister, B., Auerswald, J., Ilieva, M., & Hahn, S. (2018). Across a migratory divide: Divergent migration directions and non-breeding grounds of Eurasian reed warblers revealed by geolocators and stable isotopes. *Journal of Avian Biology,* 49(6), jav-012516. https://doi.org/10.1111/jav.01769

Procházka, P., Stokke, B. G., Jensen, H., Fainová, D., Bellinvia, E., Fossøy, F., Vikan, J. R., Bryja, J., & Soler, M. (2011). Low genetic differentiation among reed warbler *Acrocephalus scirpaceus* populations across Europe. *Journal of Avian Biology*, 42(2), 103–113.

Pujol, B., & Pannell, J. R. (2008). Reduced responses to selection after species range expansion. *Science*, 321(5885), 96–96.

Ramos, J. E., Pecl, G. T., Moltschaniwskyj, N. A., Semmens, J. M., Souza, C. A., & Strugnell, J. M. (2018). Population genetic signatures of a climate change driven marine range extension. *Scientific Reports*, 8(1), Article 1. https://doi.org/10.1038/s41598-018-27351-y

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Risberg, L. (1990). *Sveriges fåglar: Aktuell översikt över deras utbredning, numerär och flyttning samt något om svensk ornitologi.* Sveriges Ornitologiska Förening.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754.

Røed, T. (1994): *Rørsanger* — In: Norsk Fugleatlas: 382–83. Norsk Ornitologisk Forening. (Available online: https://www.birdlife.no/fuglekunnskap/fugleatlas/pdf/rorsanger.pdf)

Rózsa, J., Strand, T. M., Montadert, M., Kozma, R., & Höglund, J. (2016). Effects of a range expansion on adaptive and neutral genetic diversity in dispersal limited Hazel grouse (*Bonasa bonasia)* in the French Alps. *Conservation Genetics*, 17(2), 401–412.

Sætre, C. L. C., Eroukhmanoff, F., Rönkä, K., Kluen, E., Thorogood, R., Torrance, J., Tracey, A., Chow, W., Pelan, S., & Howe, K. (2021). A Chromosome-Level Genome Assembly of the Reed Warbler (A*crocephalus scirpaceus*). *Genome Biology and Evolution*, 13(9), evab212.

Sætre, C. L. C., Rönkä, K., Ravinet, M., Vitulano, S., Caldarella, M., Procházka, P., Stokke, B., Cuevas, A., Bergman, N., Thorogood, R., Jakobsen, K. S., Tørresen, O. K., & Eroukhmanoff, F. (2022). Genomic

consequences of range expansion and colonisation in the reed warbler (*Acrocephalus scirpaceus*) (p. 2022.11.28.518135). bioRxiv. https://doi.org/10.1101/2022.11.28.518135

Schulte, U., Veith, M., Mingo, V., Modica, C., & Hochkirch, A. (2013). Strong genetic differentiation due to multiple founder events during a recent range expansion of an introduced wall lizard population. *Biological Invasions,* 15(12), 2639–2649.

Sits, E. (1936). Kahe huvitava linnu leiust Eestis. *Loodusuurijate Seltsi Aruanded*, 42, 63–69.

Song, G., Yu, L., Gao, B., Zhang, R., Qu, Y., Lambert, D. M., Li, S., Zhou, T., & Lei, F. (2013). Gene flow maintains genetic diversity and colonization potential in recently range-expanded populations of an Oriental bird, the Light-vented Bulbul (Pycnonotus sinensis, Aves: Pycnonotidae). Diversity and Distributions, 19(10), 1248–1262. https://doi.org/10.1111/ddi.12067

Svensson, S., Svensson, M., & Tjernberg, M. (1999). *Svensk fågelatlas*. Sveriges ornitologiska fören.

Swaegers, J., Mergeay, J., Therry, L., Larmuseau, M., Bonte, D., & Stoks, R. (2013). Rapid range expansion increases genetic differentiation while causing limited reduction in genetic diversity in a damselfly. *Heredity*, 111(5), 422–429.

