# Global population structure in an arctic-breeding bird: Diversifying and stabilizing selection consistent with local adaptation

Keta Patel[1], Oliver Love[1], Kyle Wellband[2], Audrey Le Pogam[3], François Vézina[3], Noah Ashley[4], Frode Fossøy[5], Kevin Winker[6], Melanie Richter[7], H. Grant Gilchrist[8], Arne Moksnes[5], Bård Stokke[5], Jack Withrow[6], and Daniel Heath[1]

[1]University of Windsor
[2]Universite Laval
[3]Université du Québec à Rimouski
[4]Western Kentucky University
[5]Norwegian Institute for Nature Research
[6]University of Alaska Museum of the North
[7]Memphis Zoo
[8]Environment and Climate Change Canada

July 22, 2023

**Abstract**

Environmental variation among isolated populations can drive genetic differentiation by selection, while isolation alone results primarily in genetic drift. Genetic analyses can aid in identifying genetically isolated populations and population structure of a species across its range. Additionally, such analyses can provide indirect evidence of local adaptation through the comparison of allele frequencies at neutral and functional genetic markers, with the aim of identifying outlier loci consistent with the effects of selection. Here, we examine the genetic divergence and patterns of functional divergence among six breeding populations of arctic-breeding snow buntings (Plectrophenax nivalis). We genotyped 221 birds at 9 microsatellite markers and at 101 single nucleotide polymorphisms (SNPs) located within known-function genes. We identified substantial population differentiation using both marker types with relatively greater divergence and hence finer population structure using the microsatellite markers. While population structures resulting from the two marker types were in general agreement, functional SNPs showed evidence of stabilizing selection at both global and population pairwise levels, with a few key SNPs showing signatures of pairwise divergent selection, consistent with expectations of local adaptation. The observed complex and inconsistent pattern of pairwise divergence (selection) at key candidate-gene loci may reflect rapid environmental change decoupling locally adapted genotypes from actual local environmental conditions. Our work highlights microevolutionary changes that are likely to be very important not only in arctic-breeding songbirds, but in Arctic and Sub-Arctic vertebrates in general, which are experiencing strong environmental effects from accelerated climate change and human-induced stressors.

## Introduction

Local adaptation occurs when individuals from a given population exhibit higher fitness in their local environment than in other environments (Kaweki & Ebert, 2004). Because habitats are spatially and temporally variable, local environmental conditions determine which traits may be favoured by selection (Hoban et al., 2016), ultimately leading to divergent selection at the phenotypic and genotypic levels, resulting in local adaptation (Kaweki & Ebert, 2004). However, rapid environmental change generated by global climate change and other anthropogenic effects directly impact local environments and the locally adapted individuals inhabiting those changing environments (Atkins & Travis, 2010; Valladares et al., 2014). Consequently, environmental

change may decouple locally adapted allele frequencies from the current environmental conditions. Thus, assessing the genetic signatures of local adaptation in natural populations is critical for quantifying the scope of effects of changing environments on locally adapted populations (Aitken & Whitlock, 2013; Canosa et al., 2020; Lancaster et al., 2022).

Molecular genetics allows the incorporation of genetic evidence into the conservation and management of individuals, populations, and species across diverse taxonomic groups (Kirk & Freeland, 2011). Neutral molecular genetic markers are widely used to quantify genetic diversity, gene flow, and genetic differentiation among populations (Ouborg et al., 2010; Zimmerman et al., 2019); however, such data cannot inform conservation managers about potentially locally adapted or functional genetic variation. For example, the characterization of variation at functional loci (i.e., the genes that code for specific proteins) among populations provides insight into adaptive divergence (Luikart et al., 2003; Beaumont & Balding, 2004). Divergence in functional genotypes is expected to evolve relatively rapidly in response to natural selection, contrary to evolution by genetic drift alone (Kawecki & Ebert; 2004). The ideal analysis is thus a combination of rapidly evolving neutral genetic markers (e.g., microsatellite DNA, mtDNA sequence variation) and genetic polymorphisms linked to known-function genes (e.g., single nucleotide polymorphisms, or SNPs) that would characterize both gene flow and potential patterns of local adaptation.

