



**HAL**  
open science

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

Jesse R Conklin, Yvonne I Verkuil, Philip Battley, Chris J Hassell, Job Ten Horn, James A Johnson, Pavel S Tomkovich, Allan J Baker, Theunis Piersma, Michael C. Fontaine

### ► To cite this version:

Jesse R Conklin, Yvonne I Verkuil, Philip Battley, Chris J Hassell, Job Ten Horn, et al.. Global flyway evolution in red knots *Calidris canutus* and genetic evidence for a Nearctic refugium. 2021. hal-03408091v1

**HAL Id: hal-03408091**

**<https://hal.science/hal-03408091v1>**

Preprint submitted on 28 Oct 2021 (v1), last revised 9 Aug 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

3

4 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>,  
5 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  
6 C. Fontaine<sup>1,8,9</sup>

7

8 <sup>1</sup>Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen,  
9 P.O. Box 11103, 9700 CC Groningen, The Netherlands

10 <sup>2</sup>Wildlife and Ecology Group, School of Agriculture and Environment, Massey University,  
11 Palmerston North 4442, New Zealand

12 <sup>3</sup>Global Flyway Network, PO Box 3089, Broome, WA 6725, Australia

13 <sup>4</sup>Department of Coastal Systems, NIOZ Royal Netherlands Institute for Sea Research, P.O.  
14 Box 59, 1790 AB Den Burg, Texel, The Netherlands

15 <sup>5</sup>U.S. Fish & Wildlife Service, Migratory Bird Management, 1011 E. Tudor Road, MS 201,  
16 Anchorage, Alaska 99503, USA

17 <sup>6</sup>Zoological Museum, Moscow MV Lomonosov State University, Bolshaya Nikitskaya Str.  
18 6, Moscow 125009, Russia

19 <sup>7</sup>Department of Natural History, Royal Ontario Museum, 100 Queens Park, Toronto, ON  
20 M5S 2C6, Canada

21 <sup>8</sup>MIVEGEC, Univ. Montpellier, CNRS, IRD, Montpellier, France

22 <sup>9</sup>Centre de Recherche en Écologie et Évolution de la Santé (CREES), Montpellier, France

23 \*Corresponding author: conklin.jesse@gmail.com

24 †Deceased

25

26 **ABSTRACT**

27

28 Present-day ecology and population structure are the legacies of past climate and habitat  
29 perturbations, and this is particularly true for species that are widely distributed at high  
30 latitudes. The red knot, *Calidris canutus*, is an arctic-breeding, long-distance migratory  
31 shorebird with six recognized subspecies defined by differences in morphology, migration  
32 behavior, and annual-cycle phenology, in a global distribution thought to have arisen just  
33 since the Last Glacial Maximum (LGM). We used nextRAD sequencing of 10,881 single-  
34 nucleotide polymorphisms (SNPs) to assess the neutral genetic structure and  
35 phylogeographic history of 172 red knots representing all known global breeding  
36 populations. Using population genetics approaches, including model-based scenario-testing  
37 in an approximate Bayesian computation (ABC) framework, we infer that red knots derive  
38 from two main lineages that diverged ca. 34,000 years ago, and thus persisted at the LGM  
39 in both Palearctic and Nearctic refugia, followed by at least two instances of secondary  
40 contact and admixture. In two flyways, we detected clear genetic structure between  
41 population pairs with similar migrations and substantial geographic overlap in the non-  
42 breeding season. Conversely, other populations were only weakly differentiated despite  
43 clearly divergent migratory phenotypes and little or no apparent contact throughout the  
44 annual cycle. In general, the magnitude of genetic differentiation did not match that of  
45 phenotypic differences among populations, suggesting that flyway-specific phenotypes  
46 developed quite rapidly and do not necessarily impose barriers to gene flow. Our results  
47 suggest that population structure and migratory phenotypes in red knots arose from a  
48 complex interplay among phylogeography, plasticity, and selective processes.

49

## 50 1 | INTRODUCTION

51

52 The ecology and demography of species are typically viewed through the lens of present-  
53 day or very recent (i.e. decades) observations and processes. However, it is clear that

54 habitats, distributions, genetic diversity, and populations themselves are the legacies of  
55 events and conditions in historical or evolutionary time-scales, and cannot be fully

56 understood without a phylogeographic perspective (Avice et al., 1987; Knowles, 2009).

57 This is particularly true for high-latitude species, whose habitats were repeatedly

58 transformed by Pleistocene glacial cycles, and whose current distributions must have arisen

59 just since the Last Glacial Maximum (LGM, ca. 20,000 years before present, ybp; Hewitt

60 2000). Of course, the effects of glaciations on historical distributions, and on population

61 genetic diversity and structure, are strongly dependent on a species' ecological and

62 physiological attributes (Stewart, Lister, Barnes, & Dalén, 2010).

63

64 Migratory birds are characterized by seasonal movements and high dispersal capabilities,

65 with important consequences for population structure, the strength of geographical barriers,

66 and the propensity to colonize novel ranges (Winker, 2010). The effects of glacial

67 perturbations on the geographic organization of migratory bird species and subspecies are

68 profound, although the precise consequences are varied and debated (Avice & Walker,

69 1998; Johnson & Cicero, 2004; Klicka & Zink, 1997; Weir & Schluter, 2004). For example,

70 it remains unclear whether tundra-specialized populations were most restricted and isolated

71 (i.e. in 'refugia') during the LGM, when vast areas of the Arctic were under ice, or during

72 the mid-Holocene Climatic Optimum (ca. 8,000 ybp), when a period of elevated

73 temperatures limited the extent of tundra habitat (Arcones, Ponti, Ferrer, & Vieites, 2020;

74 Stewart & Dalén, 2008; Wauchope et al., 2017). Regardless of the mechanisms, the present-

75 day global distributions and ‘flyways’ (i.e. networks of routes and sites used throughout the  
76 annual journey) of many arctic-breeding migratory bird species attest to broad post-glacial  
77 colonization (Kraaijeveld & Nieboer, 2000) and a general lability of migratory behavior  
78 (Piersma, 2011).

79

80 Understanding the intertwining effects of historical and present-day processes on migratory  
81 systems is best approached through population genetic and phylogeographic approaches  
82 based on genome-wide markers, and informed by detailed ecological knowledge (Orsini,  
83 Vanoverbeke, Swillen, Mergeay, & De Meester, 2013). Among the most ecologically well-  
84 studied migratory birds is the red knot, *Calidris canutus*, a globally-distributed shorebird  
85 (order Charadriiformes, family Scolopacidae) that breeds on high-latitude (65–80°N) arctic  
86 tundra and uses temperate and tropical intertidal mudflats during the rest of the year  
87 (Piersma, 2007; Piersma & Davidson, 1992). There are six recognized subspecies of red  
88 knot (Figure 1), distinguished by their breeding and non-breeding ranges, and their  
89 migratory behavior, with one-way distances ranging from <3,000 to >14,000 km (Piersma,  
90 2007; Piersma, Rogers, et al., 2005) and including some of the longest non-stop flights  
91 recorded in birds (Conklin, Senner, Battley, & Piersma, 2017). These populations differ in  
92 body size, plumage, and degree of spatial and behavioral overlap with neighboring  
93 subspecies throughout the annual cycle (Buehler & Piersma, 2008; Tomkovich, 1992,  
94 2001). Census population sizes vary by more than an order of magnitude (<20,000 to  
95 >400,000 individuals; Wetlands International, 2021), but most have faced significant  
96 declines in recent decades (Baker et al., 2004; Boyd & Piersma, 2001; Studds et al., 2017;  
97 van Gils et al., 2016), largely due to anthropogenic impacts on non-breeding habitats (Baker  
98 et al., 2004; Piersma et al., 2016; Rakhimberdiev, van den Hout, Brugge, Spaans, &  
99 Piersma, 2015). Therefore, understanding flexibility and magnitude of genetic,

100 reproductive, and ecological separation among these populations is of significant  
101 conservation and evolutionary interest.

102

103 Shallow genetic differentiation and signals of population expansion in the mitochondrial  
104 control region (mtDNA) suggested a very recent origin of present-day red knot populations  
105 (Buehler & Baker, 2005), and perhaps a global radiation from a single LGM refugium  
106 (Buehler, Baker, & Piersma, 2006). Notably, some subspecies pairs that were  
107 indistinguishable in mtDNA (Buehler & Baker, 2005) have clear phenotypic differences in  
108 morphology and migration distance, timing, and direction. In red knots, migratory behavior  
109 is further associated with seemingly ‘hard-wired’ endogenous annual rhythms in molt and  
110 mass, apparently adapted to flyway-specific conditions (Karagicheva et al., 2016; Piersma,  
111 2011). This suggests that complex, multi-trait migratory ‘syndromes’ can arise very quickly,  
112 even in the presence of gene flow (Delmore et al., 2020; Pérez-Tris, Bensch, Carbonell,  
113 Helbig, & Tellería, 2004). However, unraveling such recent, and potentially reticulated,  
114 evolutionary histories requires genome-wide information (Brito & Edwards, 2009; Narum,  
115 Buerkle, Davey, Miller, & Hohenlohe, 2014) coupled with powerful model-based scenario-  
116 testing such as approximate Bayesian computation (ABC) (Beaumont, 2010; Bertorelle,  
117 Benazzo, & Mona, 2010; Hickerson et al., 2010).

118

119 Here, we revisit genetic population structure and phylogeography of red knots, to describe  
120 the history of divergences and degree of neutral genetic differentiation among global flyway  
121 populations. For this, we exploit recent research on red knot ecology and migration, which  
122 has clarified temporal and geographic overlap among subspecies (e.g. Atkinson et al., 2005;  
123 Carmona et al., 2013; Nebel, Piersma, van Gils, Dekinga, & Spaans, 2000; Verhoeven, van  
124 Eerbeek, Hassell, & Piersma, 2016) and has made available more comprehensive global

125 sampling associated with known migratory phenotypes and ecologies. We used nextRAD  
126 (nextera-tagmented, Reductively Amplified DNA) sequencing (Russello, Waterhouse, Etter,  
127 & Johnson, 2015) for *de novo* discovery of genome-wide single-nucleotide polymorphisms  
128 (SNPs) for population genetic analyses, and compared hypothesized evolutionary scenarios  
129 in an ABC framework using DIYABC (Cornuet et al., 2014). By reconstructing the recent  
130 evolutionary history of red knots, we: (1) revise our understanding of LGM refugia and the  
131 colonization of global flyways, (2) reveal previously unrecognized population structure, and  
132 (3) provide new insights regarding the flexibility and isolating function of migratory  
133 behavior.

134

135

## 136 **2 | MATERIALS AND METHODS**

137

### 138 **2.1 | Sampling and DNA extraction**

139 We assembled DNA samples representing all recognized and hypothesized breeding  
140 populations within the global range of red knots (Figure 1, Table S1). Where possible, we  
141 used samples collected in known breeding areas; for *C. c. roselaari* (hereafter, we refer to  
142 populations by only their sub-specific epithets) these included samples from two disjunct  
143 breeding areas (Wrangel Island, Russia and Seward Peninsula, Alaska, USA) and we treated  
144 these as separate groups (*roselaari* West (W) and East (E), respectively), to test whether  
145 *roselaari* should be considered one or two independent demographic units. Because red  
146 knots breed in low densities in remote areas of arctic tundra, for some populations there  
147 were few or no breeding samples available. In these cases, we used samples collected from  
148 non-breeding areas (i.e. sites used during migration and/or the boreal winter) when sampled  
149 individuals could be confidently assigned to breeding areas either because they were

150 remotely tracked to breeding areas using light-level geolocation, or because long-term  
151 research programs had established strong links between breeding and non-breeding areas  
152 (e.g. through mark-recapture/resight programs). All tissue samples were acquired from  
153 museum collections or collected by the authors and colleagues in the field under all  
154 requisite permits appropriate to their respective countries and institutions.

155

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

167

## 168 **2.2 | SNP genotyping using nextRAD sequencing**

169 Genomic DNA was converted into nextRAD genomic fragment libraries (SNPsaurus, LLC,  
170 USA) following the method described by Russello et al. (2015). For each sample, 20–30 ng  
171 of genomic DNA was first fragmented with Nextera reagent (Illumina, Inc., USA), which  
172 also ligates short adapter sequences to the ends of the fragments. Fragmented DNA was  
173 then amplified, with primers matching the adapter and one primer extending 10 nucleotides  
174 into the genomic DNA with the selective sequence 'GTGTAGAGCC'. Thus, only



175 fragments starting with a sequence that can be hybridized by the selective sequence of the  
176 primer were efficiently amplified. PCR amplification was done at 74°C for 27 cycles. The  
177 nextRAD libraries were sequenced on an Illumina HiSeq-4000 at the Genomics Core  
178 Facility, University of Oregon, USA.

179

180 For SNP-calling, reads were first trimmed using custom scripts (SNPsaurus, LLC, USA) in  
181 bbdduk (BBMap tools; Bushnell, 2016). Next, a *de novo* reference was created from  
182 abundant reads (after removal of low-quality (phred-scale quality <20) and very high-  
183 abundance reads) and reads that aligned to these. All 161,810,193 reads were mapped to the  
184 reference with an alignment identity threshold of 95% using bbmap (BBMap tools).

185 Genotype calling was performed using SAMtools and BCFtools (samtools mpileup -gu -Q  
186 10 -t DP, DPR -f ref.fasta -b samples.txt | bcftools call -cv - > genotypes.vcf), applying a  
187 minimum read depth filter of 7x. The genotype table was then filtered using VCFtools  
188 v.0.1.14 (Danecek et al., 2011) to remove SNPs called in <80% of samples and putative  
189 alleles with a population frequency <3%, to exclude artefactual variants. The resulting VCF  
190 file included 4,911 unique loci (150-bp sequences) containing 14,903 SNPs. Additional  
191 cleaning was performed using VCFtools to remove indels (n = 452 SNPs) and samples that  
192 failed to sequence (n = 13 individuals). Next, to exclude potential genotyping errors, SNPs  
193 deviating from Hardy-Weinberg proportions in at least 6 of 7 hypothesized populations  
194 were identified and removed using VCFtools ( $P < 0.05$ ; n = 58 SNPs). To ensure  
195 independence of loci, we used PLINK v.1.9 (Chang et al., 2015) to identify and remove all  
196 SNPs in linkage disequilibrium (LD;  $r^2 > 0.20$ ; n = 3,512 SNPs). The LD-pruned dataset  
197 included 192 individuals and 10,881 unlinked SNPs on 4,679 loci.

