

# Global flyway evolution in red knots *Calidris canutus* and genetic evidence for a Nearctic refugium

Jesse R. Conklin<sup>1</sup> | Yvonne I. Verkuil<sup>1</sup> | Phil F. Battley<sup>2</sup> | Chris J. Hassell<sup>3</sup> | Job ten Horn<sup>4</sup> | James A. Johnson<sup>5</sup> | Pavel S. Tomkovich<sup>6</sup> | Allan J. Baker<sup>7,†</sup> | Theunis Piersma<sup>1,4</sup> | Michaël C. Fontaine<sup>1,8,9</sup>

<sup>1</sup>Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, The Netherlands

<sup>2</sup>Wildlife and Ecology Group, School of Natural Sciences, Massey University, Palmerston North, New Zealand

<sup>3</sup>Global Flyway Network, Broome, Western Australia, Australia

<sup>4</sup>Department of Coastal Systems, NIOZ Royal Netherlands Institute for Sea Research, Texel, The Netherlands

<sup>5</sup>U.S. Fish & Wildlife Service, Migratory Bird Management, Anchorage, Alaska, USA

<sup>6</sup>Zoological Museum, Moscow MV Lomonosov State University, Moscow, Russia

<sup>7</sup>Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada

<sup>8</sup>MIVEGEC, CNRS, IRD, University of Montpellier, Montpellier, France

<sup>9</sup>Montpellier Ecology and Evolution of Diseases Network (MEEDiN), Montpellier, France

## Correspondence

Jesse R. Conklin, Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, The Netherlands.

Email: conklin.jesse@gmail.com

## Funding information

Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: 824.01.001

Handling Editor: Yanhua qu

## Abstract

Present-day ecology and population structure are the legacies of past climate and habitat perturbations, and this is particularly true for species that are widely distributed at high latitudes. The red knot, *Calidris canutus*, is an arctic-breeding, long-distance migratory shorebird with six recognized subspecies defined by differences in morphology, migration behavior, and annual cycle phenology, in a global distribution thought to have arisen just since the last glacial maximum (LGM). We used nextRAD sequencing of 10,881 single-nucleotide polymorphisms (SNPs) to assess the neutral genetic structure and phylogeographic history of 172 red knots representing all known global breeding populations. Using population genetics approaches, including model-based scenario-testing in an approximate Bayesian computation (ABC) framework, we infer that red knots derive from two main lineages that diverged ca. 34,000 years ago, and thus most probably persisted at the LGM in both Palearctic and Nearctic refugia, followed by at least two instances of secondary contact and admixture. Within two Beringian subspecies (*C. c. roselaari* and *rogersi*), we detected previously unknown genetic structure among sub-populations sharing a migratory flyway, reflecting additional complexity in the phylogeographic history of the region. Conversely, we found very weak genetic differentiation between two Nearctic populations (*rufa* and *islandica*) with clearly divergent migratory phenotypes and little or no apparent contact throughout the annual cycle. Together, these results suggest that relative gene flow among migratory populations reflects a complex interplay of historical, geographical, and ecological factors.

## KEYWORDS

Bird migration, Climate change, Genetic differentiation, Genotyping-by-sequencing, Glacial refugia, Phylogeography, Population genomics

†Deceased.

## 1 | INTRODUCTION

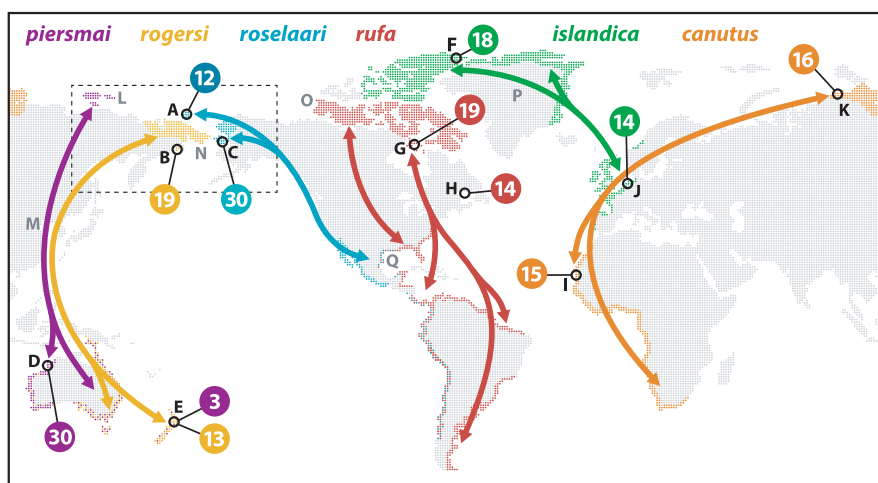
The ecology and demography of species are typically viewed through the lens of present-day or very recent (i.e., decades) observations and processes. However, it is clear that habitats, distributions, genetic diversity, and populations themselves are the legacies of events and conditions in historical or evolutionary time-scales, and cannot be fully understood without a phylogeographic perspective (Avice et al., 1987; Knowles, 2009). This is particularly true for high-latitude species, whose habitats were repeatedly transformed by Pleistocene glacial cycles, and whose current distributions may have arisen just since the last glacial maximum (LGM, c. 20,000 years before present, ybp; Hewitt, 2000). Of course, the effects of glaciations on historical distributions, and on population genetic diversity and structure, are strongly dependent on a species' ecological and physiological attributes (Stewart et al., 2010).

Migratory birds are characterized by seasonal movements and high dispersal capabilities, with important consequences for population structure, the strength of geographical barriers, and the propensity to colonize novel ranges (Winker, 2010). The effects of glacial perturbations on the geographic organization of migratory bird species and subspecies are profound, although the precise consequences are varied and debated (Avice & Walker, 1998; Johnson & Cicero, 2004; Klicka & Zink, 1997; Weir & Schluter, 2004). For example, it remains unclear whether tundra-specialized populations were most restricted and isolated (i.e., in "refugia") during the LGM, when vast areas of the Arctic were under ice, or during the mid-Holocene Climatic Optimum (c. 8,000 ybp), when a period of elevated temperatures limited the extent of tundra habitat (Arcones et al., 2020; Stewart & Dalén, 2008; Wauchope et al., 2017). Regardless, the present-day global distributions of many arctic-breeding migratory bird species attest to broad post-glacial expansions (Kraaijeveld &

Nieboer, 2000) and rapid colonization of new "flyways" (i.e., networks of routes and sites used throughout the annual journey) in response to changing conditions (Piersma, 2011).

Understanding the intertwining effects of historical and present-day processes on migratory systems is best approached through population genetic and phylogeographic approaches based on genome-wide markers, and informed by detailed ecological knowledge (Orsini et al., 2013). Among the most ecologically well-studied migratory birds is the red knot, *Calidris canutus*, a globally-distributed shorebird that breeds on high-latitude (62–80°N) arctic tundra and uses temperate and tropical intertidal mudflats during the rest of the year (Piersma, 2007; Piersma & Davidson, 1992). There are six recognized subspecies of red knot (Figure 1), distinguished by their breeding and nonbreeding ranges, and their migratory behavior, with one-way distances ranging from <3,000 to >14,000 km (Piersma, 2007; Piersma, Rogers, et al., 2005) and including some of the longest non-stop flights recorded in birds (Conklin et al., 2017). These populations differ in body size, plumage, and degree of spatial and behavioural overlap with neighbouring populations throughout the annual cycle (Buehler & Piersma, 2008; Tomkovich, 1992, 2001). Census population sizes vary by more than an order of magnitude (<20,000 to >400,000 individuals; Wetlands International, 2021), but most have faced significant declines in recent decades (Baker et al., 2004; Boyd & Piersma, 2001; Studts et al., 2017; van Gils et al., 2016), largely due to anthropogenic impacts on nonbreeding habitats (Baker et al., 2004; Piersma et al., 2016; Rakhimberdiev et al., 2015). Therefore, understanding the shared histories and magnitude of genetic, reproductive, and ecological separation among these populations is of significant conservation and evolutionary interest.

Observing shallow genetic differentiation and signals of historical bottlenecks in the mitochondrial control region (mtDNA), Buehler and Baker (2005) inferred a very recent origin of present-day red



**FIGURE 1** Global distribution and sampling of red knots. For each of six recognized subspecies (indicated by colour), arrows indicate general migration routes between breeding and (boreal) wintering areas (coloured areas). Numbers indicate total individuals sampled in each area (black circles). Letters refer to sampling sites (in black; Table S1 for details) and other locations mentioned in the text (in grey): A, Wrangel Island; B, SE Chukotka; C, Seward Peninsula; D, Roebuck Bay; E, Foxtton Beach; F, Ellesmere Island; G, Southampton Island; H, Mingan Archipelago; I, Banc d'Arguin; J, Wadden Sea; K, Taimyr Peninsula; L, New Siberian Islands; M, Yellow Sea; N, Bering Sea (Land Bridge); O, Banks Island; P, Greenland Ice Sheet; Q, Gulf of Mexico. Dashed box indicates approximate extent of the Beringia region

knot populations, and a global expansion from a single LGM refugium (Buehler et al., 2006). Notably, some subspecies pairs that were indistinguishable in mtDNA (Buehler & Baker, 2005) have phenotypic differences in morphology and migration distance, timing, and direction. In red knots, migratory behaviour is further associated with seemingly “hard-wired” endogenous annual rhythms in molt and mass, apparently adapted to flyway-specific conditions (Karagicheva et al., 2016; Piersma, 2011). This suggests that complex, multitrait migratory “syndromes” can arise in ecological timescales, and in the presence of gene flow (Delmore et al., 2020; Pérez-Tris et al., 2004). However, unravelling such recent, and potentially reticulated, evolutionary histories requires genome-wide information (Brito & Edwards, 2009; Narum et al., 2014) coupled with powerful model-based scenario-testing such as approximate Bayesian computation (ABC) (Beaumont, 2010; Bertorelle et al., 2010; Hickerson et al., 2010).

Here, we revisit population structure and phylogeography of red knots, to describe the history of divergences and degree of neutral genetic differentiation among global flyway populations. For this, we exploit recent research on red knot ecology and migration, which has clarified spatiotemporal overlap among populations (e.g. Atkinson et al., 2005; Carmona et al., 2013; Nebel et al., 2000; Verhoeven et al., 2016) and has made available more comprehensive global sampling associated with known migratory phenotypes. We used nextera-tagmented, reductively amplified DNA (nextRAD) sequencing (Russello et al., 2015) for *de novo* discovery of genome-wide single-nucleotide polymorphisms (SNPs) for population genetic analyses, and compared hypothesized evolutionary scenarios in an ABC framework using DIYABC (Cornuet et al., 2014). By reconstructing the recent evolutionary history of red knots, we (1) revise our understanding of LGM refugia and the colonization of global flyways, (2) reveal previously unrecognized population structure, and (3) provide a foundation for understanding how geography and ecology interact to regulate gene flow among migratory populations.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling and DNA extraction

We assembled DNA samples representing all recognized and hypothesized breeding populations within the global range of red knots (Figure 1, Table S1). Where possible, we used samples collected in known breeding areas; for *C. c. roselaari* (hereafter, we refer to populations by only their subspecific epithets) these included samples from two disjunct breeding areas: Wrangel Island, Russia and Seward Peninsula, Alaska, USA. We treated these as separate groups (*roselaari* West (W) and East (E), respectively), to test whether *roselaari* should be considered one or two independent demographic units. Because red knots breed in low densities in remote areas of arctic tundra, for some populations there were few or no breeding samples available. In these cases, we used samples collected from nonbreeding areas (i.e., sites used during migration and/or the boreal winter)

when sampled individuals could be confidently assigned to breeding areas either because they were remotely tracked to breeding areas using light-level geolocation, or because long-term research programmes had established strong links between breeding and nonbreeding areas (e.g., through mark-recapture/resight programmes). All tissue samples were acquired from museum collections or collected by the authors and colleagues in the field under all requisite permits appropriate to their respective countries and institutions.