Birds vary widely in their migratory patterns; from short-distance movements within seasons, to long-distance migrations covering substantial portions of the globe (Sekercioglu, 2007; Rolland et al., 2014). While such migratory life histories make them interesting candidate species for local adaptation analyses, there is limited information on patterns of genetic divergence that underlie the process of local adaptation (Kawecki & Ebert, 2004). Curiously, even though migratory birds are highly impacted by environmental changes (Both et al., 2006; Jonzén et al., 2006; Visser et al., 2015), little is understood about their genetic diversity or adaptive capacity, especially regarding the extent to which genomic variation is shaped by local environmental factors (Bay et al., 2021). Kuhn et al. (2013) used microsatellite, mitochondrial DNA, and a *Clock* gene marker in extant and historical populations of the pied flycatcher (*Ficedula hypoleuca* ), a long-distance migratory passerine, to test for potential effects of global climate change on their genetic structure. They showed stabilizing selection at the functional marker and suggested local adaptation had a greater effect on population genetic structure than recent climate change. Lehtonen et al. (2012) SNP-genotyped the same species across 17 sites across their breeding range and showed two (*follistatin* and *SWS1 opsin* ) of fourteen candidate genes involved in plumage colouration exhibited adaptive divergence – one of the few published studies of migratory passerines that quantified genetic diversity and differentiation using SNPs. To our knowledge, there has only been one published study of selection at genetic marker loci in an Arctic-breeding passerine; Tigano et al. (2017) showed that adaptation to migratory routes, or some other non-breeding ground-based environmental factor, drove the pattern of differentiation at genome-wide SNP markers in thick-billed murres (*Uria lomvia* ). Patterns of population differentiation in migratory bird species in general, and more specifically, in Arctic migratory birds, have been vastly understudied, despite the importance of population connectivity for their conservation in rapidly changing environments (Gu et al., 2021; Macdonald et al., 2012).

Snow buntings (*Plectrophenax nivalis)* are small, arctic-breeding passerines with a circumpolar distribution (Montgomerie & Lyon, 2020). Despite this species' global distribution, there are life-history differences among populations (e.g., migratory versus non-migratory; see Table 1). Most snow bunting populations migrate seasonally between high Arctic breeding grounds and temperate wintering grounds (Macdonald et al., 2012; Snell et al., 2018); however, some populations are non-migratory (e.g., Aleutian and Pribilof islands - Alaska, USA, and a high-altitude Scottish population; Montgomerie & Lyon, 2020). While globally abundant, census data suggest North American snow bunting populations have undergone substantial declines, with a reduction of 64% over the past four to five decades (Butcher & Niven, 2007). However, conservation efforts are hampered by logistic and data limitations, including a lack of information on the population structure and selection pressures on the birds.

To address population structure and functional divergence consistent with local adaptation, we assessed

global population structure among six geographically widespread breeding snow bunting populations. We first used microsatellite (presumed neutral) and transcriptome-derived SNP markers at known-function loci (potentially functional) to determine genetic divergence and hence assess population reproductive isolation. We then investigated population genetic divergence at functional marker loci, controlling for the effects of genetic drift using the neutral microsatellite marker data. As a largely migratory species, snow buntings are expected to have widely dispersed breeding populations across the globe (Montgomerie & Lyon, 2020), although current limited data suggest generally consistent migratory patterns (Lyngs, 2003; Macdonald et al., 2012; Snell et al., 2018; Montgomerie & Lyon, 2020). Hence, we predict some degree of reproductive isolation among the six breeding populations based on the expectation of consistent and separate migration routes; however, we recognized that the lack of explicit migratory data may result in unexpected gene flow and hence connectivity among some populations. We also predicted strong local selection pressures at the breeding grounds to result in patterns of local adaptation that would contribute to genetic differentiation at functional, candidate-gene loci under selection to improve local reproductive fitness. Specifically, we hypothesized that snow buntings are adapted to the local conditions on their breeding grounds, because selection pressures are strongest during the breeding period due to the high energetic demands of breeding, a short breeding season, and a correlation between local and regional climate and reproductive success (Falconer et al., 2008, Fossøy et al., 2014, Hoset et al., 2014). Although we expect patterns of local adaptation at some SNP marker loci, we predicted that the majority of our chosen functional gene SNPs would have neutral divergence patterns (consistent with genetic drift) with a few key SNPs exhibiting divergent and stabilizing selection. In this study we apply powerful population genetic approaches that can be used in future studies to facilitate the effective conservation and management of migratory species with the goal of facilitating the preservation of biodiversity.