198

199 We then used VCFtools to calculate proportion of missing SNP calls per individual (range  
200 1–87%) and removed 12 individuals with >25% missing data. Because inclusion of related  
201 individuals can bias population genetic analyses (Rodríguez-Ramilo & Wang, 2012), we  
202 estimated individual pairwise relatedness in PLINK using the Identity by Descent estimator  
203 PI HAT. Removal of 8 individuals (6 *roselaari* E, 2 *rufa*) resolved all cases of relatedness  
204 involving half-siblings or closer (PI HAT >0.20). The final dataset included 172 individuals  
205 and 10,881 SNPs. We used VCFtools, PLINK, and PGDspider v.2.1.0.3 (Lischer &  
206 Excoffier, 2012) to convert data to different formats required for analysis.

207

### 208 **2.3 | Inference of population structure and diversity**

209 To assess major axes of genetic variation and identify distinct genetic clusters, we  
210 performed a principal component analysis (PCA; Patterson et al. 2006) using the R  
211 packages *gdsfmt* v.1.14.1 and *SNPRelate* v.1.12.2 (Zheng et al., 2012). We determined the  
212 number of ancestral populations ( $K$ ) and ancestry proportions of each individual using  
213 ADMIXTURE v.1.3.0 (Alexander, Novembre, & Lange, 2009) with cross-validation (CV)  
214 (Alexander & Lange, 2011) to perform 10 replicate runs (with random seeds) for each  
215 putative number of clusters ( $K$ ) ranging from 1 to 8. We assessed how the CV error rate  
216 varied with increasing  $K$  and which value provided the lowest CV error. Then, we used the  
217 CLUMPAK (Cluster Markov Packager Across K; Kopelman et al. 2015) web server  
218 (<http://clumpak.tau.ac.il/>) with default settings to summarize estimates of individual  
219 ancestry proportions to each cluster across replicate runs, inspect potential distinct solutions  
220 identified, and visualize the most likely ancestry proportions at each value of  $K$ .

221

222 Globally and for each putative population, we characterized genetic diversity by calculating  
223 nucleotide diversity ( $\pi$ ), heterozygosity, and inbreeding coefficient ( $F_{IS}$ ) using VCFtools.

224 We evaluated evidence for deviations from mutation-drift equilibrium, indicative of  
225 departures from a null hypothesis of constant population size, by estimating the per-locus  
226 Tajima's  $D$  values for each population using VCFtools. We estimated genetic differentiation  
227 among populations by calculating pairwise  $F_{ST}$  (Weir & Cockerham, 1984) using the  
228 *diffCalc* function in the R package *diveRsity* v.1.9.90 (Keenan, McGinnity, Cross, Crozier,  
229 & Prodöhl, 2013), with 95% confidence intervals derived from 500 bootstraps.

230

## 231 **2.4 | Population evolutionary relationships and demographic history**

232 To visualize evolutionary relationships among populations, we first constructed a midpoint-  
233 rooted neighbor-joining tree based on Nei's genetic distance (Nei, 1972) using the R  
234 package *poppr* v.2.8.5 (Kamvar, Tabima, & Grünwald, 2014), with missing data replaced  
235 by mean allele counts, and node support calculated from 1,000 bootstraps.

236

237 For a probabilistic assessment of population relationships, we further investigated  
238 population branching order using the Bayesian coalescent-based approach of the program  
239 SNAPP (Bryant, Bouckaert, Felsenstein, Rosenberg, & Roychoudhury, 2012) implemented  
240 in BEAST v.2.6.1 (Bouckaert et al., 2014). For computational tractability, we reduced each  
241 population to the six individuals with highest total SNP coverage; this yielded a dataset with  
242 1,573 SNPs that contained no missing data. We performed 50 million Monte Carlo Markov  
243 Chain (MCMC) iterations, sampling parameters and trees every 1,000 iterations, after  
244 discarding the first 5% iterations as burn-in. To evaluate convergence, we used Tracer  
245 v.1.7.1 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018) to inspect posterior  
246 distributions and ensure an effective sample size greater than 4,000 for all parameters. We  
247 visualized the 50,000 resulting trees using DensiTree v.2.2.7 (Bouckaert, 2010) and plotted  
248 the maximum clade credibility (MCC) tree using FigTree v.1.4.3 (Rambaut, 2010). We

249 converted divergence time estimates to years assuming a generation time of 6 years  
250 (delayed maturity with adult annual survival ca. 0.80; Méndez, Alves, Gill, & Gunnarsson,  
251 2018) and the genome-wide mutation rate calculated by Zhang et al. (2014) for  
252 Charadriiformes:  $1.5 \times 10^{-9}$  substitutions per site per year.

253

254 To complement the coalescent-based SNAPP analysis, which does not account for possible  
255 gene flow among populations, we also used TreeMix v.1.13 (Pickrell & Pritchard, 2012) to  
256 estimate patterns of population splits and mixtures. Using genome-wide allele frequency  
257 data, TreeMix estimates how sampled populations are related to their common ancestor  
258 through a graph of ancestral populations. A maximum-likelihood tree of populations is  
259 estimated with the nodes and branch lengths representing the amount genetic variance (or  
260 drift) shared among populations and within each one, respectively. Migration edges are  
261 added in a stepwise manner among populations, minimizing the genetic covariance  
262 unexplained by the tree. To simplify this analysis, we excluded the *rogersi* 2 cluster (see  
263 *Results*), resulting in a dataset of 163 individuals and 10,881 SNPs. We performed 10  
264 replicate TreeMix runs for each value of  $m$  (migration edges) from 1 to 10, and evaluated  
265 the change in log-likelihood as  $m$  was increased ( $\Delta m$ ) using the R package *OptM* v.0.1.3  
266 (Fitak, 2019) to determine an optimal value of  $m = 4$ . We then performed 1,000 bootstrap  
267 replicates in TreeMix, using  $m = 4$  and blocks of 500 SNPs to derived a consensus  
268 population tree topology with node supports using the R package *BITE* v.1.2.0008 (Milanesi  
269 et al., 2017). A final run was conducted using the consensus tree and the optimized number  
270 of migration edges ( $m = 4$ ) using the full data set and the options ‘-m 4 -tf consensus.tree -  
271 se’, to generate the final population graph with bootstrapped node supports as described in  
272 (Milanesi et al., 2017).

273

274 Finally, we evaluated the support for different possible scenarios of population divergence  
275 and admixture using the Approximate Bayesian Computation (ABC) (Beaumont, Zhang, &  
276 Balding, 2002) Random Forest (RF) statistical framework (Pudlo et al., 2016; Raynal et al.,  
277 2019). ABC-RF can estimate posterior probabilities of historical scenarios, based on  
278 massive coalescent simulations of genetic data. Simulations are compared to observed data  
279 using summary statistics to identify the best-fitting model by calculating the number of RF  
280 votes and to derive the posterior probability for the best model. The best-fitting posterior  
281 parameter distribution values for the best model can be estimated using a RF procedure  
282 applied in a regression setting (Raynal et al., 2019). Estimated parameters include the  
283 effective size ( $N_e$ ) for each population, admixture events and their rates ( $r$ ), and split and  
284 admixture times ( $t$ ). We reduced the dataset to all 4,126 SNPs genotyped in at least one  
285 individual per population. We then evaluated potential evolutionary scenarios in a stepwise  
286 manner, as follows.

287

288 In Step 1, we compared four scenarios (Figure S1, Table S2), aimed at resolving the  
289 backbone population topology, before further exploration of evolutionary scenarios. Each of  
290 these four scenarios included two admixture events. The first scenario (Figure S1a) was  
291 based on the similar topology inferred by the NJ tree and SNAPP, and included admixed  
292 origins of *canutus* and *rogersi*, as indicated by the PCA and ADMIXTURE analyses and the  
293 uncertainty in SNAPP topologies (see *Results*). The other three scenarios (Figure S1b–d)  
294 represent three possible rootings of the unrooted topology inferred by TreeMix (see  
295 *Results*), including admixture events representing the first two inferred migration edges (i.e.  
296 admixed origin of *piersmai* and *rufa/islandica*).

297

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

307

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

318

319 We used the random forest (RF) classification procedure to compare the likelihood of the  
320 competing scenarios at each step with the R package *abcrf* v.1.8.1 (Pudlo et al., 2016). RF is  
321 a machine-learning algorithm that uses hundreds of bootstrapped decision trees to perform  
322 classification, using the summary statistics as a set of predictor variables. Some simulations

323 are not used in decision tree building at each bootstrap (i.e., the out-of-bag simulations), and  
324 are used to compute the “prior error rate,” which provides a direct method for estimating the  
325 cross-validation error rate. At each step, we built a training set of 20,000 simulated datasets  
326 per scenario, with the same number of loci and individuals as the observed dataset. We then  
327 grew a classification forest of 1,000 and 1,500 trees respectively for Step 1 and Step 2. The  
328 RF computation provides a classification vote for each scenario (i.e., the number of times a  
329 model is selected from the decision trees). We selected the scenario with the highest  
330 classification vote as the most likely scenario, and we estimated its posterior probability  
331 following the recommendation of Pudlo et al. (2016). We assessed the global performance  
332 of our chosen ABC-RF scenario, by calculating the prior error rate based on the available  
333 out-of-bag simulations and we repeated the RF analysis 10 times to ensure that the results  
334 converged.

335

336 Then, posterior distribution values of all parameters for the best model identified were  
337 estimated using a regression by RF methodology (Raynal et al., 2019), with classification  
338 forests of 1,000 decision trees, and based on a training set of 100,000 simulations. We  
339 converted estimates for timing parameters to years assuming a generation time of 6 years  
340 (as above). The simulation steps, the computation of summary statistics, and the model  
341 checking analysis were performed in DIYABC v.2.1.0. All scenario comparisons and the  
342 estimations of parameter posterior distribution values were carried out with the R package  
343 *abcrf* v.1.8.1 (Pudlo et al., 2016; Raynal et al., 2019).

344

345

346

## 347 3 | RESULTS

348

### 349 3.1 | Summary of nextRAD SNP data set

350 The final dataset comprised 172 unrelated individuals, including 7–31 individuals in each of  
351 the seven hypothesized populations (Table S1), genotyped at 10,881 unlinked high-quality  
352 SNPs. On average, individuals were genotyped at 95.5% (range: 76–99%) of SNPs and with  
353 a mean read depth of 57.3 (range: 17–137). Each SNP was genotyped in an average of  
354 164.3 (range: 149–172, or 86.6–100%) individuals. Globally, nucleotide diversity ( $\pi$ ) was  
355 0.219, and observed heterozygosity was 18.7% (Figure S5, Table S5).

356

### 357 3.2 | Population structure and diversity

358 The first three axes (PC1–3) of the principal component analysis explained 4.1% of the total  
359 genetic variation. PC1 and PC2 distributed individuals into a triangular pattern (Figures 2  
360 and S3), suggesting three major genetic pools: a Canadian Arctic group of *rufal/islandica*, a  
361 central Palearctic group of *canutus/piersmai*, and a group formed by *roselaari* E from  
362 Alaska. Individuals from *rogersi* and *roselaari* W clustered in intermediate positions  
363 between the *canutus/piersmai* and *roselaari* E clusters, suggesting these populations may  
364 represent admixed groups between the two extreme clusters. Unexpectedly, PC3 separated  
365 nine of 13 Chukotka-breeding individuals from the main *rogersi* cluster (Figures 2 and S3).  
366 Higher-order axes of variation (PC5–6) revealed a clear separation of *canutus* and *piersmai*  
367 (Figure S3); however, no clear separation of *rufa* and *islandica* was apparent in the top 8 PC  
368 axes.

369

370 Genetic ancestry analyses with ADMIXTURE for  $K = 1$  or 2 ancestral populations provided  
371 the lowest cross-validation error criterion, suggesting the best fit at these values (Figure S4).



372 However, the major clusters identified in the PCA were all clearly visible at  $K = 3$  and 4  
373 (Figure 3), and the clusters at these  $K$  values had 100% support among 10 replicate runs.  
374 Greater values of  $K$  were also informative with regard to further sub-structuring according to  
375 geography or *a priori* hypotheses of subspecies structure (see the major clusters for  $K = 2-8$   
376 in Figure 3). Consistent with the PCA, individual ancestry estimated at  $K = 3$  using  
377 ADMIXTURE identified the three major genetic clusters identified in the PCA:  
378 *rufa/islandica*, *canutus/piersmai*, and *roselaari* E, with *roselaari* W and *rogersi* displaying  
379 significant amounts of admixture. *Roselaari* W shared more common ancestry with *roselaari*  
380 E than with *canutus/piersmai*, whereas *rogersi* displayed a more balanced shared ancestry  
381 between the two groups. These results are consistent with the position of *roselaari* W and  
382 *rogersi* in the PCA relative to *roselaari* E and *canutus/piersmai* (Figures 2 and S3).  
383 Interestingly, one purported *islandica* individual, sampled in the *islandica* breeding range,  
384 was identified as a possible F1 *islandica/canutus* hybrid at  $K = 2-5$  (see also Figure S3). The  
385 nine Chukotka-breeding knots identified as distinct within the *rogersi* group on the third axis  
386 of the PCA (Figure 2) were also recognized by the ADMIXTURE analysis as a distinct  
387 genetic cluster at  $K \geq 4$ ; this cluster only weakly occurred in New Zealand *rogersi* samples.  $K$   
388 = 5 identified a cluster composed mostly of Palearctic Beringian populations (*piersmai*,  
389 *rogersi*, *roselaari* W), but absent in Alaska (*roselaari* E). Greater values of  $K$  were poorly  
390 supported and largely uninformative, but demonstrate the relatively weak differentiation of  
391 *canutus* and *piersmai*, and the even weaker structure between *rufa* and *islandica*.