We extracted genomic DNA from samples using three methods. For blood or organ tissue samples preserved in 95% ethanol, we used the DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's instructions for tissue. For blood samples preserved in Queen's lysis buffer, we used the NucleoSpin Blood QuickPure Kit (Macherey-Nagel). For feather samples and blood stored on filter paper, we used ammonium acetate precipitation (Richardson et al., 2001). Extract quality was first assessed on a 1.5% agarose gel to exclude extractions with insufficient yield or excessively degraded DNA. We then quantified DNA concentrations using a Qubit 3.0 fluorometer (Life Technologies), diluted extracts to achieve relatively even concentrations, and dried down samples in a SpeedVac concentrator. We delivered 57–168 ng of DNA of 203 individual knots (Table S1) for SNP discovery and genotyping.

### 2.2 | SNP genotyping using nextRAD sequencing

Genomic DNA was converted into nextRAD genomic fragment libraries (SNPsaurus, LLC, USA) following the method described by Russello et al. (2015). For each sample, 20–30 ng of genomic DNA was first fragmented with Nextera reagent (Illumina, Inc., USA), which also ligates short adapter sequences to the ends of the fragments. Fragmented DNA was then amplified, with primers matching the adapter and one primer extending 10 nucleotides into the genomic DNA with the selective sequence “GTGTAGAGCC”. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer were efficiently amplified. PCR amplification was done at 74°C for 27 cycles. The nextRAD libraries were sequenced on an Illumina HiSeq-4000 at the Genomics Core Facility, University of Oregon, USA.

Genotyping was performed using custom scripts (SNPsaurus, LLC). First, reads were trimmed in bbduk (BBMap tools; Bushnell, 2016). Next, a *de novo* reference was created from abundant reads (after removal of low-quality (phred-scale quality <20) and very high-abundance reads) and reads that aligned to these. All 161,810,193 reads were mapped to the reference with an alignment identity threshold of 95% using bbmap (BBMap tools). Genotype calling was performed using SAMtools and BCFtools (samtools mpileup -gu -Q 10 -t DP, DPR -f ref.fasta -b samples.txt bcftools call -cv -> genotypes.vcf), applying a minimum read depth filter of 7×. The genotype table was then filtered using VCFtools v.0.1.14 (Danecek et al., 2011) to remove SNPs called in <80% of samples and putative minor alleles with frequency <3% (allele counts <10), to exclude artefactual variants. The resulting VCF file included 4,911 unique loci (150-bp

sequences) containing 14,903 SNPs. Additional filtering was performed using VCFtools to remove indels ( $n = 452$ ) and samples that failed to sequence ( $n = 13$  individuals). Next, to exclude potential genotyping errors, SNPs deviating from Hardy-Weinberg proportions in at least six of seven hypothesized populations were identified and removed using VCFtools ( $p < .05$ ;  $n = 58$  SNPs). To minimize linkage among loci, we used PLINK v.1.9 (Chang et al., 2015) to identify and remove all SNPs in linkage disequilibrium (LD;  $r^2 > 0.20$ ;  $n = 3,512$  SNPs). The LD-pruned data set included 192 individuals and 10,881 unlinked SNPs on 4,679 loci.

We then used VCFtools to calculate proportion of missing SNP calls per individual (range 1%–87%) and removed 12 individuals with >25% missing data. Because inclusion of related individuals can bias population genetic analyses (Rodríguez-Ramilo & Wang, 2012), we estimated individual pairwise relatedness in PLINK using the identity-by-descent estimator PI HAT. Removal of eight individuals (6 *roselaari* E, 2 *rufa*) resolved all cases of relatedness involving half-siblings or closer (PI HAT >0.20). The final data set included 172 individuals and 10,881 SNPs. We used VCFtools, PLINK, and PGDspider v.2.1.0.3 (Lischer & Excoffier, 2012) to convert data to different formats required for analysis.

### 2.3 | Inference of population structure and diversity

To assess major axes of genetic variation and clustering among samples, we performed a principal component analysis (PCA; Patterson et al., 2006) using the R packages *gdsfmt* v.1.14.1 and *SNPRelate* v.1.12.2 (Zheng et al., 2012). We estimated ancestry proportions of each individual using the model-based clustering procedure of ADMIXTURE v.1.3.0 (Alexander et al., 2009). We performed 10 replicate runs (with random seeds) for each putative number of ancestral populations ( $K$ ) ranging from 1 to 8. We assessed how the cross-validation (CV) error rate varied with increasing  $K$  as a first assessment of the best  $K$  (Alexander & Lange, 2011). However, estimating  $K$  is known to be a difficult issue (Novembre, 2016) and this CV criterion alone is usually unable to separate closely related populations (Alexander & Lange, 2011; Lawson et al., 2018). Therefore, following the recommendations of Pritchard et al. (2010), and Lawson et al. (2018), we here used ADMIXTURE as an exploratory tool, inspecting newly created clusters and the stability of admixture proportions with increasing  $K$ , and compared the results with those from the PCA and with prior biological knowledge. We used the CLUMPAK (Cluster Markov Packager Across  $K$ ; Kopelman et al., 2015) web server (<http://clumpak.tau.ac.il/>) with default settings to summarize estimates of individual ancestry proportions to each cluster across replicate runs, and visualize the most likely ancestry proportions at each value of  $K$ .

Globally and for each putative population, we characterized genetic diversity by calculating per-site nucleotide diversity ( $\pi$ ), heterozygosity, and inbreeding coefficient ( $F_{IS}$ ) using VCFtools. To detect deviations from mutation-drift equilibrium, we estimated the

per-locus Tajima's  $D$  values for each population using VCFtools. We estimated genetic differentiation among populations by calculating pairwise  $F_{ST}$  (Weir & Cockerham, 1984) using the *diffCalc* function in the R package *diveRsity* v.1.9.90 (Keenan et al., 2013), with 95% confidence intervals (CI) derived from 500 bootstraps. We calculated  $p$ -values for  $F_{ST}$  estimates using the *pairwiseTest* function (1,000 permutations) in the R package *strataG* v.2.4.905 (Archer et al., 2017).

### 2.4 | Population evolutionary relationships and demographic history

To visualize evolutionary relationships among populations, we first constructed an unrooted neighbour-joining (NJ) tree based on Nei's genetic distance (Nei, 1972) using the R packages *poppr* v.2.8.5 (Kamvar et al., 2014) and *ape* v.5.3 (Paradis & Schliep, 2019), with missing data replaced by mean allele counts, and node support calculated from 1,000 bootstraps.

We used TreeMix v.1.13 (Pickrell & Pritchard, 2012) to explore the most likely population topology while accounting for possible gene flow or admixture among branches. Using genome-wide allele frequency data, TreeMix estimates a maximum-likelihood tree of populations with the nodes and branch lengths representing the amount genetic variance (or drift) shared among populations and within each one, respectively. Migration edges are added in a step-wise manner among populations, minimizing the genetic covariance unexplained by the tree. We performed 10 replicate TreeMix runs for each value of  $m$  (migration edges) from 1 to 10, and evaluated the change in log-likelihood as  $m$  was increased ( $\Delta m$ ) using the R package *OptM* v.0.1.3 (Fitak, 2019) to determine an optimal  $m$  value ranging from two to four. For each value of  $m$  (2–4), we then performed 1,000 bootstrap replicates in TreeMix using blocks of 500 SNPs to derive a consensus population tree topology with node supports using the R package *BITE* v.1.2.0008 (Milanesi et al., 2017). A final run was conducted using the consensus tree and the optimized number of migration edges ( $m = 2, 3$  or 4) using the full data set and the options “-tf consensus.tree -se”, to generate the final population graph with bootstrapped node supports as described in Milanesi et al. (2017).

We evaluated the support for different possible scenarios of population divergence and admixture using the approximate Bayesian computation (ABC) (Beaumont et al., 2002) random forest (RF) statistical framework (Pudlo et al., 2016; Raynal et al., 2019). ABC-RF can estimate posterior probabilities of historical scenarios, based on coalescent simulations of genetic data. Simulations are compared to observed data using summary statistics to identify the best-fitting model by calculating the number of RF votes and to derive the posterior probability for the best model. The best-fitting posterior parameter distribution values for the best model can be estimated using a RF procedure applied in a regression setting (Raynal et al., 2019). Estimated parameters include the effective size ( $N_e$ ) for each population, split times among populations ( $t$ ), and the timing ( $t$ ) and

rates ( $r$ ) of admixture events. We reduced the data set to all 4,126 SNPs genotyped in at least one individual per population. We then evaluated potential evolutionary scenarios in a stepwise manner, as follows.

In Step 1, we compared four scenarios (Figure S1, Table S2), aimed at resolving the backbone population topology, before further exploration of evolutionary scenarios. The first three scenarios (Figures S1a–c) represent three possible rootings of the unrooted topology inferred by TreeMix (see Results), including two admixture events representing the two best-supported migration edges (i.e., admixed origins of *piersmai* and *rufa/islandica*). The fourth scenario (Figure S1d) represents an alternative hypothesis for the uncertain relationships among Palearctic populations: using a topology consistent with the NJ tree and TreeMix, we included admixed origins of *canutus* and *rogersi*.

In Step 2, we started with the best-supported scenario from Step 1 (scenario c), and added additional complexity representing likely evolutionary scenarios (Figure S2, Table S3). First, we varied the age of the admixed origin of *piersmai* relative to other divergences (Figures S2a,b). Then, to each of these scenarios, we added admixture events representing the third (c,d) and fourth (e,f) best-supported migration edges inferred by TreeMix (i.e. admixed origins of *islandica* and *roselaari* E). For comparison, we created three additional scenarios representing alternative hypotheses for the origin of present-day populations (i.e. expert opinion); these included *islandica* diverging from *rufa* before (g,h) or after (i) admixture with the other Nearctic population *roselaari*, and an admixed origin of the Beringian populations *rogersi* (g,i) and/or *roselaari* W (g–i), rather than *piersmai* (see Figure S2).

The scenario parameters were considered as random variables drawn from prior distributions (Tables S2 and S3). We used DIYABC v.2.1.0 (Cornuet et al., 2014) to simulate 20,000 genetic data sets per scenario with the same properties as the observed data set (number of loci and proportion of missing data). Simulated and observed data sets were summarized using the whole set of summary statistics proposed by DIYABC for SNP markers, describing the genetic variation for each population (e.g., genetic diversity), pair of populations (e.g.,  $F_{ST}$  and Nei's distances), or trio of populations (e.g., admixture statistics) (see the full list and details in Table S4). Linear discriminant analysis (LDA) components were also used as additional summary statistics (Estoup et al., 2012). The total number of summary statistics was 268 and 272 for step 1 and 2, respectively.