## Methods

This project included the development and application of two types of molecular genetic markers: neutral microsatellite markers and functional gene locus SNP markers. It thus involved two types of samples: RNA samples for *de-novo* transcriptome assembly for SNP marker development, and DNA samples collected across the global breeding range of snow buntings for genotype data for the population genetic analyses. The population genetic study involved genotyping all samples at both microsatellite and SNP locus markers to determine population genetic divergence and patterns of functional divergence.

### Development of microsatellite markers

To develop snow bunting-specific microsatellite markers, multiple heterospecific primers were screened, and primers were chosen for strong amplification and high polymorphism on test samples (specifically, Mitivik Island DNA were used as a high-quality benchmark DNA for primer optimization). Some primer sequences were modified using the species-specific sequence information from an unrelated High Throughput sequencing project.

### DNA sample collection and extraction

A large-scale collaborative effort collected snow bunting tissue from populations across a wide geographic range, resulting in a total of 221 samples from six populations for DNA extraction (Figure 1, Table 2). All bird handling and sample collection was conducted under appropriate animal care permits (see Table 2). With the exception of the samples from Utqiagvik (Barrow), AK, USA, which were DNA extracted using a QIAamp DNA Mini Kit (Qiagen Inc., Toronto, ON, Canada), all samples were extracted using a solid phase reversible immobilization (SPRI) bead extraction originally optimized for bird cloacal and oral swabs (Vo & Jedlicka, 2014). Briefly, tissue was incubated in a solution containing lysis buffer, protein precipitation solution and zirconia-silica beads, followed by two rounds of homogenization and extraction of DNA from the supernatant using SPRI beads. Rather than using 200uL of lysis buffer for tissue digestion as per the original protocol, our samples (e.g., small piece of dry blood spot on filter paper for Alert and Mitivik Island samples, dried pellet containing approximately 10mg of packed red blood cells for Svalbard samples, and a grain-of-rice-sized skin tissue sample from Aleutian Islands and Pribilof Islands) were digested in 200uL of

digestion buffer (100mM NaCl, 50mM Tris-HCl pH 8.0, 10mM EDTA, 0.5% SDS) and 10uL of 20mg/mL proteinase K overnight at room temperature on a nutator. We did not include zirconia-silica beads for the homogenization step as we had soft tissue samples not requiring cellular disruption. The resulting genomic DNA was suspended in 50uL TE buffer and stored at -80°C until use.

RNA sample collection, extraction and sequencing

Sixteen snow buntings were chosen at random for RNASeq from a pool of individuals housed at the avian facility of Université du Québec à Rimouski (UQAR), QC, Canada. These individuals were captured near Rimouski, QC, Canada as wintering birds. All individuals used in the current study were humanely euthanized via cervical dislocation for a separate sequencing project (approved by Animal Care Committee at UQAR (CPA-61-15-163 and CPA-68-17-186)), their whole brain was collected and immediately preserved in a high concentration salt buffer (Final concentrations: 70g ammonium sulfate/100mL, 25mM sodium citrate, 20mM EDTA, pH 5.2), stored at -20°C, and transferred to -80°C until RNA extraction. The sampling of the 16 individuals was equally spaced out from early March to the end of April 2018 to maximize transcriptome diversity in the brain tissue samples.