392

393 Both PCA and ADMIXTURE strongly suggest that five individuals were mistakenly  
394 assigned to population based on *a priori* hypotheses: two purported *islandica* sampled  
395 during the winter in the Wadden Sea are clearly *canutus*, and three purported *rogersi* from  
396 New Zealand appear to be *piersmai* (Figures 2, 3, and S3). Therefore, we re-assigned these

397 five individuals to their ‘correct’ populations for all subsequent analyses. Also, for some  
398 further analyses, we consider the two *rogersi* clusters separately, with *rogersi* 2 comprising  
399 the nine Chukotka individuals differentiated in PC3, and *rogersi* 1 comprising the  
400 remaining 16 *rogersi* individuals. Aside from these cases, negligible differences between  
401 sampling locations presumed to represent the same population confirmed our *a priori*  
402 assumptions about population distributions (within *rufa*, *islandica*, *canutus*, and *piersmai*;  
403 see Figure 3, Table S7).

404

405 Globally, nucleotide diversity ( $\pi$ ) was 0.219, and this was similar for all populations, except  
406 for the slightly lower 0.184 in *rogersi* 2 (Figure S5a, Table S5). Mean heterozygosity was  
407 also generally uniform, but slightly higher and more variable in *rogersi* 1 and *roselaari* W  
408 (Figure S5b). We found some evidence of inbreeding (mean  $F_{IS} > 0.22$ ) in *rogersi* 2 and  
409 *roselaari* W (Figure S5c). Values of Tajima’s  $D$  (range of means: 0.006–0.154) were  
410 indistinguishable from zero for all populations (Figure S5d), suggesting no evidence of  
411 dramatic population-size changes.

412

413 Pairwise  $F_{ST}$  among the eight hypothesized populations ranged 0.005–0.058 (Table 1). The  
414 two lowest values involved pairs of previously recognized subspecies: *rufa* vs. *islandica*  
415 ( $F_{ST} = 0.005$  with 95% CI including zero) and *canutus* vs. *piersmai* ( $F_{ST} = 0.007$ ). For all  
416 populations, the greatest pairwise differences were with *rogersi* 2 ( $F_{ST} = 0.036$ –0.058),  
417 including within purported *rogersi* (*rogersi* 1 vs. *rogersi* 2:  $F_{ST} = 0.036$ ). By contrast, other  
418 pairwise differences with *rogersi* 1 were relatively low (all  $\leq 0.021$ ). Comparisons with  
419 *roselaari* W had relatively large  $F_{ST}$  estimates (0.020–0.058) but several had 95% CIs  
420 approaching or including zero, likely an effect of the small sample size for *roselaari* W ( $n =$   
421 7).

422

### 423 3.3 | Population evolutionary relationships and demographic history

424 The neighbor-joining (NJ) tree based on Nei's distance (Figure 4a) identified the sister-  
425 population pairs of *roselaari* W/E, *canutus/piersmai*, and *rufa/islandica*, and a Beringian  
426 group including *rogersi* and *roselaari*. Furthermore, it confirmed that the two clusters in  
427 *rogersi* are distinct but closely-related. The coalescent approach of SNAPP produced a  
428 maximum credibility clade (MCC) tree with essentially the same topology as the NJ tree,  
429 with two major branches diverging ca. 17,000 ybp: a Beringian clade of *rogersi/roselaari*  
430 and a second clade containing the sister-pairs of *piersma/canutus* and *rufa/islandica* (Figure  
431 4b). However, the MCC tree shows low support for nodes involving the populations  
432 *canutus*, *piersmai*, and *rogersi*, with discordant gene trees clearly illustrated by DensiTree,  
433 reflecting the effects of admixture/gene flow leading to the uncertain relationships among  
434 these three populations (Figure 4b). SNAPP, which does not account for admixture or gene  
435 flow, was thus unable to resolve clearly whether *canutus* or *rogersi* is the closest relative of  
436 *piersmai*.

437

438 The complementary approach of TreeMix provided the best fit for the data with a  
439 population tree topology that included four migration edges. This is indicated by the change  
440 of proportion of explained genetic variance reaching a plateau at 99.8% and by the change  
441 in log likelihood with increasing number of migration edges (Figure S6). With zero  
442 migration edges (not shown), TreeMix inferred the same topology recovered by the NJ tree  
443 (Figure 4a) and SNAPP (Figure 4b). The branching relationships among *canutus*, *piersmai*,  
444 and *rogersi* 1 (*rogersi* 2 excluded here) changed with the number of inferred migration  
445 edges, reflecting the uncertainty in this part of the topology (Figure S7). With four  
446 migration edges (Figure 4c), the first inferred edge was from *canutus* to *piersmai*, followed

447 by a second from the *roselaari* branch to *rufa/islandica*. The third edge reflected admixture  
448 between *canutus* and *islandica*, and in the fourth, between *rogersi* and *roselaari* E. Note the  
449 low bootstrap supports ( $\leq 50\%$ ) at the node between *piersmai* and *rogersi* 1 and the node  
450 between *canutus* and the ancestor of *rufa* and *islandica*, reflecting the impact of admixture  
451 and gene flow.

452

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

462

463 Step 2 in the ABC-RF built upon the best-supported scenario in Step 1 by comparing nine  
464 scenarios: six consisting of variations on scenario *d* from Step 1, each including additional  
465 admixture events as suggested by TreeMix, and three scenarios testing other plausible  
466 biogeographic hypotheses (Figure S2). The six variations from scenario *d* in Step 1  
467 (scenarios *a* to *f* in Step 2; Figure S2) were closely related to each other and received  
468 collectively 79% of the RF votes. In contrast, the alternative scenarios (*g* to *i*; Figure S2)  
469 were least supported, each receiving  $\leq 9\%$  of the votes.

470

471 Among Step 2 scenarios, scenario *d* (Figure 6) was identified as the best-supported scenario  
472 by the ABC-RF, with the greatest number of votes (17%), a posterior probability of 43.9%,  
473 and a prior error rate of 31.4% (Figure S2). This scenario included an older admixed origin  
474 of *piersmai* and an additional recent admixed origin of *islandica*. This reflects strong  
475 support for the third, but not the fourth, inferred migration edge from TreeMix. However,  
476 the second best-supported scenario (*f*, with 15% of votes; Figure S2), included the fourth  
477 TreeMix migration edge, providing weak support for recent admixture between *rogersi* and  
478 *roselaari*.

479

480 Each of the 25 demographic parameters of the best-supported scenario was estimated within  
481 the ABC-RF framework (Table S6). Timing parameter estimates strongly support a pre-  
482 LGM divergence of the Canadian Arctic group from the Palearctic/Beringian group (ca.  
483 34K ybp; Figure 6, Table S6), followed by post-glacial divergences and admixture in both  
484 major branches. In particular, we infer recent divergences within *roselaari* and *rogersi* (ca.  
485 6K and 3K ybp, respectively), and two instances of secondary contact between the Nearctic  
486 and Palearctic groups (via *roselaari* ca. 7K ybp and via *canutus* ca. 3K ybp; Figure 6).  
487 Estimated effective population sizes of the eight extant populations (N1–8 in Table S6)  
488 ranged from 4,231 to 50,616 and captured some expected differences among populations.  
489 For example, the three lowest *Ne* estimates were for a highly endangered population (*rufa*)  
490 and two presumably small populations with signals of inbreeding and lower than expected  
491 heterozygosity (*rogersi* 2, *roselaari* W; Figure S5b–c). Conversely, two admixed  
492 populations (*islandica*, *piersmai*) produced the highest *Ne* estimates. Estimated admixture  
493 rates were also consistent with results of ADMIXTURE and TreeMix: *islandica* derived  
494 67% of its ancestry from *rufa* and 33% from *canutus* (r1 in Table S6), whereas *piersmai*  
495 derived 33% from *canutus* and 67% from *rogersi* (r2).

#### 496 4 | DISCUSSION

497

498 In one circumpolar distribution of an arctic-breeding long-distance migratory bird, we have  
499 demonstrated numerous patterns of population structure according to migratory flyways and  
500 recognized subspecies distinctions. In one case, we discovered clear genetic structure  
501 between two breeding populations currently considered the same subspecies, and sharing a  
502 migratory route and wintering areas (*roselaari* W vs. E). At the other extreme, we found  
503 essentially no genetic differentiation between two recognized subspecies with vast  
504 differences in migratory route and distance, which spend the non-breeding season on  
505 different continents (*rufa* vs. *islandica*). Intermediate to these are: (1) detectable structure  
506 between populations with substantial ecological overlap during the non-breeding season,  
507 similar migrations but geographically distant breeding areas (*piersmai* vs. *rogersi*); and (2)  
508 apparently closely-related populations with no apparent contact during the annual cycle  
509 (*canutus* vs. *piersmai*). Lastly, we discovered curious structure within a subspecies, perhaps  
510 suggesting demographic consequences of a disjunct and recently-founded breeding area  
511 (*rogersi* 1 vs. 2). These varied cases show that there is no simple relationship between  
512 genotypic and phenotypic (i.e. migratory) structure in red knots, and that differences in  
513 migration ecology do not quickly, or necessarily, impose reproductive isolation.

514

515 Our reconstruction of their evolutionary history showed that red knots persisted at the LGM  
516 in at least two refugia, revising our understanding of the origin of their current global  
517 distribution. In addition, we identified at least three cases of historical admixture, including  
518 two instances of deeply diverged populations experiencing gene flow after recent secondary  
519 contact (*roselaari*–*rufa* and *islandica*–*canutus*). Together, uncertainty in tree topology and  
520 widespread signals of admixture imply ongoing divergence with unsorted lineages, and

521 profound responses to historical environmental changes. Below, we discuss the implications  
522 of our results for understanding: (1) the impacts of post-LGM climate perturbations, (2)  
523 present-day population structure and recognized subspecies, and (3) the flexibility and  
524 isolating function of migratory phenotypes.

525

#### 526 **4.1 | Post-glacial phylogeography of red knots**

527 Previous population-genetic work in red knots (Buehler & Baker, 2005) suggested that red  
528 knots were restricted to the Palearctic during the LGM, and only colonized the Canadian  
529 Arctic in the last few thousand years, after the Holocene Climatic Optimum (Buehler,  
530 Baker, & Piersma, 2006). This conclusion was largely based on extremely low  
531 differentiation at the mtDNA control region among the three Nearctic subspecies (Buehler  
532 & Baker, 2005). However, the apparent close relationship of *roselaari* to *rufa* and *islandica*  
533 was an artefact of sampling – in fact, there were likely no *roselaari* in that study, as the  
534 purported *roselaari* samples came from the southeastern USA, now considered an exclusive  
535 *rufa* wintering area (Atkinson et al., 2005; Carmona et al., 2013). Using verifiable breeding  
536 samples, we identified *roselaari* and *rufa* as the most differentiated neighboring subspecies  
537 pair, making a scenario of eastward colonization of the Nearctic (Buehler, Baker, &  
538 Piersma, 2006) highly unlikely. This leaves two plausible scenarios: (1) after the LGM,  
539 *canutus* spread from Europe to colonize the Nearctic breeding range of *islandica*, and then  
540 gave rise to *rufa* through a southward expansion from Ellesmere Island; or (2) the  
541 *rufa/islandica* clade arose *in situ* from a Nearctic refugium. The former scenario is possibly  
542 consistent with  $F_{ST}$  estimates (Table 1) and the topologies inferred with Nei's distance  
543 (Figure 4a) and SNAPP (Figure 4b), all of which suggest a closer relationship of  
544 *rufa/islandica* with the western Palearctic (*canutus/piersmai*) than with the Beringian group  
545 (*rogersi/roselaari*). However, the explicit consideration of admixture (Figures 3 and 4c)

546 supports the latter scenario, in which a deeply-diverged Nearctic branch came into recent  
547 secondary contact with both western Palearctic and Beringian groups. Our DIYABC  
548 analysis confirmed this, indicating a Nearctic ancestor diverging from the ancestor of the  
549 Palearctic/Beringian clade ca. 34,000 ybp, and then experiencing admixture from both the  
550 west and east in the last few thousand years (Figure 6). We note that our SNAPP analysis  
551 estimated the divergence of the two main branches at ca. 17,000 ypb (Figure 4b), which is  
552 similar to the deepest divergence estimate of 20,000 ypb from Buehler & Baker (2005) and  
553 consistent with an entirely post-LGM expansion. However, neither of these estimates  
554 consider the effect of gene flow along the tree; when this is included (as we did in  
555 DIYABC), the deepest divergence-time estimate is roughly doubled (Figure 6), a scenario  
556 requiring that red knots survived the LGM in both Palearctic and Nearctic refugia.

557

558 It is well established that a huge swath of the Arctic including most of Beringia and  
559 stretching eastward almost to the Taimyr Peninsula in central Russia was largely ice-free at  
560 the LGM (Ehlers & Gibbard, 2007; Pielou, 1991), and this refugium probably gave rise to  
561 all Palearctic and Beringian red knot subspecies. Meanwhile, the Nearctic was  
562 predominantly glaciated at the LGM, including nearly all of the present-day breeding ranges  
563 of *islandica* and *rufa*. However, there are multiple lines of evidence supporting potential  
564 refugia in this region (Dyke, 2004; Fedorov & Stenseth, 2002; Provan & Bennett, 2008),  
565 including on Banks Island near the extreme western extent of the *rufa* range (Figure 1), but  
566 also in northeast Greenland within the *islandica* range. Predictive modeling of habitats at  
567 the LGM suggested that Banks Island had the greatest potential suitability for tundra-  
568 breeding shorebirds in the Canadian Arctic, whereas ice-free areas at higher latitudes were  
569 less-suitable polar desert (Arcones, Ponti, Ferrer, & Vieites, 2020). Much of the present-day  
570 high-Arctic breeding range of *islandica* was not ice-free until ca. 6,000 ybp (Dyke, 2004).