We used the RF classification procedure to compare the likelihood of the competing scenarios at each step with the R package *abcrf* v.1.8.1 (Pudlo et al., 2016). RF is a machine-learning algorithm that uses hundreds of bootstrapped decision trees to perform classification, using the summary statistics as a set of predictor variables. Some simulations are not used in decision tree building

at each bootstrap (i.e., the out-of-bag simulations), and are used to compute the "prior error rate," which provides a direct method for estimating the CV error rate. At each step, we built a training set of 20,000 simulated data sets per scenario, with the same number of loci and individuals as the observed data set. We then grew a classification forest of 1,000 and 1,500 trees respectively for Step 1 and Step 2. The RF computation provides a classification vote for each scenario (i.e., the number of times a model is selected from the decision trees). We selected the scenario with the highest classification vote as the most likely scenario, and we estimated its posterior probability following the recommendation of Pudlo et al. (2016). We assessed the global performance of our chosen ABC-RF scenario, by calculating the prior error rate based on the available out-of-bag simulations and we repeated the RF analysis 10 times to ensure that the results converged.

Then, posterior distribution values of all parameters for the best model identified were estimated using a regression by RF methodology (Raynal et al., 2019), with classification forests of 1,000 decision trees, and based on a training set of 100,000 simulations. We converted estimates for timing parameters from generations to years assuming a generation time of 6 years (delayed maturity with adult annual survival c. 0.80; Méndez et al., 2018) and the genome-wide mutation rate calculated by Zhang et al. (2014) for Charadriiformes:  $1.5 \times 10^{-9}$  substitutions per site and per year. The simulation steps, computation of summary statistics, and model checking analysis were performed in DIYABC v.2.1.0. All scenario comparisons and estimations of parameter posterior distribution values were conducted with the R package *abcrf* v.1.8.1 (Pudlo et al., 2016; Raynal et al., 2019).

## 3 | RESULTS

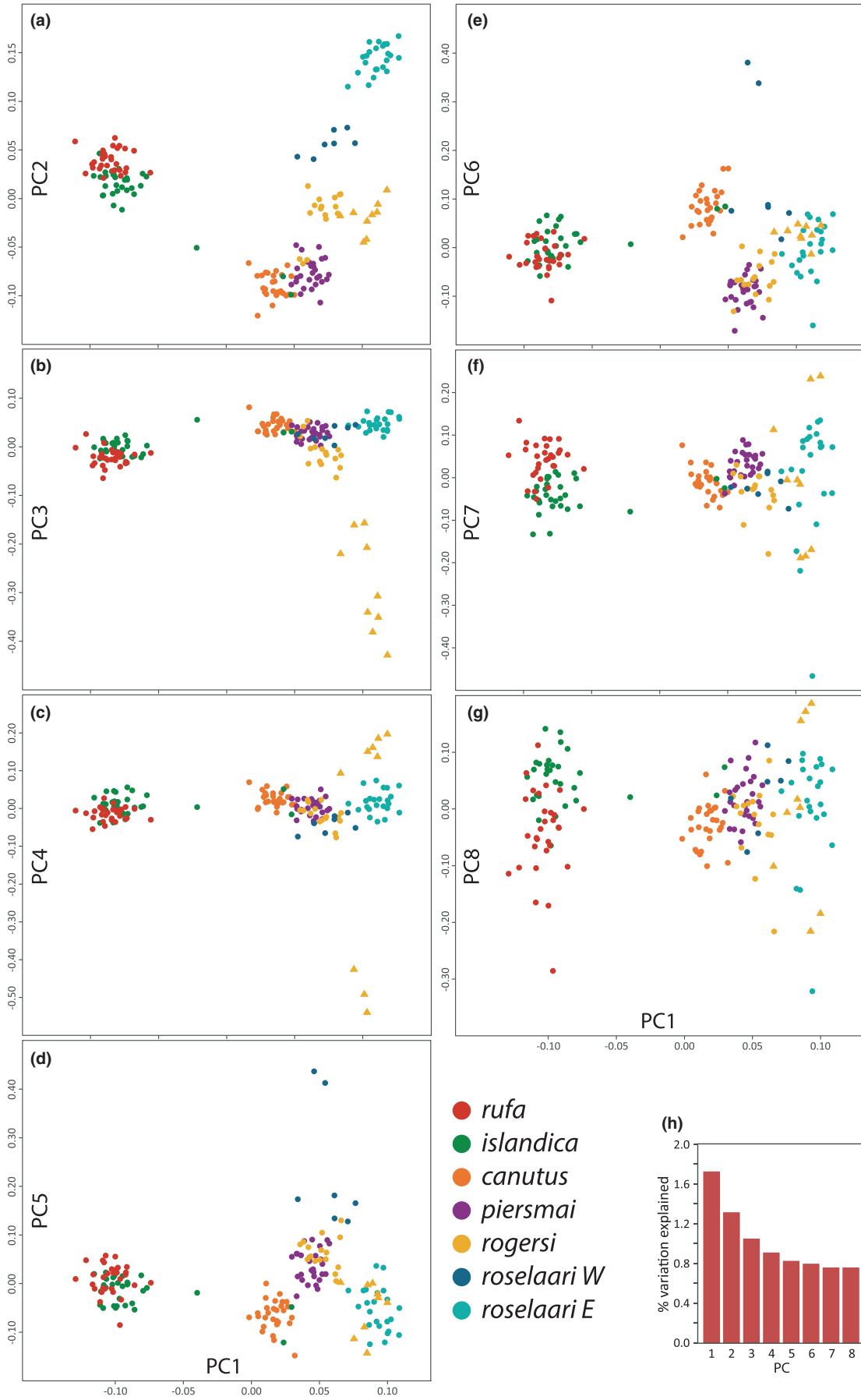
### 3.1 | Summary of nextRAD SNP data set

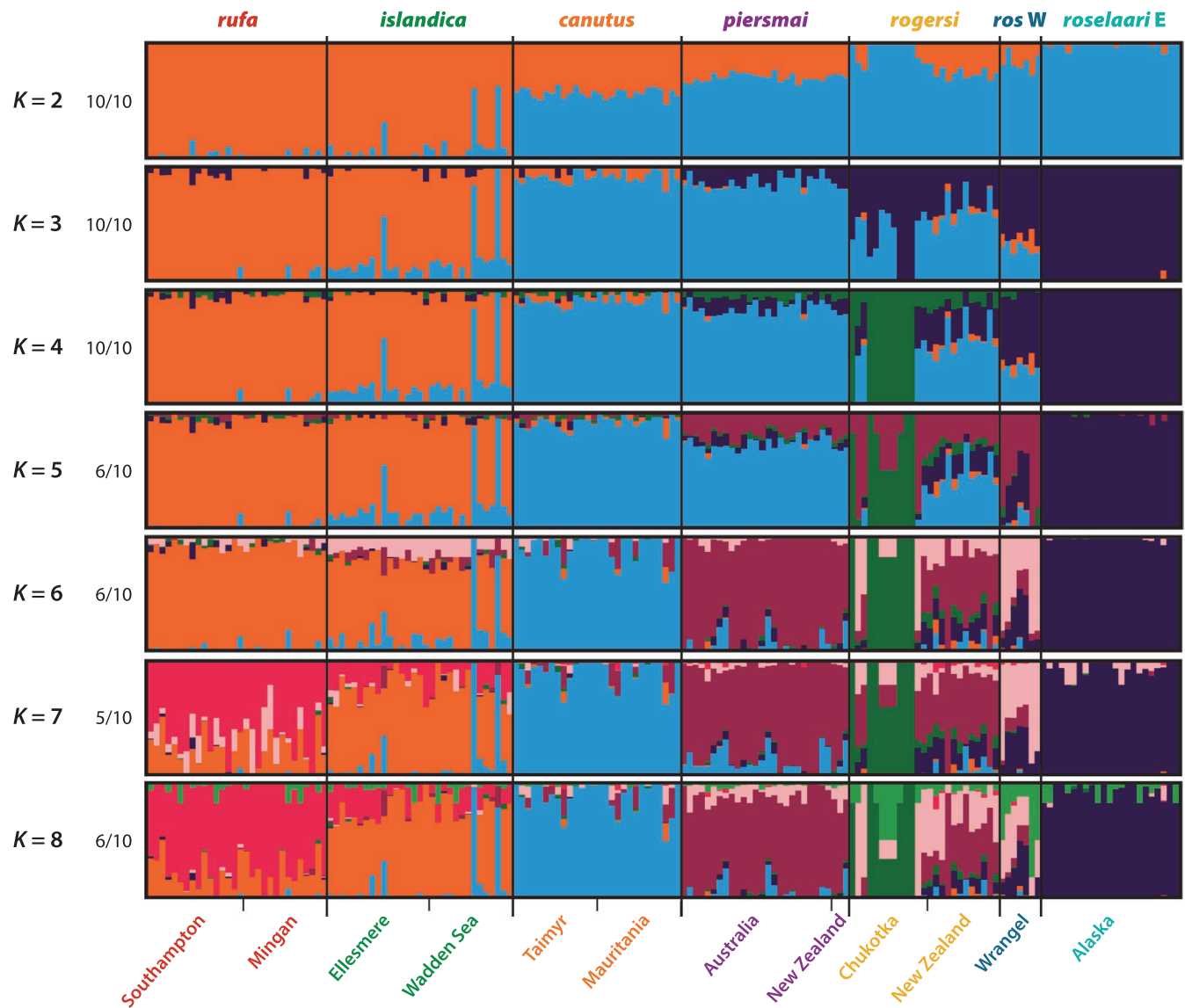
The final data set comprised 172 unrelated individuals, including 7–31 individuals in each of the seven hypothesized populations (Table S1), genotyped at 10,881 unlinked high-quality SNPs. On average, individuals were genotyped at 95.5% (range: 76%–99%) of SNPs and with a mean read depth of 57.3 (range: 17–137). Each SNP was genotyped in an average of 164.3 (range: 149–172, or 86.6%–100%) individuals. Globally, nucleotide diversity ( $\pi$ ) was 0.219, and observed heterozygosity was 18.7% (Figure S3, Table S5).

### 3.2 | Population structure and diversity

The first eight axes (PC1–8) of the principal component analysis explained 8.2% of the total genetic variation (Figure 2h). PC1 and PC2

**FIGURE 2** Population structure estimated by principal component analysis. The individual scores for the first principal component (PC1) are shown against the seven others (PC2–8) in panels (a) to (g). The scree plot (h) indicates the proportion of explained genetic variance by each PC, with PC1–8 explaining 8.2% of total variation. Note the nine individuals sampled in Chukotka (yellow triangles) that were distinguishable from the rest of *rogersi*. Also note two green and three yellow dots that fell among the *canutus/piersmai* cluster, indicating incorrect *a priori* identification of purported *islandica* and *rogersi* individuals, respectively





**FIGURE 3** Individual genetic ancestries assigned to major clusters for  $K = 2-8$  estimated using ADMIXTURE. At each value of  $K$ , the ancestry proportions for the 172 individuals for the dominant solution were determined by CLUMPAK summary of 10 replicate runs. Numbers on left indicate proportion of replicate runs contributing to the dominant solution (see Figure S4b for minor clusters for  $K = 5-8$ ). Names below plot indicate sampling locations (see Table S1)

distributed individuals into a triangular pattern (Figure 2a), suggesting three major genetic pools: a Canadian Arctic group of *rufa/islandica*, a central Palearctic group of *canutus/piersmai*, and a group formed by *roselaari E* from Alaska. Individuals from *rogersi* and *roselaari W* clustered in intermediate positions between the *canutus/piersmai* and *roselaari E* clusters, suggesting these populations may represent admixed groups between the two extreme clusters. Unexpectedly, PC3 separated nine of 13 Chukotka-breeding individuals from the main *rogersi* cluster (Figure 2b). Higher-order axes of variation (PC5–6) clearly distinguished *canutus* from *piersmai* (Figure 2d–e), and a subtle but noticeable separation of *rufa* and *islandica* was provided by all the PCs, especially on PC7–8 (Figure 2f–g).