Total RNA was extracted from brain tissue using TRIzol Reagents (Life Technologies, Mississauga, ON, Canada) according to the manufacturer's protocol. The RNA pellet was resuspended in Nuclease-Free Water (Thermo Fisher Scientific, Mississauga, ON, Canada) and RNA quality was assessed using the Eukaryotic RNA 6000 Nano assay on a 2100 Bioanalyzer (Agilent Technologies Canada Inc., Mississauga, ON, Canada). We ensured that all samples had RIN $> 8.5$ and a 28S/18S rRNA ratio $> 0.8$ when preparing the RNA for sequencing for all sixteen birds. Final RNA aliquots were sent to the Genome Quebec Innovation Centre (McGill University, Montreal, QC, Canada) for 100bp paired-end sequencing in two lanes of an Illumina HiSeq4000 sequencer (Illumina Inc., San Diego, CA, USA).

RNA sequence analyses

Following sequencing, rRNA sequence reads were removed from the total raw sequence reads using SortMeRNA v2.1 (Kopyloca et al., 2012). Non-rRNA reads were then quality filtered using the default sliding window algorithm in Trimmomatic v0.38 (Bolger et al., 2014) to remove low-quality and adapter sequences. A *de-novo* transcriptome was assembled using fourteen out of sixteen samples using the default parameters with Trinity v2.8.4 (Hass et al., 2013) which included*in-silico* normalization for all reads. In the absence of a reference genome, and to ease the computational load for downstream data processing, the final reference transcriptome was assembled with only the longest isoform per transcript. Cleaned RNA sequence reads from all sixteen individuals were mapped to the final reference transcriptome using Burrow's Wheeler Alignment (BWA) v0.7.12 (Li & Durbin, 2009) (Supplementary Material S1). Additionally, we assigned Read Group tags to all samples as unique sample IDs for each file. Resulting SAM files were converted to BAM files and sorted using SAMtools v1.3 (Li et al., 2009). We then removed PCR duplicates using Picard Tools (*http://broadinstitute.github.io/picard*), the final BAM files were merged, and low-quality mapping and supplemental alignments were removed with SAMtools v1.3 (Li et al., 2009).

SNP characterization and SNP marker development

The mapping information for all reads from the *de-novo* assembled reference transcriptome was used for nucleotide variant discovery using the Broad Institute's Genome Analysis Tools Kit (GATK) pipeline (DePristo et al., 2011; Van der Auwera et al., 2013) to characterize and develop functional gene locus SNPs. We performed quality recalibration, indel realignment and variant discovery on filtered-merged combined sequences, post-alignment, using GATK v4.1.7.0 (McKenna et al., 2010). Furthermore, we applied hard filtering parameters recommended for RNASeq experiments to detect variants (DePristo et al., 2011; Van der Auwera et al., 2013).

We used GeneMarkS-T (Besemer et al., 2001) to characterize open reading frames in our reference transcriptome and used SNPEff (Cingolani et al., 2012) to annotate variants and characterize them as missense, synonymous, upstream or downstream variants. We used the Trinotate pipeline (Bryant et al., 2017) to

annotate all genes in our reference transcriptome and used LEMONS software (Levin et al., 2015) to predict intron splice junctions. It was important for us to identify the exon/intron boundaries to ensure that the SNP primers did not span introns since our goal was to use these primers to amplify genomic DNA.

By combining the SNPs (i.e., missense, synonymous, upstream or downstream variants) with gene annotation and predicted splice junction information, we were able to identify 11,378 useable SNPs (see Supplementary Material S2). From those, we selected 192 SNP loci representing genes expected to show local selection effects among our six populations. Broadly, seven gene function categories (energetics, lipid metabolism, immune response, stress response, nervous system development, reproduction and cell-housekeeping processes) were selected*a priori* based on their relevance and importance for the our study species (justifications for gene categories are shown in Supplementary Material S3, gene function categories for selected loci shown in Supplementary Material S4). We designed SNP primers to amplify a 100bp-150bp region surrounding the SNP of interest for the 192 loci using default settings with Primer3 v4.1.0 (Untergasser et al., 2012). Forward and reverse universal adapters (ACCTGCCTGCC & ACGCCACCGAGC, respectively) were added to the 5' end of the designed primers to allow for the addition of sequencing adapters and sample-specific barcodes for High Throughput Sequencing (HTS). All primers were tested in 12.5uL reactions containing 20mM Tris-HCl pH 8.0, 10mM KCl, 10mM $(NH_4)_2SO_4$, 2mM $MgSO_4$, 0.1% Triton X-100, 0.1mg/mL bovine serum albumin (BSA), 200 μM of each dNTP, 200nM of forward and reverse primers, 0.5U of Taq polymerase (Bio Basic Canada Inc., Markham, ON, Canada), and 0.5uL of genomic DNA. The PCR cycling conditions were: 2 min at 95°C; 20s at 95°C, 20s at 58°C, 30s at 72°C (32 cycles); and 2 mins at 72°C. Of the 192 primer sets, 72 either did not amplify genomic DNA, yielded non-specific amplification or produced an amplicon larger than 350bp: all of these were discarded from subsequent analyses. Details for the remaining 117 SNP primers are provided in Supplementary Material S4 in Supplementary Data.