571 Our DIYABC analysis inferred that *islandica* arose ca. 3,200 ybp, after Nearctic red knots  
572 had established contact with *roselaari* ca. 6,400 ybp (Figure 6), perhaps through non-  
573 breeding overlap in temperate North America. Thus, the most likely scenario is that red  
574 knots persisted at the LGM in a refugium on or near Banks Island, and then expanded  
575 northeastward as glaciers retreated, colonizing the present-day *islandica* range. After this, a  
576 new migration route to Europe was established, promoting contact and admixture with  
577 *canutus* in the last 3,000–4,000 years (Figure 6).

578

579 Red knots breeding near Banks Island at the LGM and migrating southward to temperate or  
580 tropical non-breeding areas would be required to migrate >2,500 km over the Laurentide Ice  
581 Sheet, which covered much of central North America (Dyke, 2004; Ehlers & Gibbard,  
582 2007). The present-day migrations of all red knot subspecies involve non-stop flights of  
583 4,000–10,000 km (Conklin, Senner, Battley, & Piersma, 2017; Kok et al., 2020). In  
584 particular, post-breeding *islandica* individuals can fly 4,000 km from the Canadian Arctic to  
585 western Europe, including up to 1,700 km across the Greenland Ice Sheet (Kok et al., 2020).  
586 Therefore, we find a migration from Banks Island entirely plausible.

587

588 The historical migration of the westernmost Palearctic red knots (i.e. present-day *canutus*)  
589 may have similarly involved flights across the ice sheets that covered northwestern Europe  
590 (Batchelor et al., 2019), which could explain their contemporary non-stop flights between  
591 Taimyr and the Wadden Sea coast (Figure 1), both of which were largely ice-free at the  
592 LGM (Ehlers & Gibbard, 2007). However, it is also possible that *canutus* colonized this  
593 flyway more recently, after ice sheets retreated from Europe (Patton et al., 2017), and that  
594 this westward expansion facilitated a post-LGM divergence from other Palearctic red knots  
595 (Figure 6).

596

597 As recognized by Buehler et al. (2006), the post-LGM diversification of  
598 Palearctic/Beringian red knots into four recognized subspecies (Figure 6) is consistent with  
599 a scenario of increasingly isolated patches of tundra breeding habitat in the warming period  
600 prior to the Holocene Climatic Optimum (Arcones, Ponti, Ferrer, & Vieites, 2020;  
601 Kraaijeveld & Nieboer, 2000; Stewart & Dalén, 2008; Wauchope et al., 2017). However,  
602 the inferred admixed origin of *piersmai* (Figure 6) suggests the additional possibility that  
603 the breeding population on the New Siberian Islands (Figure 1) was founded later by  
604 colonization from two previously isolated mainland populations (*canutus* and *rogersi*).  
605 Interestingly, unlike much of the present-day *rogersi* breeding range, the region of our  
606 Chukotka sampling site was likely glaciated at the LGM (Ehlers & Gibbard, 2007),  
607 although the precise extent is disputed (Gualtieri, Glushkova, & Brigham-Grette, 2000),  
608 suggesting that the unexpected genetic cluster we detected (*rogersi* 2) could reflect a recent  
609 colonization of this disjunct breeding area (Figure 1). The warming period of retreating  
610 glaciers also featured rising sea levels which inundated the Bering Land Bridge, gradually  
611 isolating Alaska from Chukotka (and thus *roselaari* E from *rogersi*) by ca. 12,000 ybp, and  
612 eventually Wrangel Island from mainland Russia (i.e. *roselaari* W from *rogersi*), ca. 10,000  
613 ybp (Dyke, 2004; Manley, 2002). Clearly, the population structure and uncertainty in tree  
614 topology we found in red knots of the eastern Palearctic and particularly Beringia, attests to  
615 a recent history of climate and habitat upheavals in the region (McLaughlin, Faircloth,  
616 Glenn, & Winker, 2020). Expanded and more targeted genetic sampling may reveal further  
617 insights into this complicated regional history.

618

619 The scenario of recent global expansion from the Palearctic proposed by Buehler et al.  
620 (2006) was partly based on mtDNA evidence for dramatic population growth after a

621 demographic bottleneck, including both Tajima's  $D$  and mismatch analysis (Buehler &  
622 Baker, 2005). By contrast, we found no non-zero mean values of  $D$  (Figure S5d) and  
623 therefore no support for recent population-size changes globally or in any subspecies. We  
624 filtered out alleles with a frequency <3%, which may have slightly dampened our Tajima's  
625  $D$  estimates. However, even the most extreme single-population value reported by Buehler  
626 & Baker (2005),  $D = -1.72$  in *rogersi*, falls within the wide range of estimates across  
627 individual SNPs in our study (Figure S5d), demonstrating how single-locus approaches can  
628 provide misleading characterizations of genome-wide patterns (Brito & Edwards, 2009).

629

#### 630 **4.2 | Further insights for global population structure**

631 We confirm the extremely close relationship between the Nearctic subspecies *rufa* and  
632 *islandica*, which were effectively indistinguishable in mtDNA (Buehler & Baker, 2005).  
633 With genome-wide SNPs, we found similarly weak differentiation, which may reflect a very  
634 recent northeastward expansion (see above), but also raises the question of whether the two  
635 populations are demographically independent at present. Currently, the extent and potential  
636 overlap of the breeding ranges of *rufa* and *islandica* are poorly described (Lathrop et al.,  
637 2018), and mark-resight studies have recorded instances of apparently 'switching' of  
638 individual knots between the two flyways (Wilson, Aubry, Buidin, Rochepault, & Baker,  
639 2010). We sampled the two subspecies at the geographic extremes of the Canadian breeding  
640 range (Figure 1), and thus intermediate sampling might reveal more clinal variation by  
641 latitude across the ranges of *rufa* and *islandica*. Moreover, indications of genetic structure  
642 within the non-breeding range of *rufa* have been detected in amplified fragment length  
643 polymorphism (AFLP) markers (Verkuil et al., in revision), raising the possibility of further  
644 longitudinal structure across the low Canadian Arctic. The *rufa* individuals in our study  
645 formed a relatively homogeneous group (Figures 2 and 3), but we note that our samples

646 came from easterly breeding and passage sites and therefore may miss potential within-  
647 subspecies differences, particularly with *rufa* that migrate through central North America  
648 (Figure 1; Newstead et al., 2013). This may also exaggerate the observed genetic  
649 differentiation between *rufa* and *roselaari* E, which are known to share non-breeding areas  
650 in the Gulf of Mexico (Figure 1; (Carmona et al., 2013; Newstead et al., 2013). Elucidating  
651 this recent and perhaps still-developing Nearctic structure requires a greater breadth of  
652 sampling across breeding areas in the Canadian Arctic and Greenland, which may be best  
653 accomplished through tracking the migrations of knots captured at non-breeding sites.

654

655 Red knots breeding in Alaska and on Wrangel Island have been considered the same  
656 subspecies, *roselaari*, due to morphological similarity and a common migration route along  
657 the west coast of North America (Carmona et al., 2013; Tomkovich, 1992). Although we  
658 confirmed their close relationship by descent (Figure 4), we observed genetic differentiation  
659 at a magnitude similar to some geographically distant subspecies-level comparisons (Table  
660 1). Marked individuals from both breeding areas have been detected as far south as Sonora  
661 and Baja California, Mexico (Carmona et al., 2013), suggesting that *roselaari* W and E may  
662 largely overlap during the non-breeding season, and that population estimates and  
663 conservation efforts therefore conflate two potentially independent demographic units.  
664 Given the distinct clustering of the two populations, large-scale molecular population  
665 assignment of individuals from the *roselaari* non-breeding range may be possible, to  
666 elucidate the population sizes and degree of year-round ecological and geographic overlap  
667 of Wrangel Island and Alaskan knots. Also, broader geographic sampling in Alaska may  
668 help describe whether structure within the currently-accepted *roselaari* is as sharply defined  
669 as our results suggest, or perhaps more clinal.

670

671 The apparent structure we detected within *rogersi* was quite unexpected and warrants more  
672 investigation. Our expectation was that non-breeding *rogersi* samples from New Zealand  
673 and breeding samples from Meinypilgyno, in southeast Chukotka, would equally represent  
674 the population that breeds widely across (predominantly northern) Chukotka. This  
675 assumption was based on demonstrated links between Meinypilgyno and New Zealand  
676 through color-band resighting and geolocator-tracking (Tomkovich, Porter, Loktionov, &  
677 Niles, 2013; Zöckler & O’Sullivan, 2005). However, in both PCA and ADMIXTURE  
678 analysis, only 4 of 13 Meinypilgyno individuals grouped with the New Zealand birds, while  
679 the remaining 9 formed a clearly distinct cluster (Figures 2 and 3), and pairwise  $F_{ST}$   
680 between the two sampling sites argued against a homogeneous population (Table S7).  
681 Meinypilgyno is located in a small, disjunct part of the breeding range, separated by >200  
682 km from the closest known *rogersi* breeding areas (Figure 1; Lappo, Tomkovich, &  
683 Syroechkovskiy, 2012) and likely colonized after retreat of glaciers in the region (see  
684 above). In this light, the modest signals of inbreeding and elevated expected heterozygosity  
685 we found in *rogersi* 2 (Figure S5) could reflect a recent founder effect. There is some  
686 independent evidence from geolocator-tracking of associated migratory differences: two  
687 knots tracked from Meinypilgyno made >10,000 km non-stop northward flights from New  
688 Zealand to the Yellow Sea (Tomkovich, Porter, Loktionov, & Niles, 2013), whereas seven  
689 individuals tagged in New Zealand all made *en route* stops in northeastern Australia (P.F.  
690 Battley, unpubl.). Further genetic sampling and migration tracking of knots from the main  
691 *rogersi* breeding range are needed to understand this intriguing structure.

692

693 Due to the lack of sufficient breeding samples for all populations, we used non-breeding  
694 samples when current understanding allowed confident assignment to subspecies. These  
695 assumptions were generally confirmed (Table S7), but two exceptions highlighted by PCA

696 and ADMIXTURE analyses provide insight about unrecognized subspecies overlap in the  
697 non-breeding season. First, it is accepted that red knots performing post-breeding molt and  
698 spending the winter in Europe are *islandica*, because *canutus* individuals only pass through  
699 this region *en route* to (boreal) wintering sites in western Africa (Dick, Piersma, &  
700 Prokosch, 1987). However, we identified two apparent *canutus* individuals among knots  
701 sampled in the Dutch Wadden Sea in mid-winter (14% of samples). This implies that the  
702 wintering range of *canutus* extends into western Europe, and the period of annual contact  
703 between the subspecies is greater than previously understood. This has important  
704 implications for estimating populations sizes and trends based on distinct wintering areas  
705 (van Roomen et al., 2015), and provides a simpler potential pathway for gene flow, as  
706 indicated by TreeMix and the one apparent F1 hybrid (*canutus* x *islandica*) sampled at the  
707 *islandica* breeding area on Ellesmere Island (see Figure 3).

708

709 Similarly, we identified three apparent *piersmai* in our sample of purported *rogersi* from  
710 New Zealand. These subspecies overlap clinally during the boreal winter, such that sites in  
711 western Australia contain >80% *piersmai*, whereas sites in New Zealand are ~80% *rogersi*  
712 (Piersma et al., 2021; Tomkovich & Riegen, 2000; Verhoeven, van Eerbeek, Hassell, &  
713 Piersma, 2016). Plumage differences are considered sufficient to distinguish the subspecies  
714 (Hassell, Southey, Boyle, & Yang, 2011; Tomkovich, 2001), and we individually verified  
715 the samples of *piersmai* from Broome, Australia, based on plumage recorded during  
716 resights of color-marked individuals in Bohai Bay, China (Rogers et al., 2010). This was  
717 entirely successful, as we found no evidence of subspecies mixing in our Broome samples.  
718 We had also correctly identified three *piersmai* from our New Zealand site, two based on  
719 plumage observed before northward departure and one based on a geolocator track (Table  
720 S1). However, PCA and ADMIXTURE identified three more apparent *piersmai*, which we

721 had initially included in our sample of purported *rogersi* (Figures 2 and 3). This  
722 demonstrates that not all *piersmai* individuals attain a distinguishable plumage prior to  
723 northward migration, and can thus be mistaken for the duller-plumaged *rogersi* subspecies  
724 at this time of year.

725

#### 726 **4.3 | Implications for the flexibility and isolating function of migratory phenotypes**

727 In migratory birds, fitness is dependent on a multi-trait phenotype (including flyway-specific  
728 physiological adaptations for flight, fueling, navigation, molt, etc.; Piersma, Pérez-Tris,  
729 Mouritsen, Bauchinger, & Bairlein, 2005) that ensures well-timed reproduction in seasonal  
730 conditions (Åkesson et al., 2017). Some components, such as migration direction and  
731 circannual rhythms, are to some extent heritable and endogenously entrained (Berthold &  
732 Helbig, 1992; Piersma, 2011), implying that migration may be relatively inflexible and  
733 subject to strong divergent selection. Further, migratory differences may impose and maintain  
734 reproductive isolation, through assortative mating based on behavior, timing, or plumage  
735 (Taylor & Friesen, 2017; Uy, Irwin, & Webster, 2018) or selection against intermediate  
736 phenotypes (Delmore & Irwin, 2014). Despite these expectations, novel migration behavior  
737 can evolve rapidly in new circumstances (Able & Belthoff, 1998; Berthold, Helbig, Mohr, &  
738 Querner, 1992), and migratory differences may arise and persist in sympatry without  
739 precluding gene flow (Delmore et al., 2020; Pérez-Tris et al., 2004).