Similar to the three major axes of variation identified by PCA, the CV error in the genetic ancestry analysis of ADMIXTURE conservatively suggested that there could be up to three ancestral genetic pools (Figure

S4a). Major clusters at  $K = 2-4$  had 100% support among 10 replicate runs, and greater values of  $K$  demonstrated further sub-structuring consistent with the PCA analysis, geography, or *a priori* hypotheses of population structure (see major clusters for  $K = 2-8$  in Figure 3; minor clusters for  $K = 5-8$  are shown in Figure S4b). Consistent with the PCA, individual ancestry estimated at  $K = 3$  identified three major genetic clusters: *rufa/islandica*, *canutus/piersmai*, and *roselaari E*, with *rogersi* and *roselaari W* displaying significant amounts of admixture. Interestingly, one individual sampled in the *islandica* breeding range displayed admixed genetic ancestry between *islandica* and *canutus* at  $K = 2-5$ , and therefore represents a possible F1 *islandica/canutus* hybrid (see also Figure 2). The nine Chukotka-breeding knots identified as distinct within the *rogersi* group on the third axis of the PCA (Figure 2b) were also recognized by the ADMIXTURE analysis as a distinct genetic cluster at  $K \geq 4$ ; other *rogersi* samples from Chukotka and New Zealand

contained only small proportions of this cluster.  $K = 5$  identified a cluster strongly present in Palearctic Beringian populations (*piersmai*, *rogersi*, *roselaari* W), but almost absent in Alaska (*roselaari* E). Greater values of  $K (\geq 6)$  illustrated the noticeable differentiation of *canutus* and *piersmai*, and the weak but noticeable genetic structure between *rufa* and *islandica* that was also identified in the PCA.

Both PCA and ADMIXTURE indicated that five individuals were mistakenly assigned to population based on *a priori* hypotheses, without any evidence of admixture: two purported *islandica* sampled during the winter in the Wadden Sea group with *canutus*, and three purported *rogersi* from New Zealand appear to be *piersmai* (Figures 2 and 3). Although these individuals have unknown breeding areas, we judge these to reflect errors in our assignment of nonbreeding individuals, rather than cases of true dispersal and potential gene flow. As there is no evidence these individuals contributed to the gene pool where they were *a priori* identified, we reassigned these five individuals to their "correct" or "original" populations for all subsequent analyses. Also, we hereafter consider the two *rogersi* clusters separately, with *rogersi* 2 comprising the nine Chukotka individuals differentiated in PC3, and *rogersi* 1 comprising the remaining 16 *rogersi* individuals. Aside from these cases, negligible differences between sampling locations presumed to represent the same population confirmed our *a priori* assumptions about population distributions (within *rufa*, *islandica*, *canutus*, and *piersmai*; see Figure 3, Table S6).

Globally, nucleotide diversity ( $\pi = 0.219$ ) was similar for all populations, but slightly lower in *rogersi* 2 ( $\pi = 0.184$ , Figure S3a, Table S5). Mean observed heterozygosity was also generally uniform among populations, but slightly higher and more variable in *rogersi* 1 and *roselaari* W (Figure S3b). We found some evidence of inbreeding (mean  $F_{IS} > 0.22$ ) in *rogersi* 2 and *roselaari* W (Figure S3c). Values of Tajima's  $D$  (range of means: 0.006–0.154) were indistinguishable from zero for all populations (Figure S3d), suggesting no strong deviations from neutral expectations.

Pairwise  $F_{ST}$  among the eight hypothesized populations ranged 0.005–0.058, and all differed significantly from zero ( $p < .01$ ; Table 1). The two lowest estimates involved pairs of previously recognized subspecies: *rufa* vs. *islandica* ( $F_{ST} = 0.005$ ) and *canutus* versus *piersmai* ( $F_{ST} = 0.007$ ). For all populations, the greatest pairwise differences were with *rogersi* 2 ( $F_{ST} = 0.036$ –0.058), including within purported *rogersi* (*rogersi* 1 vs. *rogersi* 2:  $F_{ST} = 0.036$ ). By contrast, other pairwise differences with *rogersi* 1 were lower (all  $\leq 0.021$ ). Comparisons with *roselaari* W also produced relatively large  $F_{ST}$  estimates (0.020–0.058).

### 3.3 | Population evolutionary relationships and demographic history

The neighbour-joining (NJ) tree based on Nei's distance (Figure 4a) identified four closely-related pairs of populations: *canutus/piersmai*, *rufa/islandica*, *roselaari* W/E, and *rogersi* 1/2, the last two pairs forming a Beringian group including *rogersi* and *roselaari*. These pairs clustered with high node support ( $>99.9\%$ ), except *canutus/piersmai* that had slightly

lower support (84.0%). Furthermore, the long branch length between the two clusters in *rogersi* confirmed that they are distinct but closely-related.

We further explored the population evolutionary relationships using TreeMix (Figures 4b, S5 and S6). To understand how the unexpected *rogersi* 2 cluster would affect topologies inferred by TreeMix, we ran this analysis with *rogersi* 2 either included or excluded. In both cases, change in likelihood and proportion of explained genetic variance with increasing number of migration edges ( $m$ ) indicated strong support for at least two migrations edges ( $m = 2$ ). Adding 3 or 4 migration edges further improved the explained variance and likelihood of the model, although the improvement was less pronounced (Figures S5 and S6); Figure 4b shows the topology inferred at  $m = 2$  with *rogersi* 2 included. Across all scenarios, the most consistently supported migration edge was from the *roselaari* branch to *rufa/islandica*. The branching relationships among *canutus*, *piersmai*, and *rogersi* varied across scenarios, and featured low node support ( $\leq 50\%$ ), but uncertainty in this part of the topology was consistently addressed with the inference of a migration edge to *piersmai* from either *canutus* or the *rogersi* branch. The third best-supported edge indicated migration from *canutus* to *islandica* (Figures S5 and S6). At  $m = 4$ , the scenarios with and without *rogersi* 2 differed in the inferred topology and migration edges, but both included a weakly-supported migration edge from the *rogersi* branch to *roselaari* E.

Finally, we compared the likelihood of alternative scenarios describing different population branching topologies and admixture events suggested by the above analyses using the approximate Bayesian computation – random forest approach (ABC-RF). Step 1 of the ABC-RF analysis (Figures 5 and S1) supported scenario *c*, representing the topology inferred by TreeMix, rooted such that the Canadian Arctic group (*rufa/islandica*) descended from a lineage that was the first to split from a group comprising all Palearctic and Alaskan populations (Figure 5c). This scenario received the greatest number of RF votes (33.0%), with a posterior probability of 45.5% and a prior error rate of 39.7% (Figure 5). The second best-fitting scenario, receiving 25.7% of the RF votes, was scenario *d*.

Step 2 in the ABC-RF built upon the best-supported scenario in Step 1 by comparing nine scenarios: six consisting of variations on scenario *c* from Step 1, each including additional admixture events as suggested by TreeMix, and three scenarios testing other plausible biogeographic hypotheses (Figure S2 and Table S3). The six closely-related variations from scenario *c* in Step 1 (scenarios *a* to *f* in Step 2; Figure S2) received collectively 79% of the RF votes. In contrast, the alternative scenarios (*g* to *i*; Figure S2) were least supported, each receiving  $\leq 9\%$  of the votes. The single best-supported scenario was *d* (Figure 6), with the greatest number of votes (17%), a posterior probability of 43.9%, and a prior error rate of 31.4% (Figure S2). This scenario included an older admixed origin of *piersmai* and an additional recent admixed origin of *islandica*; this reflects support for the third, but not the fourth, inferred migration edge from TreeMix. However, the second best-supported scenario (*f*, with 15% of votes; Figure S2), included the fourth TreeMix migration edge, providing weak support for recent admixture between *rogersi* and *roselaari*.



TABLE 1 Mean population pairwise  $F_{ST}$  and 95% confidence intervals of estimates. All comparisons are significantly different from zero (1,000 permutations; \* $p < .01$ , \*\* $p < .001$ )

	<i>rufa</i>	<i>islandica</i>	<i>canutus</i>	<i>piersmai</i>	<i>rogersi 1</i>	<i>rogersi 2</i>	<i>roselaari W</i>
<i>islandica</i>	0.005** [0–0.012]						
<i>canutus</i>	0.020** [0.014–0.028]	0.016** [0.009–0.023]					
<i>piersmai</i>	0.021** [0.016–0.029]	0.018** [0.012–0.026]	0.007** [0.001–0.014]				
<i>rogersi 1</i>	0.021** [0.009–0.043]	0.019** [0.007–0.040]	0.013** [0–0.032]	0.009** [0–0.031]			
<i>rogersi 2</i>	0.055** [0.035–0.086]	0.056** [0.037–0.086]	0.047** [0.028–0.078]	0.045** [0.025–0.075]	0.036** [0.006–0.076]		
<i>roselaari W</i>	0.032** [0.005–0.071]	0.031** [0.006–0.072]	0.029** [0.004–0.066]	0.026** [0.002–0.071]	0.020** [0–0.067]	0.058* [0.016–0.114]	
<i>roselaari E</i>	0.031** [0.024–0.044]	0.031** [0.023–0.043]	0.027** [0.019–0.039]	0.024** [0.017–0.037]	0.018** [0.004–0.042]	0.052** [0.031–0.084]	0.024** [0–0.063]