Szpiech, Z. A., Jakobsson, M., & Rosenberg, N. A. (2008). ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, 24(21), 2498–2504. https://doi.org/10.1093/bioinformatics/btn478

Szűcs, M., Melbourne, B. A., Tuff, T., Weiss-Lehman, C., & Hufbauer, R. A. (2017). Genetic and demographic founder effects have long-term fitness consequences for colonising populations. *Ecology Letters*, 20(4), 436–444. https://doi.org/10.1111/ele.12743

Thorogood, R., Mustonen, V., Aleixo, A., Aphalo, P. J., Asiegbu, F. O., Cabeza, M., Cairns, J., Candolin, U., Cardoso, P., Eronen, J. T., Hällfors, M., Hovatta, I., Juslén, A., Kovalchuk, A., Kulmuni, J., Kuula, L., Mäkipää, R., Ovaskainen, O., Pesonen, A.-K., … Vanhatalo, J. (2023). Understanding and applying biological resilience, from genes to ecosystems. *Npj Biodiversity,* 2(1), 1–13. https://doi.org/10.1038/s44185-023-00022-6

Tischler, F. (1914). *Die Vögel der Provinz Ostpreussen.* Springer.

Tischler, F. (1941). *Die Vögel Ostpreußens und seiner Nachbargebiete.* Ost-Europa-Verlag.

Virkkala, R., Luoto, M., Heikkinen, R. K., & Leikola, N. (2005). Distribution patterns of boreal marshland birds: Modelling the relationships to land cover and climate. *Journal of Biogeography*, 32(11), 1957–1970.

von Numers, M. (2011). Sea shore plants of the SW archipelago of Finland—Distribution patterns and long-term changes during the 20th century. *Annales Botanici Fennici*, 48, 1–46.

Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., Mock, T., Schmid, K., Schmitt, I., & Bataillon, T. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters*, 4(1), 4–18.

Wallingford, P. D., Morelli, T. L., Allen, J. M., Beaury, E. M., Blumenthal, D. M., Bradley, B. A., Dukes, J. S., Early, R., Fusco, E. J., & Goldberg, D. E. (2020). Adjusting the lens of invasion biology to focus on the impacts of climate-driven range shifts. *Nature Climate Change*, 10(5), 398–405.

Wang, J. (2018). Effects of sampling close relatives on some elementary population genetics analyses. *Molecular Ecology Resources*, 18(1), 41–54. https://doi.org/10.1111/1755-0998.12708

Weatherhead, P. J., & Forbes, M. R. L. (1994). Natal philopatry in passerine birds: Genetic or ecological influences? *Behavioral Ecology,* 5(4), 426–433. https://doi.org/10.1093/beheco/5.4.426

Welles, S. R., & Dlugosch, K. M. (2019). Population genomics of colonization and invasion. *Population Genomics: Concepts, Approaches and Applications*, 655–683.

White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus)* in Ireland. *Molecular Ecology,* 22(11), 2971–2985. https://doi.org/10.1111/mec.12343

Wikström, D. (1945). Lampikerttu, *Acrocephalus s. scirpaceus* (Herm.) pesinyt Suomessa jo v. 1922. *Ornis Fennica,* 22(1), 29-31-29–31.

Wynn, J., Padget, O., Mouritsen, H., Morford, J., Jaggers, P., & Guilford, T. (2022). Magnetic stop signs signal a European songbird's arrival at the breeding site after migration. *Science*, 375(6579), 446–449. https://doi.org/10.1126/science.abj4210

Wynn, J., Fandos, G., Delmore, K., Van Doren, B. M., Fransson, T., & Liedvogel, M. (2024). Could bi-axial orientation explain range expansion in a migratory songbird? *Journal of Avian Biology*, e03196.

Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., & Yang, T.-L. (2019). PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, 35(10), 1786–1788. https://doi.org/10.1093/bioinformatics/bty875