740

741 Neither the phylogeographic history nor current population structure we inferred in red knots  
742 is consistent with migratory phenotypes imposing limitations to innovation (i.e. colonization  
743 of new flyways) or reproductive barriers between populations. In general, the recently  
744 established global distribution attests to a lack of geographic barriers, and a general flexibility  
745 to adjust or invent migratory phenotypes based on novel and changing conditions, either

746 through phenotypic plasticity or rapid micro-evolutionary processes involving selection (e.g.  
747 Delmore et al., 2020). More specifically, we found evidence for multiple cases of post-glacial  
748 admixture (Figure 6), implying relatively common and successful mixing of populations with  
749 very different migratory phenotypes. For example, the admixture of Nearctic and Palearctic  
750 red knots that led to present-day *islandica* must have involved interbreeding between  
751 individuals with markedly different migration directions and distances, and with annual  
752 cycles adapted to non-breeding conditions on separate continents (Buehler & Piersma, 2008).  
753

754 Conversely, reproductive barriers should be least likely between populations with very  
755 similar migrations, because gene flow would involve fewer necessary adjustments to the  
756 annual cycle. It is easy to envision ‘flyway switching’ between populations that share  
757 wintering or migratory staging sites, because their migratory phenotypes necessarily share  
758 some key aspects (e.g. timing, navigation), and social interactions may lead individuals to  
759 join migrating flocks of another breeding population and arrive at the ‘wrong’ breeding area.  
760 This is perhaps most likely for ‘naïve’ young birds on their first northward migration, which  
761 is often considered the most likely source of innovation and gene flow in migratory birds.  
762 Therefore, it is surprising that we detected clear structure *within* flyways, where we would  
763 least expect barriers to gene flow. For example, *roselaari* red knots breeding on Wrangel  
764 Island and Alaska are thought to inter-mix throughout migration and winter in western North  
765 America (Carmona et al., 2013), and the final northward flight to the breeding grounds is  
766 essentially the same direction (Figure 1). Similarly, *piersmai* and *rogersi* share wintering sites  
767 throughout Australia and New Zealand, and then use a small number of shared staging sites  
768 in the Yellow Sea during migration (Rogers et al., 2010). Since very low levels of  
769 immigration are expected to effectively homogenize populations (Slatkin, 1985), detectable  
770 structure between these population pairs in such a brief time frame implies that successful



771 interbreeding is surprisingly rare. This suggests that the greatest barriers to gene flow are not  
772 necessarily related to migration, but occur in the breeding season.

773

774 Intriguingly, we found the weakest neutral genetic differentiation precisely where  
775 neighboring populations display the greatest phenotypic differences: between *rufa* and  
776 *islandica*. These subspecies migrate, respectively, the longest (up to ca. 15,000 km) and  
777 shortest (ca. 3,000–4,000 km) distances in the species (Figure 1), and also differ markedly in  
778 direction of migration (Piersma, 2011). Further, they differ in body size and plumage and  
779 largely spend the non-breeding season in opposite hemispheres. Because this phenotypic  
780 divergence appears quite recent (<5,000 ypb; Figure 6), it is possible that reproductive  
781 isolation does occur and has yet to manifest in significant neutral genetic differentiation,  
782 perhaps involving specific portions of the genome under selection and not included among  
783 our genetic markers. However, the adjacent, and possibly overlapping, breeding areas in the  
784 Canadian Arctic may represent a ‘migratory divide’, in which phenotypic divergence occurs  
785 across a geographic threshold or cline, with no necessary interruption of gene flow (Bensch,  
786 Andersson, & Akesson, 1999; Delmore et al., 2020).

787

788 Taken together, our results suggest that reproductive isolation in red knots is not primarily a  
789 product of migratory phenotypes *per se*, but rather of adaptations to breeding conditions that  
790 vary geographically. Scolopacid shorebirds are well-known for high site-fidelity and tightly-  
791 scheduled annual routines (Piersma & Baker, 2000), adapted to ensure timely arrival at  
792 breeding sites (Åkesson et al., 2017), and perhaps entrained by early-life conditions at the  
793 natal site (e.g. Ciarleglio, Axley, Strauss, Gamble, & McMahon, 2011). Accordingly, we  
794 believe that timing of breeding is the most likely driver of reproductive isolation in red knots,  
795 rather than lack of dispersal, assortative mating by size, plumage, or pre-nuptial behavior, or

796 endogenous pre- or post-zygotic incompatibility. Therefore, we hypothesize that both  
797 historical admixture events and present-day gene flow are regulated by the degree of  
798 similarity in endogenous processes that guide readiness for breeding (Karagicheva et al.,  
799 2016).

800

801 In this study, we have used genome-wide neutral genetic variation to reconstruct the global  
802 population structure and phylogeographic time-frame in which divergent migratory  
803 phenotypes have evolved in red knots. The general lack of correspondence between the  
804 degree of phenotypic and neutral genetic differentiation in our study leaves unanswered the  
805 question of whether selection is primarily driving the observed differences. The global  
806 colonization by red knots suggests a complex history involving incomplete lineage sorting,  
807 vicariance, secondary contact, gene flow (perhaps asymmetrical), and some interplay of  
808 phenotypic plasticity and selective processes. Future work including improved geographic  
809 sampling and a higher-resolution representation of the genome will allow us to jointly assess  
810 the contributions of selection and neutral evolutionary forces to global phenotypic  
811 diversification in red knots.

812

### 813 **ACKNOWLEDGEMENTS**

814 This work is dedicated to the memory of Allan Baker. We thank Per Palsbøll and Ritsert  
815 Jansen for guidance and help to develop the funding proposal, and to members of the Palsbøll  
816 and Fontaine labs at the University of Groningen for helpful discussions during project  
817 conception and analysis, respectively. We thank Anneke Bol at NIOZ Royal Netherlands  
818 Institute for Sea Research for DNA extractions of Mauritania samples, and Marco van der  
819 Velde for lab assistance in Groningen. We thank Eric Johnson and Paul Etter (SNPsaurus) for  
820 sequencing, bioinformatics, and technical expertise. We thank the Center for Information

821 Technology of the University of Groningen (particularly Bob Dröge and Cristian Marocico)  
822 for their support and for providing access to the Peregrine high-performance computing  
823 cluster. For providing samples, we thank Massey University, Royal Ontario Museum (Oliver  
824 Haddrath and Mark Peck), Moscow Lomonosov State University Zoological Museum, NIOZ,  
825 U.S. Fish & Wildlife Service (especially Lucas DeCicco and Nicholas Hajdukovich), and  
826 University of Groningen. Australia sampling was made possible by the support of BirdLife  
827 Netherlands at the initiative of its director, the late Adri de Gelder, during the first few years  
828 of operation of Global Flyway Network. This project was supported by a Dutch Research  
829 Council (NWO) grant to TP (ALW-Open Programme grant, ‘Ecological drivers of global  
830 flyway evolution’ 824.01.001).

831

#### 832 **AUTHOR CONTRIBUTIONS**

833 JRC, YIV, MCF and TP conceived and designed the study. PFB, CJH, JtH, JAJ, and PST  
834 organized and performed field sampling and maintained individual resight-history databases.  
835 JRC and YIV curated samples and conducted the labwork. JRC and MCF analyzed and  
836 interpreted the data, with assistance from YIV. JRC and MCF wrote the manuscript, with  
837 major contributions from YIV and TP. All authors provided feedback and edited the  
838 manuscript.

839

#### 840 **DATA AVAILABILITY STATEMENT**

841 Demultiplexed nextRAD short-read data were deposited in NCBI’s SRA archives under  
842 BioProject ID *to be announced* (TBA). Associated files (VCF and metadata) were deposited  
843 in DRYAD (doi: TBA).

844

## 845 REFERENCES

- 846 Able, K. P., & Belthoff, J. R. (1998). Rapid “evolution” of migratory behaviour in the  
847 introduced house finch of eastern North America. *Proceedings of the Royal Society B:*  
848 *Biological Sciences*, 265, 2063–2071. doi: 10.1098/rspb.1998.0541
- 849 Åkesson, S., Ilieva, M., Karagicheva, J., Rakhimberdiev, E., Tomotani, B., & Helm, B.  
850 (2017). Timing avian long-distance migration: From internal clock mechanisms to  
851 global flights. *Philosophical Transactions of the Royal Society B: Biological Sciences*,  
852 372(1734). doi: 10.1098/rstb.2016.0252
- 853 Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for  
854 individual ancestry estimation. *BMC Bioinformatics*, 12, 246. doi: 10.1186/1471-2105-  
855 12-246
- 856 Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of  
857 ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. doi:  
858 10.1101/gr.094052.109.vidual
- 859 Arcones, A., Ponti, R., Ferrer, X., & Vieites, D. R. (2020). Pleistocene glacial cycles as  
860 drivers of allopatric differentiation in Arctic shorebirds. *Journal of Biogeography*, 48(4),  
861 747–759. doi: 10.1111/jbi.14023
- 862 Atkinson, P. W., Baker, A. J., Bevan, R. M., Clark, N. a., Cole, K. B., González, P. M., ...  
863 Robinson, R. A. (2005). Unravelling the migration and moult strategies of a long-  
864 distance migrant using stable isotopes: Red Knot *Calidris canutus* movements in the  
865 Americas. *Ibis*, 147(4), 738–749. doi: 10.1111/j.1474-919x.2005.00455.x
- 866 Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., ... Saunders,  
867 N. C. (1987). Intraspecific phylogeography—the mitochondrial-DNA bridge between  
868 population genetics and systematics. *Annual Review of Ecology, Evolution, and*  
869 *Systematics*, 40, 593–612.
- 870 Avise, J. C., & Walker, D. (1998). Pleistocene phylogeographic effects on avian populations  
871 and the speciation process. *Proceedings of the Royal Society B: Biological Sciences*,  
872 265, 457–463.
- 873 Baker, A. J., González, P. M., Piersma, T., Niles, L. J., De Lima Serrano Do Nascimento, I.,  
874 Atkinson, P. W., ... Aarts, G. (2004). Rapid population decline in red knots: Fitness  
875 consequences of decreased refuelling rates and late arrival in Delaware Bay.  
876 *Proceedings of the Royal Society B: Biological Sciences*, 271(1541), 875–882. doi:  
877 10.1098/rspb.2003.2663
- 878 Batchelor, C. L., Margold, M., Krapp, M., Murton, D. K., Dalton, A. S., Gibbard, P. L., ...  
879 Manica, A. (2019). The configuration of Northern Hemisphere ice sheets through the  
880 Quaternary. *Nature Communications*, 10(1), 3713. doi: 10.1038/s41467-019-11601-2
- 881 Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology.  
882 *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 379–406. doi:  
883 10.1146/annurev-ecolsys-102209-144621
- 884 Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation  
885 in population genetics. *Genetics*, 162(4), 2025–2035.
- 886 Bensch, S., Andersson, T., & Akesson, S. (1999). Morphological and molecular variation  
887 across a migratory divide in Willow Warblers, *Phylloscopus trochilus*. *Evolution*, 53(6),  
888 1925–1935.
- 889 Berthold, P., & Helbig, A. J. (1992). The genetics of bird migration: stimulus, timing, and  
890 direction. *Ibis*, 134, 35–40. doi: 10.1111/j.1474-919X.1992.tb04731.x
- 891 Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid microevolution of  
892 migratory behaviour in a wild bird species. *Nature*, 360, 668–670. doi:  
893 10.1038/360668a0

- 894 Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate  
895 demography over space and time: Some cons, many pros. *Molecular Ecology*, *19*(13),  
896 2609–2625. doi: 10.1111/j.1365-294X.2010.04690.x
- 897 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... Drummond, A. J.  
898 (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS*  
899 *Computational Biology*, *10*(4), 1–6. doi: 10.1371/journal.pcbi.1003537
- 900 Bouckaert, R. R. (2010). DensiTree: Making sense of sets of phylogenetic trees.  
901 *Bioinformatics*, *26*(10), 1372–1373. doi: 10.1093/bioinformatics/btq110
- 902 Boyd, H., & Piersma, T. (2001). Changing balance between survival and recruitment explains  
903 population trends in Red Knots *Calidris canutus islandica* wintering in Britain, 1969–  
904 1995. *Ardea*, *89*(2), 301–317.
- 905 Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using  
906 sequence-based markers. *Genetica*, *135*(3), 439–455. doi: 10.1007/s10709-008-9293-3
- 907 Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & Roychoudhury, A. (2012).  
908 Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a  
909 full coalescent analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932. doi:  
910 10.1093/molbev/mss086
- 911 Buehler, D. M., & Baker, A. J. (2005). Population divergence times and historical  
912 demography in Red Knots and Dunlins. *Condor*, *107*, 497–513.
- 913 Buehler, D. M., Baker, A. J., & Piersma, T. (2006). Reconstructing palaeoflyways of the late  
914 Pleistocene and early Holocene Red Knot *Calidris canutus*. *Ardea*, *94*(3), 485–498.
- 915 Buehler, D. M., & Piersma, T. (2008). Travelling on a budget: predictions and ecological  
916 evidence for bottlenecks in the annual cycle of long-distance migrants. *Philosophical*  
917 *Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1490),  
918 247–266. doi: 10.1098/rstb.2007.2138
- 919 Bushnell, B. (2016). BMap. Retrieved from <http://sourceforge.net/projects/bbmap>
- 920 Carmona, R., Arce, N., Ayala-Perez, V., Buchanan, J. B., Salzer, L. J., Tomkovich, P. S., ...  
921 Newstead, D. (2013). Red Knot *Calidris canutus roselaari* migration connectivity,  
922 abundance and non-breeding distribution along the Pacific coast of the Americas. *Wader*  
923 *Study Group Bulletin*, *120*(3), 168–180.
- 924 Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J.  
925 (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets.  
926 *GigaScience*, *4*(1), 1–16. doi: 10.1186/s13742-015-0047-8
- 927 Ciarleglio, C. M., Axley, J. C., Strauss, B. R., Gamble, K. L., & McMahon, D. G. (2011).  
928 Perinatal photoperiod imprints the circadian clock. *Nature Neuroscience*, *14*(1), 25–27.  
929 doi: 10.1038/nn.2699
- 930 Conklin, J. R., Senner, N. R., Battley, P. F., & Piersma, T. (2017). Extreme migration and the  
931 individual quality spectrum. *Journal of Avian Biology*, *48*(1), 19–36. doi:  
932 10.1111/jav.01316
- 933 Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., ... Estoup,  
934 A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation  
935 inferences about population history using single nucleotide polymorphism, DNA  
936 sequence and microsatellite data. *Bioinformatics*, *30*(8), 1187–1189. doi:  
937 10.1093/bioinformatics/btt763
- 938 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin,  
939 R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.  
940 doi: 10.1093/bioinformatics/btr330
- 941 Delmore, K. E., & Irwin, D. E. (2014). Hybrid songbirds employ intermediate routes in a  
942 migratory divide. *Ecology Letters*, *17*(10), 1211–1218. doi: 10.1111/ele.12326
- 943 Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Ramos, J. S. L., Durieux, G., ...