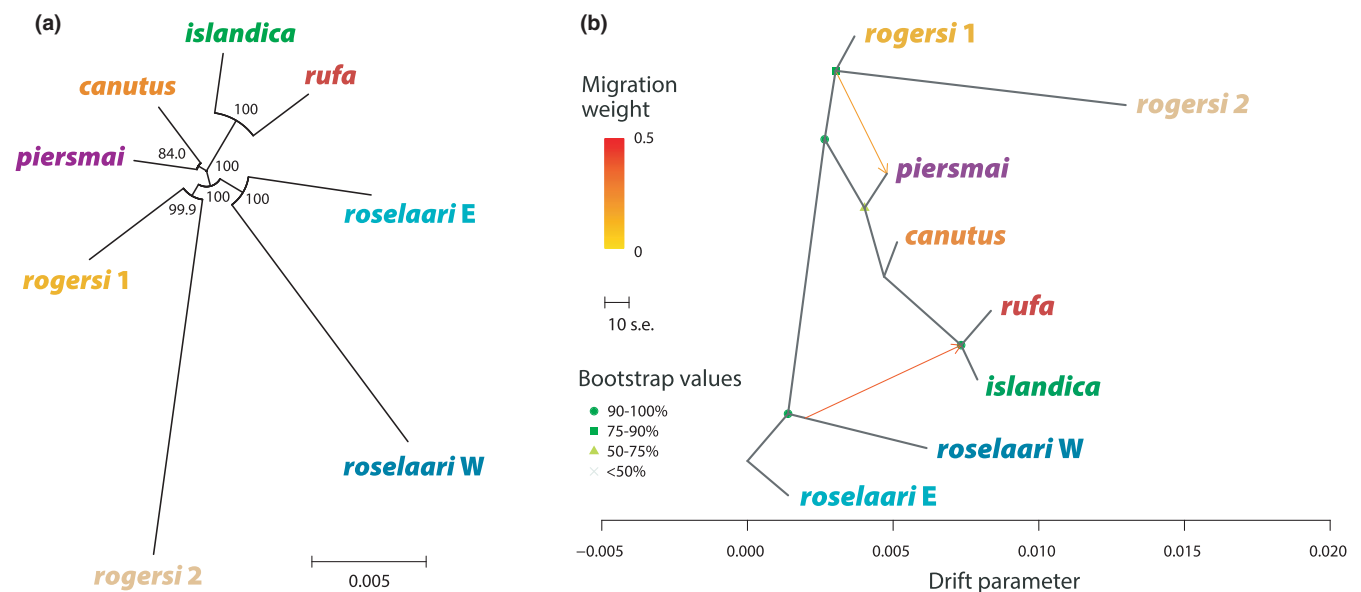
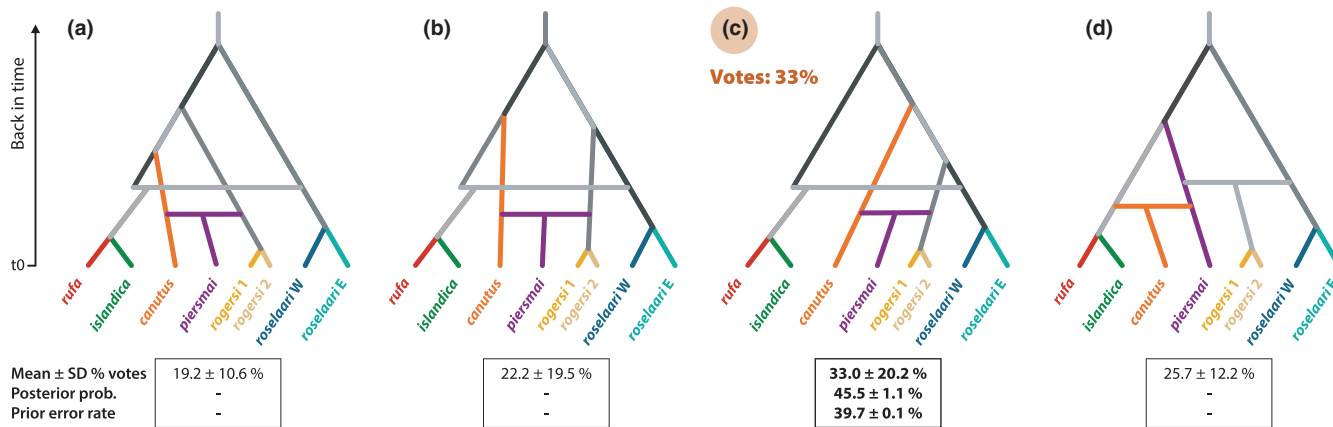


FIGURE 4 Evolutionary relationships among populations of red knots. (a) Unrooted neighbour-joining tree, based on Nei's minimum distance with bootstrapped node confidence (%). (b) Maximum-likelihood (ML) tree inferred by TreeMix, including two migration edges. See Figures S5 and S6 for TreeMix results with two to four migration edges

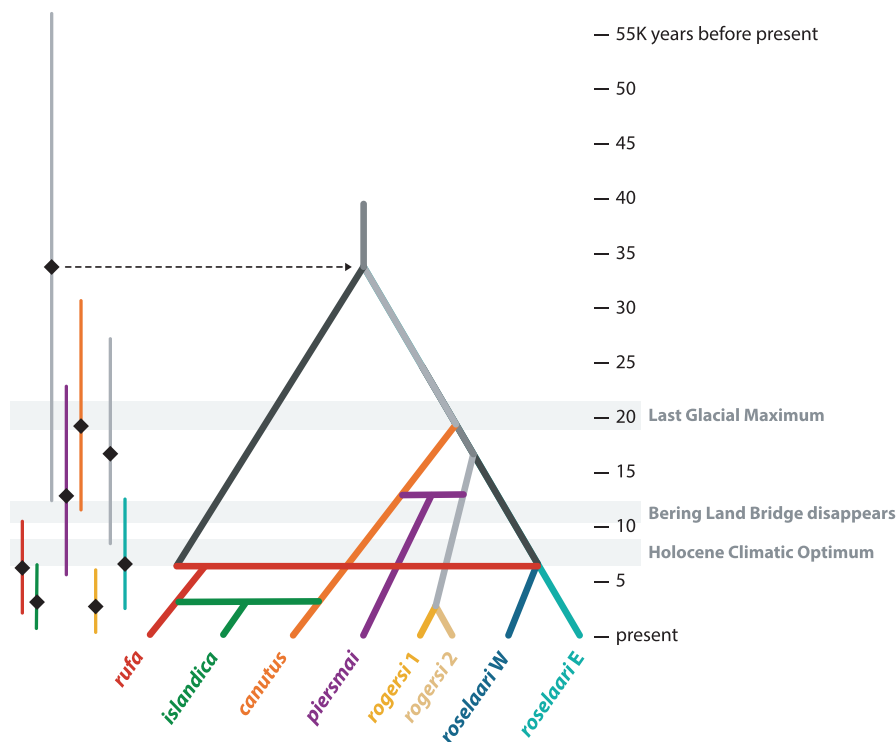
Each of the 25 demographic parameters of the best-supported scenario was estimated within the ABC-RF framework (Table S7). Timing parameter estimates support a pre-LGM divergence of the Canadian Arctic group from the Palearctic/Beringian group (mean estimate 33,718 ybp, 95% CI: 12,436–56,718; Figure 6, Table S7), followed by post-glacial divergences and admixture in both major branches. In particular, we infer recent divergences within *roselaari* (6,722 ybp, CI: 2,677–12,516) and *rogersi* (2,830 ybp, CI: 534–6,078), and two instances of secondary contact between the Nearctic and Palearctic groups, via *roselaari* (6,356 ybp, CI: 2,291–10,518) and via *canutus* (3,242 ybp, CI: 870–6,528; Figure 6). Estimated effective

population sizes of the eight extant populations (N1–8 in Table S7) ranged from 4,231 to 50,616 and captured some expected differences among populations. For example, the three lowest  $N_e$  estimates were for a highly endangered population (*rufa*) and two presumably small populations with signals of inbreeding and lower than expected heterozygosity (*rogersi 2*, *roselaari W*; Figures S3b–c). Conversely, the two admixed populations (*islandica*, *piersmai*) produced the highest  $N_e$  estimates. Estimated admixture rates were also consistent with results of ADMIXTURE and TreeMix: *islandica* derived 67% of its ancestry from *rufa* and 33% from *canutus* ( $r1$  in Table S7), whereas *piersmai* derived 33% from *canutus* and 67% from *rogersi* ( $r2$ ).



**FIGURE 5** Scenarios tested in Step 1 of DIYABC analysis: (a–c) three possible rootings of the population topology inferred by TreeMix, including two admixture events (horizontal branches); (d) an alternative scenario with hypothesized admixed origin of *canutus* and *rogersi*. Extant (sampled) populations are indicated by colours; inferred historical populations are shown in grey. The best-supported scenario was (c), as inferred by the proportion of Random Forest classification votes; for this scenario, posterior probabilities, and classification error rates are also indicated. See Figure S1 and Table S2 for further details on the model parameters

**FIGURE 6** Best-supported scenario in Step 2 of DIYABC analysis (see Figure S2), scaled to relative time-parameter estimates (converted to years assuming a generation time of 6 years) for five divergence events (branches) and three admixture events (horizontal bars). On left, each time parameter is indicated by an estimate (black diamond) and 95% confidence interval (vertical line)



## 4 | DISCUSSION

This study substantially revises our understanding of the nature, age, and origins of the current global distribution of red knots. Our historical reconstruction strongly supports the persistence of the species in two refugia at the LGM, followed by at least two instances of admixture after recent secondary contact between Nearctic and Palearctic lineages. In addition, we detected unrecognized structure within two populations currently regarded as subspecies. Below, we discuss the implications of our results for understanding: (1) the impacts of post-LGM environmental changes, (2) present-day

population structure, and (3) the potential drivers of isolation and gene flow among flyway populations.

### 4.1 | Post-glacial phylogeography of red knots

Previous population-genetic work in red knots (Buehler & Baker, 2005) suggested that the species was restricted to the Palearctic during the LGM, and only colonized the Canadian Arctic in the last few thousand years, after the Holocene Climatic Optimum (Buehler et al., 2006). This conclusion was partly based on extremely

low differentiation at the mtDNA control region among the three Nearctic subspecies (Buehler & Baker, 2005). However, the apparent close relationship of *roselaari* to *rufa* and *islandica* was an artefact of sampling – in fact, there were likely no *roselaari* in that study, as the purported *roselaari* samples came from the southeastern USA, now considered an exclusive *rufa* wintering area (Atkinson et al., 2005; Carmona et al., 2013; Verkuil et al., 2021). Using verifiable breeding samples, we identified *roselaari* and *rufa* as the most differentiated neighbouring population pair, making a scenario of eastward colonization of the Nearctic (Buehler et al., 2006) highly unlikely. This leaves two plausible scenarios: (1) after the LGM, *canutus* spread from Europe to colonize the Nearctic breeding range of *islandica*, and then gave rise to *rufa* through a southward expansion from Ellesmere Island; or (2) the *rufa/islandica* clade arose *in situ* from a Nearctic refugium. The former scenario is possibly consistent with estimates of  $F_{ST}$  (Table 1) and Nei's distance (Figure 4a), which suggest a closer relationship of *rufa/islandica* with the western Palearctic (*canutus/piersmai*) than with the Beringian group (*rogersi/roselaari*). However, the explicit consideration of admixture (Figures 3 and 4b) supports the latter scenario, in which a previously diverged Nearctic branch came into recent secondary contact with both western Palearctic and Beringian groups. Our DIYABC analysis corroborated this, indicating that a Nearctic ancestor diverged from the ancestor of the Palearctic/Beringian clade c. 34,000 ybp, and then experienced admixture from both the west and east in the last few thousand years (Figure 6). For comparison, the oldest population divergence estimated by Buehler and Baker (2005) was 20,000 ybp (95% CI: 5,600–58,000), consistent with an entirely post-LGM expansion from the Palearctic. The deeper history of divergences we describe suggests that red knots survived the LGM in both Palearctic and Nearctic refugia.

It is well established that a huge swath of the Arctic including most of Beringia and stretching westward almost to the Taimyr Peninsula in central Russia was largely ice-free at the LGM (Ehlers & Gibbard, 2007; Pielou, 1991), and this refugium probably gave rise to all Palearctic and Beringian red knot populations. Meanwhile, the Nearctic was predominantly glaciated at the LGM, including nearly all of the present-day breeding ranges of *islandica* and *rufa*. However, multiple lines of evidence support potential refugia in this region (Dyke, 2004; Fedorov & Stenseth, 2002; Provan & Bennett, 2008), including on Banks Island near the western end of the *rufa* range (Figure 1) and in northeast Greenland within the *islandica* range. Predictive modeling of habitats at the LGM suggested that Banks Island was most suitable for tundra-breeding shorebirds in the Canadian Arctic, whereas ice-free areas at higher latitudes were less suitable polar desert (Arcones et al., 2020). Much of the present-day high-Arctic breeding range of *islandica* was not ice-free until c. 6,000 ybp (Dyke, 2004). Our DIYABC analysis inferred that *islandica* arose c. 3,200 ybp, after Nearctic red knots had established contact with *roselaari* ca. 6,400 ybp (Figure 6), perhaps through non-breeding overlap in temperate North America. Thus, it is most likely that red knots persisted at the LGM in a refugium on or near Banks Island, and then expanded northeastward as glaciers retreated, colonizing

the present-day *islandica* range. After this, a new migration route to Europe was established, promoting contact and admixture with *canutus* in the last 3,000–4,000 years (Figure 6).