Microsatellite and SNP marker genotyping

Microsatellite DNA marker data were used to assess population genetic structure (population connectivity), and they were used as the neutral marker controls for assessing divergence at the SNP loci. Briefly, all DNA samples were amplified at nine microsatellite loci with three PCR reactions: i) a first round of 20-cycle multiplex PCR (all primers combined) for preamplification of the DNA (this was done due to the small amount of DNA recovered from some samples) followed by ii) a second round of 30-cycle PCR with individual microsatellite primers, and iii) a final round of 5-cycle PCR to add fluorescent tags for fluorescence-based capillary electrophoresis. For each individual, we conducted the multiplex PCR in a 5uL reaction mixture containing 2.5uL of 2x Multiplex PCR Master mix (Qiagen Inc., Toronto, ON, Canada), 0.5uL of primer pool (10x primer mix containing 2uM each of all 9 primer pairs), and 1.0uL each of RNase-Free Water and template DNA. The amplification conditions were: 5 min at 95°C followed by 20 cycles of 30s at 95°C, 90s at 57°C, 30s at 72°C; and ending with 30 mins at 60°C. We diluted the PCR products 20-fold by adding 95uL of $ddH_20$. For the second round PCR, we amplified 2-4uL of the diluted multiplexed PCR product in a single-PCR reaction of 25uL which contained 2.5uL of 10x Taq buffer (20mM Tris-HCl pH 8.0, 10mM KCl, 10mM $(NH_4)_2SO_4$; Bio Basic Canada Inc., Markham, ON, Canada), 200uM each of dNTP, $MgSO_4$ (2uM), forward and reverse primers (2uM each), and 0.5U of Taq Polymerase (Bio Basic Canada Inc., Markham, ON, Canada). Thermocycling conditions were 95°C for 2 min; followed by 30 cycles of 95°C for 20s, locus-specific annealing temperature for 20s (56°C for CAM17, Lox8, Indigo29, SNBU682, and SNBU705; 58°C for Cuu28, POCC6, Ecit2, and CAM17), and 72°C for 30s, ending with 72°C for 2 min. For the final round of PCR, we used a PCR-based labelling technique where products from 1-4 loci were labelled with different dyes (6FAM, VIC, PET and NED; PCR conditions were identical to that of the second round of PCR with the exception of 5 cycles instead of 30) and combined with Hi-Di formamide (Applied Biosystems, Foster City, CA, USA) and a GeneScan LIZ600 size standard (Applied Biosystems, Foster City, CA, USA) for separation on a SeqStudio Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Each sample was genotyped using GeneMapper software v3.5 and verified by eye.

We genotyped all individuals at the selected SNP loci using HTS. The HTS library preparation was completed using two rounds of PCR; multiplex followed by barcoding (ligation) PCR. We first amplified the 117 SNP