944 Liedvogel, M. (2020). The evolutionary history and genomics of european blackcap  
945 migration. *ELife*, 9, 1–24. doi: 10.7554/eLife.54462

946 Dick, W. J., Piersma, T., & Prokosch, P. (1987). Spring migration of the Siberian Knots  
947 *Calidris canutus canutus*: results of a co-operative Wader Study Group project. *Ornis*  
948 *Scandinavica*, 18, 5–16.

949 Dyke, A. S. (2004). An outline of North American deglaciation with emphasis on central and  
950 northern Canada. *Developments in Quaternary Science*, 2(Part B), 373–424. doi:  
951 10.1016/S1571-0866(04)80209-4

952 Ehlers, J., & Gibbard, P. L. (2007). The extent and chronology of Cenozoic Global  
953 Glaciation. *Quaternary International*, 164–165, 6–20. doi: 10.1016/j.quaint.2006.10.008

954 Estoup, A., Lombaert, E., Marin, J.-M., Guillemaud, T., Pudlo, P., Robert, C. P., & Cornuet,  
955 J.-M. (2012). Estimation of demo-genetic model probabilities with Approximate  
956 Bayesian Computation using linear discriminant analysis on summary statistics.  
957 *Molecular Ecology Resources*, 12(5), 846–855. doi: 10.1111/j.1755-0998.2012.03153.x

958 Fedorov, V. B., & Stenseth, N. C. (2002). Multiple glacial refugia in the North American  
959 Arctic: inference from phylogeography of the collared lemming (*Dicrostonyx*  
960 *groenlandicus*). *Proceedings of the Royal Society B: Biological Sciences*, 269(1505),  
961 2071–2077. doi: 10.1098/rspb.2002.2126

962 Fitak, R. R. (2019). optM: an R package to optimize the number of migration edges using  
963 threshold models. Retrieved from <https://github.com/cran/OptM>

964 Gualtieri, L., Glushkova, O., & Brigham-Grette, J. (2000). Evidence for restricted ice extent  
965 during the last glacial maximum in the Koryak Mountains of Chukotka, far eastern  
966 Russia. *Bulletin of the Geological Society of America*, 112(7), 1106–1118. doi:  
967 10.1130/0016-7606(2000)112<1106:EFRIED>2.0.CO;2

968 Hassell, C., Southey, I., Boyle, A., & Yang, H.-Y. (2011). Red Knot *Calidris canutus*:  
969 subspecies and migration in the East Asian-Australasian flyway—where do all the Red  
970 Knot go? *BirdingASIA*, 16, 89–93.

971 Hewitt, G. M. (2000). The genetic legacy of the Quarternary ice ages. *Nature*, 405, 907–913.  
972 doi: 10.1038/35016000

973 Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H.,  
974 Johnson, J. B., ... Yoder, A. D. (2010). Phylogeography's past, present, and future: 10  
975 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54(1), 291–301. doi:  
976 10.1016/j.ympev.2009.09.016

977 Johnson, N. K., & Cicero, C. (2004). New mitochondrial DNA data affirm the importance of  
978 pleistocene speciation in North American birds. *Evolution*, 58(5), 1122–1130. doi:  
979 10.1111/j.0014-3820.2004.tb00445.x

980 Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic  
981 analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*,  
982 2, e281. doi: 10.7717/peerj.281

983 Karagicheva, J., Rakhimberdiev, E., Dekinga, A., Brugge, M., Koolhaas, A., ten Horn, J., &  
984 Piersma, T. (2016). Seasonal time keeping in a long-distance migrating shorebird.  
985 *Journal of Biological Rhythms*, 31(5), 509–521. doi: 10.1177/0748730416655929

986 Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). DiveRsity:  
987 An R package for the estimation and exploration of population genetics parameters and  
988 their associated errors. *Methods in Ecology and Evolution*, 4(8), 782–788. doi:  
989 10.1111/2041-210X.12067

990 Klicka, J., & Zink, R. M. (1997). The importance of recent ice ages in speciation: A failed  
991 paradigm. *Science*, 277(5332), 1666–1669. doi: 10.1126/science.277.5332.1666

992 Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and*  
993 *Systematics*, 40, 593–612.

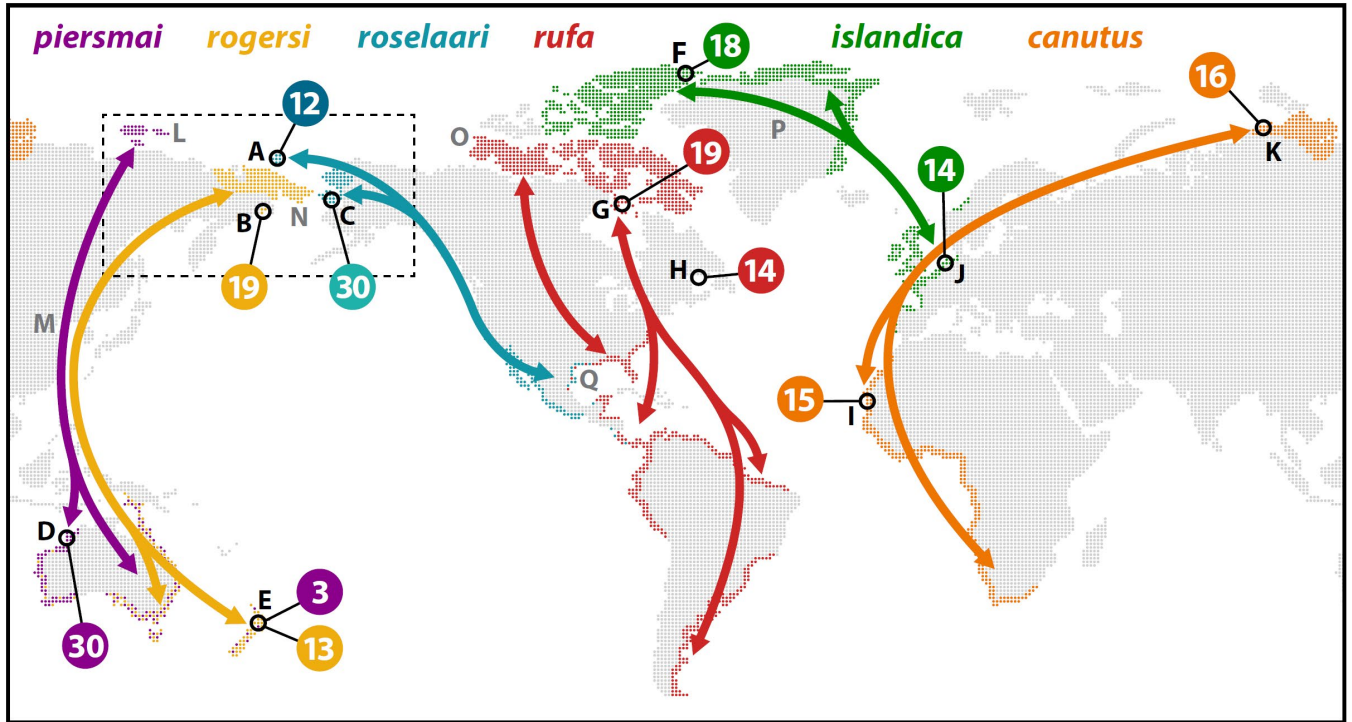
- 994 Kok, E. M. A., Tibbitts, T. L., Douglas, D. C., Howey, P. W., Dekinga, A., Gnep, B., &  
 995 Piersma, T. (2020). A red knot as a black swan: how a single bird shows navigational  
 996 abilities during repeat crossings of the Greenland Icecap. *Journal of Avian Biology*,  
 997 *51*(8), 1–11. doi: 10.1111/jav.02464
- 998 Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015).  
 999 Clumpak: A program for identifying clustering modes and packaging population  
 1000 structure inferences across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. doi:  
 1001 10.1111/1755-0998.12387
- 1002 Kraaijeveld, K., & Nieboer, E. N. (2000). Late Quaternary paleogeography and evolution of  
 1003 arctic breeding waders. *Ardea*, *88*(2), 193–205.
- 1004 Lappo, E. G., Tomkovich, P. S., & Syroechkovskiy, E. E. (2012). *Atlas of breeding waders in*  
 1005 *the Russian Arctic*. Moscow, Russia: UF Ofsetnaya Pechat.
- 1006 Lathrop, R. G., Niles, L., Smith, P., Peck, M., Dey, A., Sacatelli, R., & Bognar, J. (2018).  
 1007 Mapping and modeling the breeding habitat of the Western Atlantic Red Knot (*Calidris*  
 1008 *canutus rufa*) at local and regional scales. *Condor*, *120*(3), 650–665. doi:  
 1009 10.1650/CONDOR-17-247.1
- 1010 Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for  
 1011 connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299.  
 1012 doi: 10.1093/bioinformatics/btr642
- 1013 Manley, W. F. (2002). *Postglacial flooding of the Bering Land Bridge: A geospatial*  
 1014 *animation*. INSTAAR, University of Colorado, v1. Retrieved June 14, 2021, from  
 1015 [http://instaar.colorado.edu/groups/QGISL/bering\\_land\\_bridge/](http://instaar.colorado.edu/groups/QGISL/bering_land_bridge/)
- 1016 McLaughlin, J. F., Faircloth, B. C., Glenn, T. C., & Winker, K. (2020). Divergence, gene  
 1017 flow, and speciation in eight lineages of trans-Beringian birds. *Molecular Ecology*,  
 1018 *29*(18), 3526–3542. doi: 10.1111/mec.15574
- 1019 Méndez, V., Alves, J. A., Gill, J. A., & Gunnarsson, T. G. (2018). Patterns and processes in  
 1020 shorebird survival rates: a global review. *Ibis*, *160*(4), 723–741. doi: 10.1111/ibi.12586
- 1021 Milanese, M., Capomaccio, S., Vajana, E., Bomba, L., Garcia, J. F., Ajmone-Marsan, P., &  
 1022 Colli, L. (2017). BITE: an R package for biodiversity analyses. *BioRxiv*, 181610. doi:  
 1023 10.1101/181610
- 1024 Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2014).  
 1025 Genotyping-by-sequencing in ecological and conservation genomics. *Molecular*  
 1026 *Ecology*, *22*(11), 2841–2847. doi: 10.1111/mec.12350. Genotyping-by-sequencing
- 1027 Nebel, S., Piersma, T., van Gils, J. A., Dekinga, A., & Spaans, B. (2000). Length of stopover,  
 1028 fuel storage and a sex-bias in the occurrence of Red Knots *Calidris c. canutus* and *C.c.*  
 1029 *islandica* in the Wadden Sea during southward migration. *Ardea*, *88*(2), 165–176.
- 1030 Nei, M. (1972). Genetic distance between populations. *American Naturalist*, *106*(949), 283–  
 1031 292.
- 1032 Newstead, D. J., Niles, L. J., Porter, R. R., Dey, A. D., Burger, J., & Fitzsimmons, O. N.  
 1033 (2013). Geolocation reveals mid-continent migratory routes and Texas wintering areas  
 1034 of Red Knots *Calidris canutus rufa*. *Wader Study Group Bulletin*, *120*, 53–59.
- 1035 Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of  
 1036 population genetic differentiation in the wild: Isolation by dispersal limitation, isolation  
 1037 by adaptation and isolation by colonization. *Molecular Ecology*, *22*(24), 5983–5999.  
 1038 doi: 10.1111/mec.12561
- 1039 Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS*  
 1040 *Genetics*, *2*(12), 2074–2093. doi: 10.1371/journal.pgen.0020190
- 1041 Patton, H., Hubbard, A., Andreassen, K., Auriac, A., Whitehouse, P. L., Stroeven, A. P., ...  
 1042 Hall, A. M. (2017). Deglaciation of the Eurasian ice sheet complex. *Quaternary Science*  
 1043 *Reviews*, *169*, 148–172. doi: 10.1016/j.quascirev.2017.05.019