At the LGM, a southward migration from Banks Island to temperate or tropical non-breeding areas would require a flight of >2,500 km over the Laurentide Ice Sheet, which covered much of central North America (Dyke, 2004; Ehlers & Gibbard, 2007). Present-day migrations of all red knot subspecies involve non-stop flights of 4,000–10,000 km (Conklin et al., 2017). In particular, *islandica* individuals can fly 4,000 km from the Canadian Arctic to western Europe, including up to 1,700 km across the Greenland Ice Sheet (Kok et al., 2020). Therefore, we find a migration from Banks Island entirely plausible. Similarly, the historical migration of the westernmost Palearctic red knots (i.e., present-day *canutus*) likely involved flights across the ice sheets that covered northwestern Europe (Batchelor et al., 2019), which could explain their contemporary non-stop flights between Taimyr and the Wadden Sea coast (Figure 1), both of which were largely ice-free at the LGM (Ehlers & Gibbard, 2007).

As recognized by Buehler et al. (2006), the post-LGM diversification of at least four Palearctic/Beringian red knot populations (Figure 6) is consistent with a scenario of increasingly isolated patches of tundra breeding habitat in the warming period prior to the Holocene Climatic Optimum (Arcones et al., 2020; Kraaijeveld & Nieboer, 2000; Stewart & Dalén, 2008; Wauchope et al., 2017). The inferred admixed origin of *piersmai* (Figure 6) suggests that a population on the New Siberian Islands (Figure 1) received subsequent immigration from two previously isolated mainland populations (*canutus* and *rogersi*). Interestingly, unlike much of the present-day *rogersi* breeding range, the region of our Chukotka sampling site was at least partially glaciated at the LGM (Ehlers & Gibbard, 2007; Gualtieri et al., 2000), suggesting that the unexpected genetic cluster we detected (*rogersi* 2) could reflect a recent colonization of this disjunct breeding area (Figure 1). The warming period of retreating glaciers also featured rising sea levels which inundated the Bering Land Bridge, gradually isolating Alaska from Chukotka (and thus *roselaari* E from *rogersi*) by c. 12,000 ybp, and eventually Wrangel Island from mainland Russia (i.e., *roselaari* W from *rogersi*), ca. 10,000 ybp (Dyke, 2004; Manley, 2002). The population structure and uncertain tree topology we found in the eastern Palearctic attests to a recent history of climate and habitat upheavals in the region (McLaughlin et al., 2020). Expanded and more targeted genetic sampling may reveal further insights into this complicated regional history.

## 4.2 | Further implications for global population structure

We confirm the close relationship between the Nearctic subspecies *rufa* and *islandica*, which were effectively indistinguishable in mtDNA (Buehler & Baker, 2005). With genome-wide SNPs, we found weak but significant differentiation, which may reflect a

recent northeastward expansion (see above), but raises the question of whether the populations are demographically independent. Currently, the extent and potential overlap of the breeding ranges of *rufa* and *islandica* are poorly described (Lathrop et al., 2018), and mark-resight studies have recorded individuals apparently “switching” between the two flyways (Wilson et al., 2010). We sampled the two populations at the geographic extremes of the Canadian breeding range (Figure 1), and thus intermediate sampling might reveal more clinal variation by latitude across the ranges of *rufa* and *islandica*. Moreover, amplified fragment length polymorphism (AFLP) markers have indicated structure across the nonbreeding range of *rufa* (Verkuil et al., 2021), raising the possibility of longitudinal structure across the low Canadian Arctic. We note that our *rufa* samples came from easterly breeding and passage sites, and therefore do not include *rufa* that migrate through central North America (Figure 1; Newstead et al., 2013). This may also exaggerate the observed genetic differentiation between *rufa* and *roselaari* E, which are known to share nonbreeding areas in the Gulf of Mexico and perhaps western South America (Figure 1; Carmona et al., 2013; Gherardi-Fuentes et al., 2021). Elucidating this recent, perhaps dynamic, structure requires a greater breadth of sampling across the Canadian Arctic and Greenland, which is most feasible by tracking the migrations of knots captured at nonbreeding sites.

Red knots breeding in Alaska and on Wrangel Island have been considered one subspecies, *roselaari*, due to morphological similarity and a common migration route along the west coast of North America (Carmona et al., 2013; Tomkovich, 1992). We confirmed their close relationship by descent (Figure 4), but also observed substantial genetic differentiation (Table 1). Marked individuals from both breeding areas have been detected as far south as Sonora and Baja California, Mexico (Carmona et al., 2013), suggesting that *roselaari* W and E may largely overlap during the nonbreeding season, and that population estimates and conservation efforts therefore conflate two potentially independent demographic units. Given the distinct clustering of the two populations (Figures 2 and 3), large-scale molecular population assignment of individuals from the *roselaari* non-breeding range may help to elucidate the population sizes and degree of year-round spatial overlap of Wrangel Island and Alaskan knots.

The unexpected structure we detected within *rogersi* warrants further investigation. Our expectation was that nonbreeding *rogersi* samples from New Zealand and breeding samples from Meinypilgyno, in southeast Chukotka, would equally represent the population that breeds widely across northern Chukotka (see Tomkovich et al., 2013; Zöckler & O'Sullivan, 2005). However, only four of 13 Meinypilgyno individuals clustered with the New Zealand birds, while the remaining nine formed a distinct cluster (Figures and 3), and pairwise  $F_{ST}$  argued against a homogeneous population (Table S7). Meinypilgyno is located in a small, disjunct part of the breeding range, separated by >200 km from the closest known *rogersi* breeding areas (Figure 1; Lappo et al., 2012) and likely colonized after retreat of glaciers in the region (see above). In this light, the signals of inbreeding (Figure S3) and drift (i.e., long branch length in Figure 4) we found in *rogersi* 2 suggest a local founder effect, perhaps followed by recent

immigration to the area by *rogersi* 1. Further genetic sampling and migration tracking of knots from the main *rogersi* breeding range are needed to understand this intriguing structure.

Due to the lack of sufficient breeding samples for all populations, we used nonbreeding samples when current understanding allowed confident assignment to population. These assumptions were generally confirmed (Table S7), but two exceptions provide insight about unrecognized subspecies overlap in the nonbreeding season. First, it is thought that red knots spending the winter in Europe are *islandica*, because *canutus* only passes through this region en route to nonbreeding sites in western Africa (Dick et al., 1987). However, we identified two apparent *canutus* individuals among knots sampled in the Dutch Wadden Sea in mid-winter (14% of samples; Figures 2 and 3), implying that the wintering range of *canutus* extends into western Europe and overlaps with that of *islandica*. This has important implications for estimating population sizes and trends based on distinct wintering areas (van Roomen et al., 2015), and provides a simpler potential pathway for gene flow, as indicated by TreeMix and the one apparent F1 hybrid (*canutus* × *islandica*) sampled at the *islandica* breeding area on Ellesmere Island (see Figure 3).

Similarly, we identified three apparent *piersmai* in our sample of purported *rogersi* from New Zealand (Figures 2 and 3). These subspecies overlap clinally during the boreal winter, such that sites in western Australia contain >80% *piersmai*, whereas sites in New Zealand are ~80% *rogersi* (Piersma et al., 2021; Tomkovich & Riegen, 2000; Verhoeven et al., 2016). Plumage differences are considered sufficient to distinguish the subspecies (Hassell et al., 2011; Tomkovich, 2001), and we correctly verified the samples of *piersmai* from Broome, Australia, based on plumage recorded during re-sights of these individuals when migrating through Bohai Bay, China (Rogers et al., 2010). The identification of three unrecognized *piersmai* in our New Zealand sample of purported *rogersi* (Figures 2 and 3) demonstrates that not all *piersmai* individuals attain a distinguishable plumage prior to northward migration, and can thus be mistaken for the duller-plumaged *rogersi* subspecies at that time of year.

### 4.3 | Implications for the flexibility and isolating function of migratory phenotypes

The study of population structure and differentiation in migratory taxa is particularly intriguing, due to the presumably contradictory influences of high mobility and increased phenotypic specialization (Winker, 2010). On one hand, high dispersal and the ability to cross potential geographic barriers, such as oceans and mountain ranges, should promote gene flow and weaken structure among migratory bird populations. On the other hand, selection for flyway-specific adaptations should reduce both gene flow and successful colonization of new geographic areas. In migratory birds, fitness depends on a multi-trait phenotype (encompassing flight, fueling, navigation, timing, molt, etc.; Åkesson et al., 2017; Piersma et al., 2005), of which some components, such as migration direction and circannual rhythms, are to some extent

heritable and endogenously entrained (Berthold & Helbig, 1992; Piersma, 2011). Despite the inflexibility this may imply, novel migration behavior can evolve rapidly in new circumstances (Able & Belthoff, 1998; Berthold et al., 1992), and migratory differences may arise and persist in sympatry without precluding gene flow (Delmore et al., 2020; Pérez-Tris et al., 2004).

In red knots, we found the weakest genetic differentiation precisely where neighboring populations display the greatest phenotypic differences: between *rufa* and *islandica*. These subspecies migrate, respectively, the longest (up to c. 15,000 km one-way) and shortest (c. 3,000–4,000 km) distances in the species (Figure 1), differ markedly in migration direction, body size, and plumage (Buehler & Piersma, 2008; Piersma, 2011), and spend most of the year in opposite hemispheres. Conversely, we found clear structure between *roselaari* W and E, which have no obvious barriers to gene flow, as they are morphologically similar (Tomkovich, 1992) and share a flyway, directly meeting at both passage and wintering sites (Carmona et al., 2013) and migrating in essentially the same direction (Figure 1). Further understanding of the spatial organization of neutral genetic variation in red knots and other migratory birds will require quantification of the interacting contributions of history, geography, ecology, and behaviour to the formation of corridors and barriers to gene flow. This would be best accomplished in a comparative framework including multiple globally-distributed species, and with explicit consideration of the life stage (i.e., naïve first-time migrants vs. experienced adults) during which dispersal is most likely to occur.

## ACKNOWLEDGEMENTS

This work is dedicated to the memory of Allan Baker, who built its conceptual and practical foundations. We thank Per Palsbøll and Ritsert Jansen for guidance and help to develop the funding proposal, and to members of the Palsbøll and Fontaine laboratories at the University of Groningen for helpful discussions during project conception and analysis, respectively. We thank Anneke Bol at NIOZ Royal Netherlands Institute for Sea Research for DNA extractions of Mauritania samples; Marco van der Velde for laboratory assistance in Groningen; Eric Johnson and Paul Etter (SNPsaurus) for sequencing, bioinformatics, and technical expertise; and the Centre for Information Technology of the University of Groningen (particularly Bob Dröge and Cristian Marocico) for technical support and access to the Peregrine high-performance computing cluster. For providing samples, we thank Massey University, Royal Ontario Museum (Oliver Haddrath and Mark Peck), Moscow Lomonosov State University Zoological Museum, NIOZ, U.S. Fish & Wildlife Service (especially Lucas DeCicco and Nicholas Hajdukovich), and University of Groningen. Australia sampling was made possible by the support of BirdLife Netherlands at the initiative of its director, the late Adri de Gelder, during the early years of Global Flyway Network. This project was supported by a Dutch Research Council (NWO) grant to TP (ALW-Open Programme grant, "Ecological drivers of global flyway evolution" 824.01.001).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

J.R.C, Y.I.V, M.C.F and T.P conceived and designed the study. P.F.B, C.J.H, J.t.H, James J.A.M, and P.S.T organized and performed field sampling and maintained individual resight-history databases. J.R.C and Y.I.V curated samples and conducted the labwork. J.R.C and M.C.F analysed and interpreted the data, with assistance from Y.I.V. J.R.C and M.C.F wrote the manuscript, with major contributions from Y.I.V and T.P. All authors provided feedback and edited the manuscript.