- 1044 Pérez-Tris, J., Bensch, S., Carbonell, R., Helbig, A. J., & Tellería, J. L. (2004). Historical  
 1045 diversification of migration patterns in a passerine bird. *Evolution*, 58(8), 1819–1832.  
 1046 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15446433>
- 1047 Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from  
 1048 genome-wide allele frequency data. *PLoS Genetics*, 8(11), e1002967. doi:  
 1049 10.1371/journal.pgen.1002967
- 1050 Pielou, E. C. (1991). *After the Ice Age: The Return of Life to Glaciated North America*.  
 1051 Chicago, IL: University of Chicago Press.
- 1052 Piersma, T. (2007). Using the power of comparison to explain habitat use and migration  
 1053 strategies of shorebirds worldwide. *Journal of Ornithology*, 148(S1), 45–59. doi:  
 1054 10.1007/s10336-007-0240-3
- 1055 Piersma, T. (2011). Flyway evolution is too fast to be explained by the modern synthesis:  
 1056 proposals for an ‘extended’ evolutionary research agenda. *Journal of Ornithology*, 152  
 1057 (suppl.), S151–S159. doi: 10.1007/s10336-011-0716-z
- 1058 Piersma, T., & Baker, A. J. (2000). Life history characteristics and the conservation of  
 1059 migratory shorebirds. In L. M. Gosling & W. J. Sutherland (Eds.), *Behaviour and*  
 1060 *Conservation* (pp. 105–124). Cambridge, UK: Cambridge University Press.
- 1061 Piersma, T., & Davidson, N. (1992). The migration of Knots. *Wader Study Group Bulletin*,  
 1062 64(suppl. 1), 1–209.
- 1063 Piersma, T., Kok, E. M. A., Hassell, C. J., Peng, H., Verkuil, Y. I., Lei, G., ... Chan, Y.  
 1064 (2021). When a typical jumper skips: itineraries and staging habitats used by Red Knots  
 1065 (*Calidris canutus piersmai*) migrating between northwest Australia and the New  
 1066 Siberian Islands. *Ibis*. doi: 10.1111/ibi.12964
- 1067 Piersma, T., Lok, T., Chen, Y., Hassell, C. J., Yang, H. Y., Boyle, A., ... Ma, Z. (2016).  
 1068 Simultaneous declines in summer survival of three shorebird species signals a flyway at  
 1069 risk. *Journal of Applied Ecology*, 53(2), 479–490. doi: 10.1111/1365-2664.12582
- 1070 Piersma, T., Pérez-Tris, J., Mouritsen, H., Bauchinger, U., & Bairlein, F. (2005). Is there a  
 1071 “migratory syndrome” common to all migrant birds? *Annals of the New York Academy*  
 1072 *of Sciences*, 1046, 282–293. doi: 10.1196/annals.1343.026
- 1073 Piersma, T., Rogers, D. I., González, P. M., Zwarts, L., Niles, L. J., Donascimento, I. de L.  
 1074 S., ... Baker, A. J. (2005). Fuel storage rates before northward flights in Red Knots  
 1075 worldwide: Facing the severest ecological constraint in tropical intertidal environments?  
 1076 In R. Greenberg & P. P. Marra (Eds.), *Birds of two worlds: the ecology and evolution of*  
 1077 *migratory birds* (pp. 262–274). Baltimore, USA: Johns Hopkins University Press.
- 1078 Provan, J., & Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia.  
 1079 *Trends in Ecology and Evolution*, 23(10), 564–571. doi: 10.1016/j.tree.2008.06.010
- 1080 Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2016).  
 1081 Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866. doi:  
 1082 10.1093/bioinformatics/btv684
- 1083 Rakhimberdiev, E., van den Hout, P. J., Brugge, M., Spaans, B., & Piersma, T. (2015).  
 1084 Seasonal mortality and sequential density dependence in a migratory bird. *Journal of*  
 1085 *Avian Biology*, 46(4), 332–341. doi: 10.1111/jav.00701
- 1086 Rambaut, A. (2010). FigTree v1.3.1. Retrieved from <http://tree.bio.ed.ac.uk/software/figtree/>
- 1087 Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior  
 1088 summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5),  
 1089 901–904. doi: 10.1093/sysbio/syy032
- 1090 Raynal, L., Marin, J. M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2019). ABC  
 1091 random forests for Bayesian parameter inference. *Bioinformatics*, 35(10), 1720–1728.  
 1092 doi: 10.1093/bioinformatics/bty867
- 1093 Richardson, D. S., Jury, F. L., Blaakmeer, K., Komdeur, J., & Burke, T. (2001). Parentage



- 1094 assignment and extra-group paternity in a cooperative breeder: The Seychelles warbler  
 1095 (*Acrocephalus sechellensis*). *Molecular Ecology*, 10(9), 2263–2273. doi:  
 1096 10.1046/j.0962-1083.2001.01355.x
- 1097 Rodríguez-Ramilo, S. T., & Wang, J. (2012). The effect of close relatives on unsupervised  
 1098 Bayesian clustering algorithms in population genetic structure analysis. *Molecular*  
 1099 *Ecology Resources*, 12(5), 873–884. doi: 10.1111/j.1755-0998.2012.03156.x
- 1100 Rogers, D. I., Yang, H.-Y., Hassell, C. J., Boyle, A. N., Rogers, K. G., Chen, B., ... Piersma,  
 1101 T. (2010). Red Knots (*Calidris canutus piersmai* and *C. c. rogersi*) depend on a small  
 1102 threatened staging area in Bohai Bay, China. *Emu*, 110, 307–315.
- 1103 Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to  
 1104 practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, 3, e1106.  
 1105 doi: 10.7717/peerj.1106
- 1106 Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and*  
 1107 *Systematics*, 16, 393–430. doi: 10.1146/annurev.ecolsys.16.1.393
- 1108 Stewart, J. R., & Dalén, L. (2008). Is the glacial refugium concept relevant for northern  
 1109 species? A comment on Pruett and Winker 2005. *Climatic Change*, 86, 19–22. doi:  
 1110 10.1007/s10584-007-9366-9
- 1111 Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: Individualistic  
 1112 responses of species in space and time. *Proceedings of the Royal Society B: Biological*  
 1113 *Sciences*, 277(1682), 661–671. doi: 10.1098/rspb.2009.1272
- 1114 Studds, C. E., Kendall, B. E., Murray, N. J., Wilson, H. B., Rogers, D. I., Clemens, R. S., ...  
 1115 Fuller, R. A. (2017). Rapid population decline in migratory shorebirds relying on  
 1116 Yellow Sea tidal mudflats as stopover sites. *Nature Communications*, 8, 14895. doi:  
 1117 10.1038/ncomms14895
- 1118 Taylor, R. S., & Friesen, V. L. (2017). The role of allochry in speciation. *Molecular*  
 1119 *Ecology*, 26(13), 3330–3342. doi: 10.1111/mec.14126
- 1120 Tomkovich, P. S. (1992). An analysis of the geographic variability in Knots *Calidris canutus*  
 1121 based on museum skins. *Wader Study Group Bulletin*, 64 (suppl), 17–23.
- 1122 Tomkovich, P. S. (2001). A new subspecies of red knot *Calidris canutus* from the New  
 1123 Siberian islands. *Bulletin of the British Ornithologists' Club*, 121(4), 257–263.
- 1124 Tomkovich, P. S., Porter, R. R., Loktionov, E. Y., & Niles, L. J. (2013). Pathways and  
 1125 staging areas of Red Knots *Calidris canutus rogersi* breeding in southern Chukotka, Far  
 1126 Eastern Russia. *Wader Study Group Bulletin*, 120(3), 181–193.
- 1127 Tomkovich, P. S., & Riegen, A. C. (2000). Mixing of red knot populations in Australia: some  
 1128 thoughts. *Stilt*, 37, 25–27.
- 1129 Uy, J. A. C., Irwin, D. E., & Webster, M. S. (2018). Behavioral isolation and incipient  
 1130 speciation in birds. *Annual Review of Ecology, Evolution, and Systematics*, 49, 1–24.  
 1131 doi: 10.1146/annurev-ecolsys-110617-062646
- 1132 van Gils, J. A., Lisovski, S., Lok, T., Meissner, W., Ożarowska, A., de Fouw, J., ... Klaassen,  
 1133 M. (2016). Body shrinkage due to Arctic warming reduces red knot fitness in tropical  
 1134 wintering range. *Science*, 352(6287), 819–822. doi: 10.1126/science.aad6351
- 1135 van Roomen, M., Nagy, S., Foppen, R., Dodman, T., Citegetse, G., & Ndiaye, A. (2015).  
 1136 *Status of coastal waterbird populations in the East Atlantic Flyway 2014: With special*  
 1137 *attention to flyway populations making use of the Wadden Sea*. Leeuwarden, The  
 1138 Netherlands: Programme Rich Wadden Sea.
- 1139 Verhoeven, M. A., van Eerbeek, J., Hassell, C. J., & Piersma, T. (2016). Fuelling and moult  
 1140 in Red Knots before northward departure: A visual evaluation of differences between  
 1141 ages, sexes and subspecies. *Emu*, 116(2), 158–167. doi: 10.1071/MU15035
- 1142 Wauchope, H. S., Shaw, J. D., Varpe, Ø., Lappo, E. G., Boertmann, D., Lanctot, R. B., &  
 1143 Fuller, R. A. (2017). Rapid climate-driven loss of breeding habitat for Arctic migratory

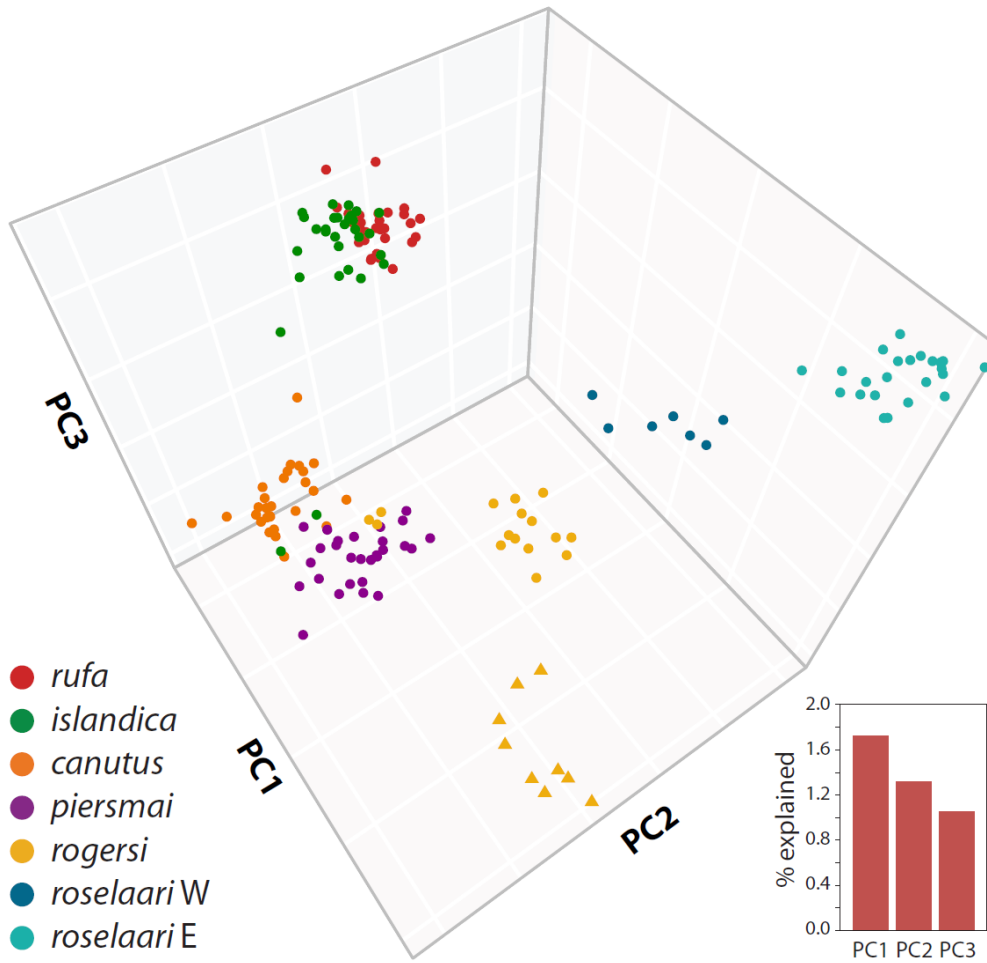
- 1144 birds. *Global Change Biology*, 23(3), 1085–1094. doi: 10.1111/gcb.13404
- 1145 Weir, B. S., & Cockerham, C. C. (1984). Estimating  $F$ -statistics for the analysis of population  
1146 structure. *Evolution*, 38(6), 1358–1370. doi: 10.2307/2409936
- 1147 Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings*  
1148 *of the Royal Society B: Biological Sciences*, 271(1551), 1881–1887. doi:  
1149 10.1098/rspb.2004.2803
- 1150 Wetlands International. (2021). *Waterbird Population Estimates*. Retrieved January 20, 2021,  
1151 from wpe.wetlands.org
- 1152 Wilson, J. I. M., Aubry, Y., Buidin, C., Rochepault, Y., & Baker, A. J. (2010). Three records  
1153 of Red Knots *Calidris canutus* possibly changing flyways. *Wader Study Group Bulletin*,  
1154 117(3), 192–193.
- 1155 Winker, K. (2010). On the origin of species through heteropatric differentiation: A review  
1156 and a model of speciation in migratory animals. *Ornithological Monographs*, 69, 1–30.  
1157 doi: 10.1525/om.2010.69.1.1.1
- 1158 Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., ... Wang, J. (2014). Comparative  
1159 genomics reveals insights into avian genome evolution and adaptation. *Science*,  
1160 346(6215), 1311–1321.
- 1161 Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-  
1162 performance computing toolset for relatedness and principal component analysis of SNP  
1163 data. *Bioinformatics*, 28(24), 3326–3328. doi: 10.1093/bioinformatics/bts606
- 1164 Zöckler, C., & O’Sullivan, J. (2005). New Zealand Red Knot breeding in Meinopylgino,  
1165 Chukotka, NE Russia. *Wader Study Group Bulletin*, 108, 76.
- 1166



1168  
1169

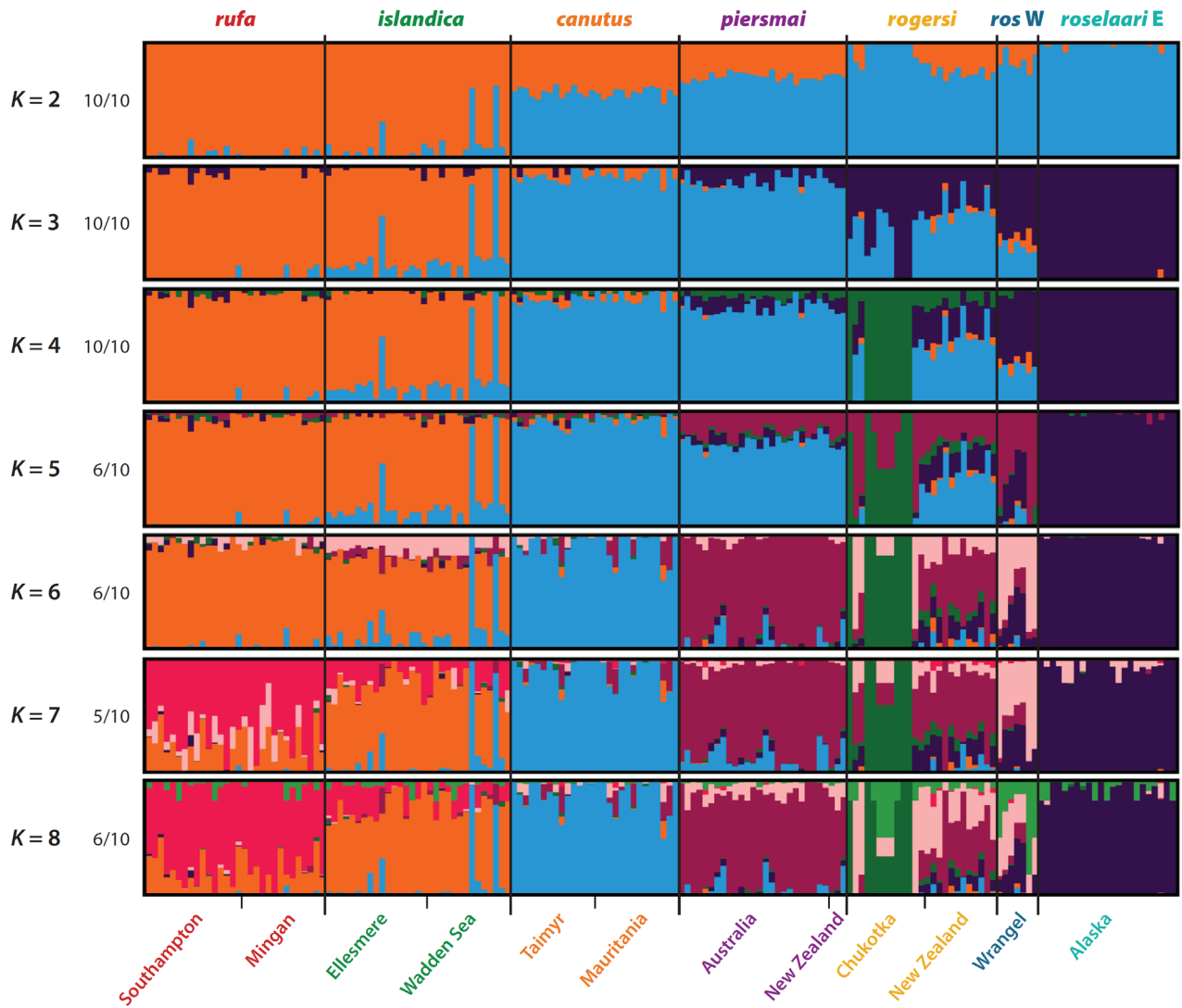
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179

**Figure 1.** Global distribution and sampling of red knots. For each of six recognized subspecies (indicated by color), arrows indicate general migration routes between breeding and (boreal) wintering areas (colored 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 gray): A = Wrangel Island, B = SE Chukotka, C = Seward Peninsula, D = Roebuck Bay, E = Foxton 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.



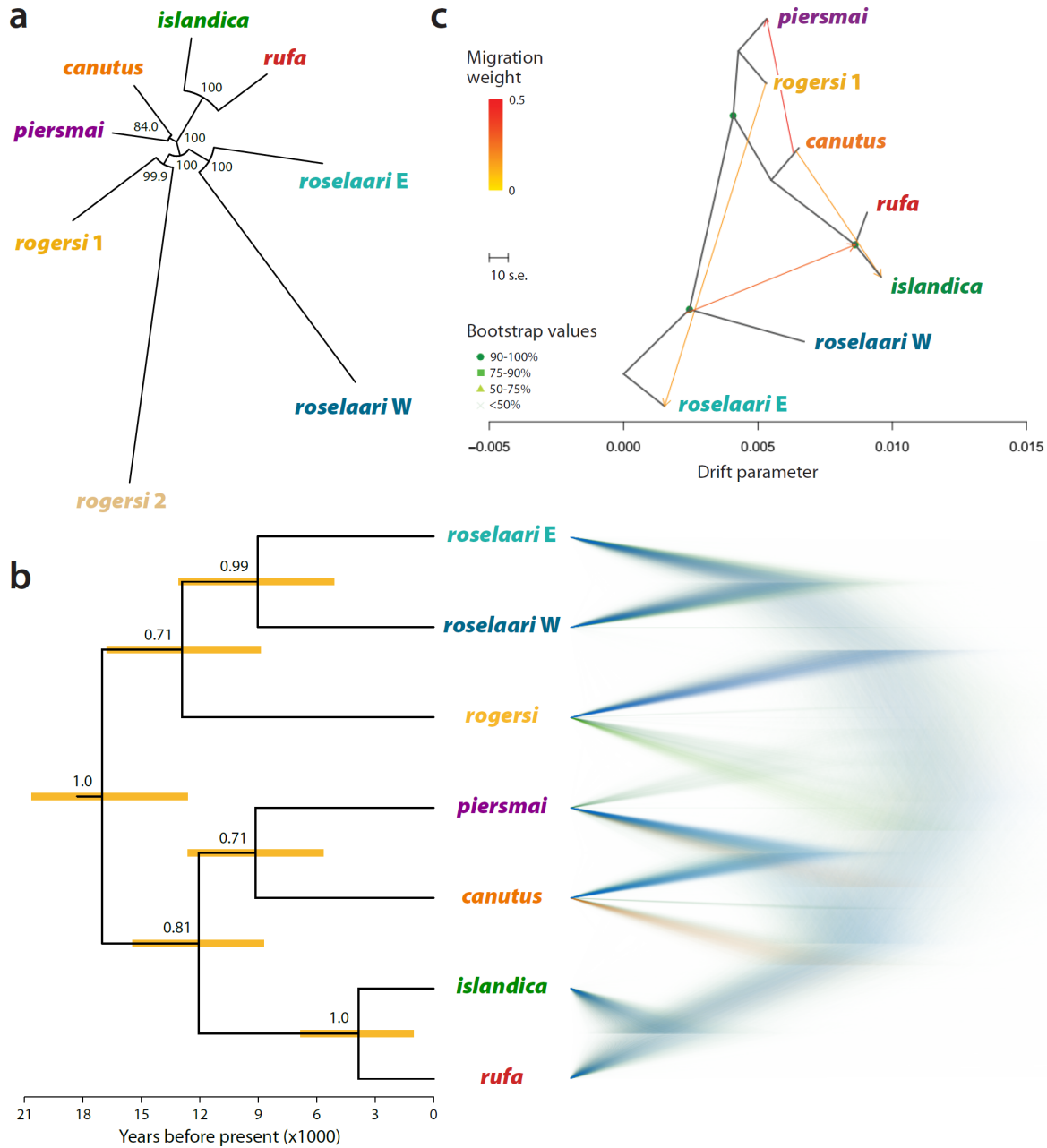
1181  
1182

1183 **Figure 2.** Genetic structure identified with the first three principal components of PCA  
 1184 analysis (explaining 4.1% of total variation). Note the nine individuals sampled in Chukotka  
 1185 (yellow triangles) that were distinguishable from the rest of *rogersi* in PC3. Also note two  
 1186 green and three yellow dots that fell among the *canutus/piersmai* cluster, indicating  
 1187 incorrect *a priori* identification of purported *islandica* and *rogersi* individuals, respectively.  
 1188 See Figure S3 for the 2D scatter plots for PC1 to PC8.



1190  
 1191  
 1192  
 1193  
 1194  
 1195  
 1196  
 1197

**Figure 3.** Individual genetic ancestries to 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. Names below plot indicate sampling locations (see Table S1).



1199

1200

1201

1202

1203

1204

1205

1206

1207

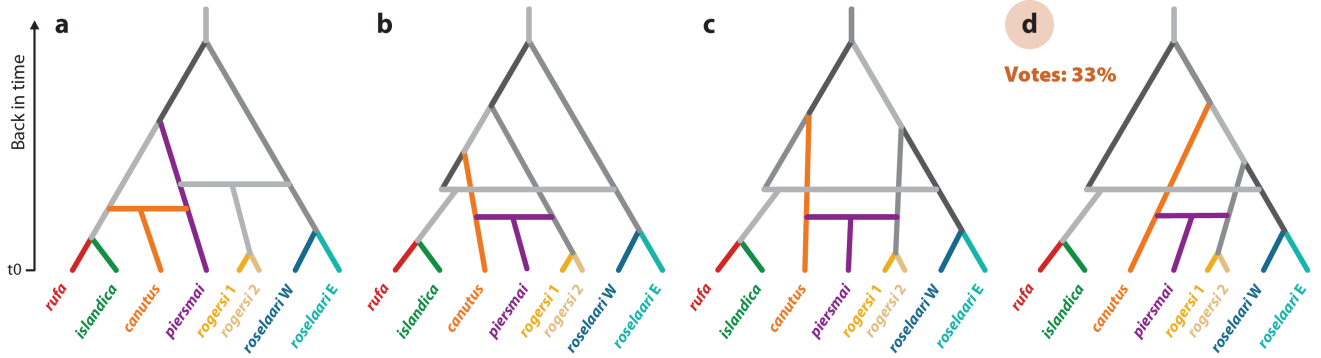
1208

1209

1210

**Figure 4.** Evolutionary relationships among populations of red knots. **(a)** Unrooted neighbor-joining tree of 8-population scenario, based on Nei’s minimum distance with bootstrapped node confidence. **(b)** Midpoint-rooted Bayesian coalescent-based trees derived with SNAPP from a reduced dataset of six individuals per population and 1,573 unlinked SNPs with no missing data. At left, the maximum clade credibility tree, with node support (numbers) and 95% highest posterior density intervals for node ages (orange bars). At right, DensiTree visualization of 50,000 trees, indicating frequency of the best-supported (in blue), next-supported (in green), and least-supported (in orange) topologies. **(c)** Maximum-likelihood (ML) tree inferred by TreeMix, including four migration edges. See Figure S7 for the TreeMix ML topologies with two to four migration edges.

1211



1212

1213

1214

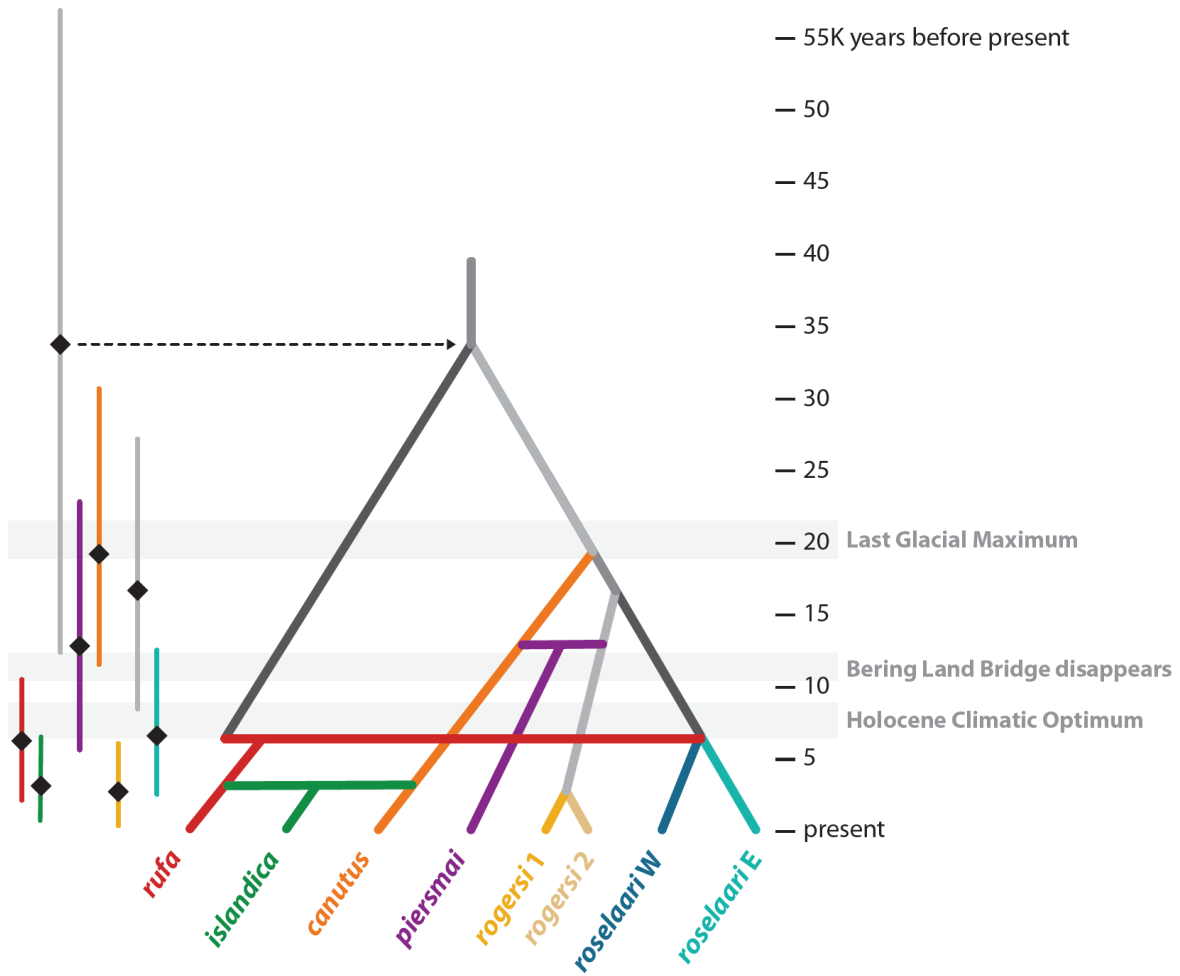
1215

1216

1217

1218

**Figure 5.** Scenarios tested in Step 1 of DIYABC analysis: **(a)** population topology suggested by NJ tree and SNAPP; **(b–d)** three possible rootings of the topology inferred by TreeMix. Each topology includes two admixture events (horizontal branches). Extant (sampled) populations are indicated by colors; inferred historical populations are shown in gray. The best-supported scenario was **(d)**.



1220  
 1221  
 1222  
 1223  
 1224  
 1225  
 1226  
 1227

**Figure 6.** Best-supported scenario in Step 2 of DIYABC analysis, 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).



1228  
1229  
1230

**Table 1.** Mean population pairwise  $F_{ST}$  and 95% confidence intervals of estimates. Non-significant  $F_{ST}$  values (CI includes zero) are in red.

	<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 [-0.001–0.012]						
<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 [-0.002–0.032]	0.009 [-0.003–0.031]			
<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 [-0.008–0.067]	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 [-0.006–0.063]

1231