## OPEN RESEARCH BADGES



This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at 10.5061/dryad.j3tx95xgb.

## DATA AVAILABILITY STATEMENT

Demultiplexed nextRAD short-read data with sample metadata have been deposited in NCBI's SRA archives under BioProject ID PRJNA799587 (accession numbers SAMN25275563–SAMN25275765). Associated files (VCF and metadata) have been deposited in DRYAD (<https://doi.org/10.5061/dryad.j3tx95xgb>).

## ORCID

Jesse R. Conklin <https://orcid.org/0000-0002-1414-0587>

Yvonne I. Verkuil <https://orcid.org/0000-0001-7080-3026>

Phil F. Battley <https://orcid.org/0000-0002-8590-8098>

James A. Johnson <https://orcid.org/0000-0002-2312-0633>

Pavel S. Tomkovich <https://orcid.org/0000-0002-1563-2196>

Theunis Piersma <https://orcid.org/0000-0001-9668-466X>

Michaël C. Fontaine <https://orcid.org/0000-0003-1156-4154>

## REFERENCES

- Able, K. P., & Belthoff, J. R. (1998). Rapid "evolution" of migratory behaviour in the introduced house finch of eastern North America. *Proceedings of the Royal Society B: Biological Sciences*, 265, 2063–2071. <https://doi.org/10.1098/rspb.1998.0541>
- Åkesson, S., Ilieva, M., Karagicheva, J., Rakhimberdiev, E., Tomotani, B., & Helm, B. (2017). Timing avian long-distance migration: From internal clock mechanisms to global flights. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1734), <https://doi.org/10.1098/rstb.2016.0252>
- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12, 246. <https://doi.org/10.1186/1471-2105-12-246>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109.vidual>
- Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). strataG: An R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, 17(1), 5–11. <https://doi.org/10.1111/1755-0998.12559>
- Arcones, A., Ponti, R., Ferrer, X., & Vieites, D. R. (2020). Pleistocene glacial cycles as drivers of allopatric differentiation in Arctic shorebirds. *Journal of Biogeography*, 48(4), 747–759. <https://doi.org/10.1111/jbi.14023>

- Atkinson, P. W., Baker, A. J., Bevan, R. M., Clark, N. A., Cole, K. B., Gonzalez, P. M., Newton, J., Niles, L. J., & Robinson, R. A. (2005). Unravelling the migration and moult strategies of a long-distance migrant using stable isotopes: Red Knot *Calidris canutus* movements in the Americas. *Ibis*, 147(4), 738–749. <https://doi.org/10.1111/j.1474-919x.2005.00455.x>
- Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography—the mitochondrial-DNA bridge between population genetics and systematics. *Annual Review of Ecology, Evolution, and Systematics*, 40, 593–612. <https://doi.org/10.1146/annurev.es.18.110187.002421>
- Avise, J. C., & Walker, D. (1998). Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society B: Biological Sciences*, 265, 457–463. <https://doi.org/10.1098/rspb.1998.0317>
- Baker, A. J., González, P. M., Piersma, T., Niles, L. J., de Lima Serrano do Nascimento, I., Atkinson, P. W., Clark, N. A., Minton, C. D. T., Peck, M. K., & Aarts, G. (2004). Rapid population decline in red knots: Fitness consequences of decreased refuelling rates and late arrival in Delaware Bay. *Proceedings of the Royal Society B: Biological Sciences*, 271(1541), 875–882. <https://doi.org/10.1098/rspb.2003.2663>
- Batchelor, C. L., Margold, M., Krapp, M., Murton, D. K., Dalton, A. S., Gibbard, P. L., Stokes, C. R., Murton, J. B., & Manica, A. (2019). The configuration of Northern Hemisphere ice sheets through the Quaternary. *Nature Communications*, 10(1), 3713. <https://doi.org/10.1038/s41467-019-11601-2>
- Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4), 2025–2035. <https://doi.org/10.1093/genetics/162.4.2025>
- Berthold, P., & Helbig, A. J. (1992). The genetics of bird migration: stimulus, timing, and direction. *Ibis*, 134, 35–40. <https://doi.org/10.1111/j.1474-919x.1992.tb04731.x>
- Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid micro-evolution of migratory behaviour in a wild bird species. *Nature*, 360, 668–670. <https://doi.org/10.1038/360668a0>
- Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: Some cons, many pros. *Molecular Ecology*, 19(13), 2609–2625. <https://doi.org/10.1111/j.1365-294x.2010.04690.x>
- Boyd, H., & Piersma, T. (2001). Changing balance between survival and recruitment explains population trends in Red Knots *Calidris canutus islandica* wintering in Britain, 1969–1995. *Ardea*, 89(2), 301–317.
- Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, 135(3), 439–455. <https://doi.org/10.1007/s10709-008-9293-3>
- Buehler, D. M., & Baker, A. J. (2005). Population divergence times and historical demography in Red Knots and Dunlins. *Condor*, 107, 497–513. <https://doi.org/10.1093/condor/107.3.497>
- Buehler, D. M., Baker, A. J., & Piersma, T. (2006). Reconstructing palaeoflyways of the late Pleistocene and early Holocene Red Knot *Calidris canutus*. *Ardea*, 94(3), 485–498.
- Buehler, D. M., & Piersma, T. (2008). Travelling on a budget: predictions and ecological evidence for bottlenecks in the annual cycle of long-distance migrants. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1490), 247–266. <https://doi.org/10.1098/rstb.2007.2138>
- Bushnell, B. (2016). *BBMap*. Retrieved from <http://sourceforge.net/projects/bbmap>
- Carmona, R., Arce, N., Ayala-Perez, V., Buchanan, J. B., Salzer, L. J., Tomkovich, P. S., & Newstead, D. (2013). Red Knot *Calidris canutus roselaari* migration connectivity, abundance and non-breeding distribution along the Pacific coast of the Americas. *Wader Study Group Bulletin*, 120(3), 168–180.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), 1–16. <https://doi.org/10.1186/s13742-015-0047-8>
- Conklin, J. R., Senner, N. R., Battley, P. F., & Piersma, T. (2017). Extreme migration and the individual quality spectrum. *Journal of Avian Biology*, 48(1), 19–36. <https://doi.org/10.1111/jav.01316>
- Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., & Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30(8), 1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>
- 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., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., & Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *Elife*, 9, 1–24. <https://doi.org/10.7554/eLife.54462>
- Dick, W. J., Piersma, T., & Prokosch, P. (1987). Spring migration of the Siberian Knots *Calidris canutus canutus*: results of a co-operative Wader Study Group project. *Ornis Scandinavica*, 18, 5–16.
- Dyke, A. S. (2004). An outline of North American deglaciation with emphasis on central and northern Canada. *Developments in Quaternary Science*, 2(Part, B), 373–424. [https://doi.org/10.1016/S1571-0866\(04\)80209-4](https://doi.org/10.1016/S1571-0866(04)80209-4)
- Ehlers, J., & Gibbard, P. L. (2007). The extent and chronology of Cenozoic Global Glaciation. *Quaternary International*, 164–165, 6–20. <https://doi.org/10.1016/j.quaint.2006.10.008>
- Estoup, A., Lombaert, E., Marin, J.-M., Guillemaud, T., Pudlo, P., Robert, C. P., & Cornuet, J.-M. (2012). Estimation of demographic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources*, 12(5), 846–855. <https://doi.org/10.1111/j.1755-0998.2012.03153.x>
- Fedorov, V. B., & Stenseth, N. C. (2002). Multiple glacial refugia in the North American Arctic: inference from phylogeography of the colored lemming (*Dicrostonyx groenlandicus*). *Proceedings of the Royal Society B: Biological Sciences*, 269(1505), 2071–2077. <https://doi.org/10.1098/rspb.2002.2126>
- Fitak, R. R. (2019). *optM: an R package to optimize the number of migration edges using threshold models*. Retrieved from: <https://github.com/cran/OptM>
- Gherardi-Fuentes, C., Ruiz, J. O. R. G. E., & Navedo, J. G. (2021). Insights into migratory connectivity and conservation concerns of Red Knots *Calidris canutus* in the austral Pacific coast of the Americas. *Bird Conservation International*. doi: <https://doi.org/10.1017/S0959270921000095>
- Gualtieri, L., Glushkova, O., & Brigham-Grette, J. (2000). Evidence for restricted ice extent during the last glacial maximum in the Koryak Mountains of Chukotka, far eastern Russia. *Bulletin of the Geological Society of America*, 112(7), 1106–1118.
- Hassell, C., Southey, I., Boyle, A., & Yang, H.-Y. (2011). Red Knot *Calidris canutus*: subspecies and migration in the East Asian-Australasian flyway—where do all the Red Knot go? *BirdingASIA*, 16, 89–93.
- Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913. <https://doi.org/10.1038/35016000>
- Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F., & Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years

- after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54(1), 291–301. <https://doi.org/10.1016/j.ympev.2009.09.016>
- Johnson, N. K., & Cicero, C. (2004). New mitochondrial DNA data affirm the importance of pleistocene speciation in North American birds. *Evolution*, 58(5), 1122–1130. <https://doi.org/10.1111/j.0014-3820.2004.tb00445.x>
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>
- Karagicheva, J., Rakhimberdiev, E., Dekinga, A., Brugge, M., Koolhaas, A., ten Horn, J., & Piersma, T. (2016). Seasonal time keeping in a long-distance migrating shorebird. *Journal of Biological Rhythms*, 31(5), 509–521. <https://doi.org/10.1177/0748730416655929>
- Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). DiveRsim: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782–788. <https://doi.org/10.1111/2041-210X.12067>
- Klicka, J., & Zink, R. M. (1997). The importance of recent ice ages in speciation: A failed paradigm. *Science*, 277(5332), 1666–1669. <https://doi.org/10.1126/science.277.5332.1666>
- Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, 40, 593–612. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095702>
- Kok, E. M. A., Tibbitts, T. L., Douglas, D. C., Howey, P. W., Dekinga, A., Gnep, B., & Piersma, T. (2020). A red knot as a black swan: how a single bird shows navigational abilities during repeat crossings of the Greenland Icecap. *Journal of Avian Biology*, 51(8), 1–11. <https://doi.org/10.1111/jav.02464>
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5), 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Kraaijeveld, K., & Nieboer, E. N. (2000). Late Quaternary paleogeography and evolution of arctic breeding waders. *Ardea*, 88(2), 193–205.
- Lappo, E. G., Tomkovich, P. S., & Syroechkovskiy, E. E. (2012). *Atlas of breeding waders in the Russian Arctic*. UF Ofsetnaya Pechat.
- Lathrop, R. G., Niles, L., Smith, P., Peck, M., Dey, A., Sacatelli, R., & Bogner, J. (2018). Mapping and modeling the breeding habitat of the Western Atlantic Red Knot (*Calidris canutus rufa*) at local and regional scales. *Condor*, 120(3), 650–665. <https://doi.org/10.1650/CONDOR-17-247.1>
- 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, 3258. <https://doi.org/10.1038/s41467-018-05257-7>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Manley, W. F. (2002). *Postglacial flooding of the Bering Land Bridge: A geospatial animation*. INSTAAR, University of Colorado, v1. Retrieved from [http://instaar.colorado.edu/groups/QGISL/bering\\_land\\_bridge/](http://instaar.colorado.edu/groups/QGISL/bering_land_bridge/) Accessed June 14, 2021
- McLaughlin, J. F., Faircloth, B. C., Glenn, T. C., & Winker, K. (2020). Divergence, gene flow, and speciation in eight lineages of trans-Beringian birds. *Molecular Ecology*, 29(18), 3526–3542. <https://doi.org/10.1111/mec.15574>
- Méndez, V., Alves, J. A., Gill, J. A., & Gunnarsson, T. G. (2018). Patterns and processes in shorebird survival rates: a global review. *Ibis*, 160(4), 723–741. <https://doi.org/10.1111/ibi.12586>
- Milanesi, M., Capomaccio, S., Vajana, E., Bomba, L., Garcia, J. F., Ajmone-Marsan, P., & Colli, L. (2017). BITE: an R package for biodiversity analyses. *BioRxiv*, 181610. <https://doi.org/10.1101/181610>
- Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2014). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22(11), 2841–2847. <https://doi.org/10.1111/mec.12350> Genotyping-by-sequencing
- Nebel, S., Piersma, T., van Gils, J. A., Dekinga, A., & Spaans, B. (2000). Length of stopover, fuel storage and a sex-bias in the occurrence of Red Knots *Calidris c. canutus* and *C.c. islandica* in the Wadden Sea during southward migration. *Ardea*, 88(2), 165–176.
- Nei, M. (1972). Genetic distance between populations. *American Naturalist*, 106(949), 283–292. <https://doi.org/10.1086/282771>
- Newstead, D. J., Niles, L. J., Porter, R. R., Dey, A. D., Burger, J., & Fitzsimmons, O. N. (2013). Geolocation reveals mid-continent migratory routes and Texas wintering areas of Red Knots *Calidris canutus rufa*. *Wader Study Group Bulletin*, 120, 53–59.
- Novembre, J. (2016). Pritchard, Stephens, and Donnelly on population structure. *Genetics*, 204(2), 391–393. <https://doi.org/10.1534/genetics.116.195164>
- 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>
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528. <https://doi.org/10.1093/bioinformatics/bty633>
- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2(12), 2074–2093. <https://doi.org/10.1371/journal.pgen.0020190>
- Pérez-Tris, J., Bensch, S., Carbonell, R., Helbig, A. J., & Tellería, J. L. (2004). Historical diversification of migration patterns in a passerine bird. *Evolution*, 58(8), 1819–1832. <https://doi.org/10.1554/03-731>
- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8(11), e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
- Pielou, E. C. (1991). *After the ice age: The return of life to glaciated North America*. University of Chicago Press.
- Piersma, T. (2007). Using the power of comparison to explain habitat use and migration strategies of shorebirds worldwide. *Journal of Ornithology*, 148(S1), 45–59. <https://doi.org/10.1007/s10336-007-0240-3>
- Piersma, T. (2011). Flyway evolution is too fast to be explained by the modern synthesis: proposals for an 'extended' evolutionary research agenda. *Journal of Ornithology*, 152 (suppl), S151–S159. <https://doi.org/10.1007/s10336-011-0716-z>
- Piersma, T., & Davidson, N. (1992). The migration of Knots. *Wader Study Group Bulletin*, 64(suppl. 1), 1–209.
- Piersma, T., Kok, E. M. A., Hassell, C. J., Peng, H., Verkuil, Y. I., Lei, G., & Chan, Y. (2021). When a typical jumper skips: itineraries and staging habitats used by Red Knots (*Calidris canutus piersmai*) migrating between northwest Australia and the New Siberian Islands. *Ibis*, <https://doi.org/10.1111/ibi.12964>
- Piersma, T., Lok, T., Chen, Y., Hassell, C. J., Yang, H.-Y., Boyle, A., Slaymaker, M., Chan, Y.-C., Melville, D. S., Zhang, Z.-W., & Ma, Z. (2016). Simultaneous declines in summer survival of three shorebird species signals a flyway at risk. *Journal of Applied Ecology*, 53(2), 479–490. <https://doi.org/10.1111/1365-2664.12582>
- Piersma, T., Pérez-Tris, J., Mouritsen, H., Bauchinger, U., & Bairlein, F. (2005). Is there a “migratory syndrome” common to all migrant birds? *Annals of the New York Academy of Sciences*, 1046, 282–293. <https://doi.org/10.1196/annals.1343.026>
- Piersma, T., Rogers, D. I., González, P. M., Zwarts, L., Niles, L. J., de Donascimento, I., ... Baker, A. J. (2005). Fuel storage rates before northward flights in Red Knots worldwide: Facing the severest ecological constraint in tropical intertidal environments? In R. Greenberg, & P. P. Marra (Eds.), *Birds of two worlds: the ecology and evolution of migratory birds* (pp. 262–274). Johns Hopkins University Press.

- Pritchard, J. K., Wen, X., & Falush, D. (2010). *Documentation for structure software: Version 2.3* (pp. 1–39). Retrieved from [http://pritch.bsd.uchicago.edu/structure\\_software/release\\_versions/v2.3.4/html/structure.html](http://pritch.bsd.uchicago.edu/structure_software/release_versions/v2.3.4/html/structure.html)
- Provan, J., & Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23(10), 564–571. <https://doi.org/10.1016/j.tree.2008.06.010>
- Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866. <https://doi.org/10.1093/bioinformatics/btv684>
- Rakhimberdiev, E., van den Hout, P. J., Brugge, M., Spaans, B., & Piersma, T. (2015). Seasonal mortality and sequential density dependence in a migratory bird. *Journal of Avian Biology*, 46(4), 332–341. <https://doi.org/10.1111/jav.00701>
- Raynal, L., Marin, J. M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2019). ABC random forests for Bayesian parameter inference. *Bioinformatics*, 35(10), 1720–1728. <https://doi.org/10.1093/bioinformatics/bty867>
- Richardson, D. S., Jury, F. L., 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(9), 2263–2273. <https://doi.org/10.1046/j.0962-1083.2001.01355.x>
- Rodríguez-Ramilo, S. T., & Wang, J. (2012). The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. *Molecular Ecology Resources*, 12(5), 873–884. <https://doi.org/10.1111/j.1755-0998.2012.03156.x>
- Rogers, D. I., Yang, H.-Y., Hassell, C. J., Boyle, A. N., Rogers, K. G., Chen, B., & Piersma, T. (2010). Red Knots (*Calidris canutus piersma* and *C. c. rogersi*) depend on a small threatened staging area in Bohai Bay, China. *Emu*, 110, 307–315.
- Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, 3, e1106. <https://doi.org/10.7717/peerj.1106>
- Stewart, J. R., & Dalén, L. (2008). Is the glacial refugium concept relevant for northern species? A comment on Pruett and Winker 2005. *Climatic Change*, 86, 19–22. <https://doi.org/10.1007/s10584-007-9366-9>
- Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: Individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, 277(1682), 661–671. <https://doi.org/10.1098/rspb.2009.1272>
- Studds, C. E., Kendall, B. E., Murray, N. J., Wilson, H. B., Rogers, D. I., Clemens, R. S., Gosbell, K., Hassell, C. J., Jessop, R., Melville, D. S., Milton, D. A., Minton, C. D. T., Possingham, H. P., Riegen, A. C., Straw, P., Woehler, E. J., & Fuller, R. A. (2017). Rapid population decline in migratory shorebirds relying on Yellow Sea tidal mudflats as stopover sites. *Nature Communications*, 8, 14895. <https://doi.org/10.1038/ncomms14895>
- Tomkovich, P. S. (1992). An analysis of the geographic variability in Knots *Calidris canutus* based on museum skins. *Wader Study Group Bulletin*, 64(suppl), 17–23.
- Tomkovich, P. S. (2001). A new subspecies of red knot *Calidris canutus* from the New Siberian islands. *Bulletin of the British Ornithologists' Club*, 121(4), 257–263.
- Tomkovich, P. S., Porter, R. R., Loktionov, E. Y., & Niles, L. J. (2013). Pathways and staging areas of Red Knots *Calidris canutus rogersi* breeding in southern Chukotka, Far Eastern Russia. *Wader Study Group Bulletin*, 120(3), 181–193.
- Tomkovich, P. S., & Riegen, A. C. (2000). Mixing of red knot populations in Australia: some thoughts. *Stilt*, 37, 25–27.
- van Gils, J. A., Lisovski, S., Lok, T., Meissner, W., Ożarowska, A., de Fouw, J., Rakhimberdiev, E., Soloviev, M. Y., Piersma, T., & Klaassen, M. (2016). Body shrinkage due to Arctic warming reduces red knot fitness in tropical wintering range. *Science*, 352(6287), 819–822. <https://doi.org/10.1126/science.aad6351>
- van Roomen, M., Nagy, S., Foppen, R., Dodman, T., Citegetse, G., & Ndiaye, A. (2015). *Status of coastal waterbird populations in the East Atlantic Flyway 2014: With special attention to flyway populations making use of the Wadden Sea*. Programme Rich Wadden Sea.
- Verhoeven, M. A., van Eerbeek, J., Hassell, C. J., & Piersma, T. (2016). Fuelling and moult in Red Knots before northward departure: A visual evaluation of differences between ages, sexes and subspecies. *Emu*, 116(2), 158–167. <https://doi.org/10.1071/MU15035>
- Verkuil, Y. I., Tavares, E., González, P. M., Choffe, K., Haddrath, O., Peck, M., & Conklin, J. R. (2021). Genetic structure in the nonbreeding range of rufa Red Knots suggests distinct Arctic breeding populations. *Ornithological Applications*, 124, duab053.
- Wauchope, H. S., Shaw, J. D., Varpe, Ø., Lappo, E. G., Boertmann, D., Lanctot, R. B., & Fuller, R. A. (2017). Rapid climate-driven loss of breeding habitat for Arctic migratory birds. *Global Change Biology*, 23(3), 1085–1094. <https://doi.org/10.1111/gcb.13404>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370. <https://doi.org/10.2307/2409936>
- Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society B: Biological Sciences*, 271(1551), 1881–1887. <https://doi.org/10.1098/rspb.2004.2803>
- Wetlands International (2021). *Waterbird Population Estimates*. Retrieved from [wpe.wetlands.org](http://wpe.wetlands.org) Accessed January 20, 2021
- Wilson, J., Aubry, Y., Buidin, C., Rochepault, Y., & Baker, A. J. (2010). Three records of Red Knots *Calidris canutus* possibly changing flyways. *Wader Study Group Bulletin*, 117(3), 192–193.
- Winker, K. (2010). On the origin of species through heteropatric differentiation: A review and a model of speciation in migratory animals. *Ornithological Monographs*, 69, 1–30. <https://doi.org/10.1525/om.2010.69.1.1.1>
- Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., & Wang, J. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, 346(6215), 1311–1321.
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>
- Zöckler, C., & O'Sullivan, J. (2005). New Zealand Red Knot breeding in Meinopylgino, Chukotka, NE Russia. *Wader Study Group Bulletin*, 108, 76.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Conklin, J. R., Verkuil, Y. I., Battley, P. F., Hassell, C. J., ten Horn, J., Johnson, J. A., Tomkovich, P. S., Baker, A. J., Piersma, T., & Fontaine, M. C. (2022). Global flyway evolution in red knots *Calidris canutus* and genetic evidence for a Nearctic refugium. *Molecular Ecology*, 31, 2124–2139. <https://doi.org/10.1111/mec.16379>