loci using five separate multiplex PCRs for each sample (bird). Each multiplex PCR included 17-25 primer pairs (SNP locus groups shown in Supplementary Material S4). Multiplex PCR used the Qiagen Multiplex PCR Plus Kit (Qiagen Inc., Toronto, ON, Canada). For each multiplex group, we first made 10x primer pools containing all primers within that group at equimolar concentration of 0.2uM. Each 7uL multiplex reaction contained 3.5uL Multiplex PCR Plus Master mix, 0.7uL of the 10x primer pool, 1.3uL $ddH_2O$, and 1.5uL genomic DNA. The amplification conditions were: 5 min at 95°C followed by 28 cycles of 30s at 95°C, 90s at 58°C and 30s at 72°C followed by 10 mins at 68°C. We diluted the multiplexed PCR product 10-fold with $ddH_2O$. Next, PCR products from each of the five multiplex reactions were pooled for each individual and cleaned using Sera-Mag Speed Beads (Cytiva, Mississauga, ON, Canada) to remove unincorporated dNTPs, primers, primer dimers and PCR buffers. We then ligated individual barcode sequences and HTS adaptor sequences to the PCR products in a second (ligation) short-cycle PCR. The 20uL PCR reaction included: 10x Taq buffer (20mM Tris-HCl pH 8.0, 10mM KCl, 10mM $(NH_4)_2SO_4$; Bio Basic Canada Inc., Markham, ON, Canada), 2mM $MgSO_4$, 0.1mg/mL bovine serum albumin (BSA), 200uM of each dNTP, 200nM of forward and reverse primers, 0.5 U of Taq polymerase (Bio Basic Canada Inc., Markham, ON, Canada), and 10uL of pooled and cleaned multiplex PCR product. The PCR conditions for the ligation PCR were: 94°C for 2 min, followed by 6 cycles of 94°C for 30s, 60°C for 30s and 72°C for 60s, followed by 72°C for 5 min. This second PCR ligated a "barcode" sequence that allowed us to identify each sample for allocating sequence data to specific individuals post-sequencing. The barcoded products were pooled and gel-extracted using the GenCatch Gel Extraction Kit (Epoch Life Science Inc., Sugar Land, TX, USA) as per manufacturer's instructions. Purified pooled product was analyzed on an Agilent 2100 Bioanalyzer using a High Sensitivity chip (Agilent Technologies Canada Inc., Mississauga, ON, Canada) to verify the size and concentration of the library amplicons. Finally, the library was diluted to approximately 60pM and sequenced using Ion PGM Hi-Q chemistry in an Ion Chef System (Thermo Fisher Scientific Inc., Streetsville, ON, Canada). Specifically, the library was sequenced using an Ion 318 Chip Kit with an Ion PGM Sequencing 400 Kit (Thermo Fisher Scientific, Mississauga, ON).

SNP HTS Bioinformatics

After HTS sequencing of the pooled SNP PCR amplicons, we used the FASTX Toolkit (Gordon & Hannon, 2010) and its Barcode Splitter script to demultiplex the sequences. We then trimmed off the sequencing adapters and barcodes from all reads using CUTADAPT v1.11 (Martin, 2011) and subsequently mapped the resulting PCR-amplified sequences to our reference transcriptome using BWA v0.7.12 (Li & Durbin, 2009) to identify the genes containing the amplified SNP regions. To genotype all individuals at target SNP loci, we used FreeBayes (Garrison & Marth, 2012), a Bayesian genetic variant detector. Since FreeBayes detects many other variants such as small multi-nucleotide polymorphisms (MNPs), insertions and deletions (indels), composite insertions, and substitutions, we discarded such variants using VCFtools (Danecek et al., 2011). Next, we refined the VCF file through a filtration step which excluded the SNP locus markers that were called in less than 30% of individuals (16 out of 117 SNPs) and the individuals that were missing more than 10% of their genotypes (2 out of 221 individuals). Lastly, we only kept one SNP per amplicon (i.e., the original SNP used to design the primers for that amplicon) for further analyses to avoid any bias resulting from including multiple (linked) SNPs per amplicon. As a result, we had 101 SNP genotypes across 219 individuals for further analyses.

Population genetic analyses

*Testing for temporal effects*

Because we had individuals collected across multiple years for most of our study populations, we first tested for temporal effects (i.e., a year effect) on allele frequencies. We conducted separate Fisher's exact tests of allele frequency variation for the microsatellite marker data for multi-year samples from Alert, Svalbard, Utqiagvik and Mitivik Island using the genepop package (Rousset, 2008) in R v1.2.5 (R Core Team 2016). Since $p$ -values ranged from 0.08-0.50 for each population, we concluded that there were no temporal effects, hence we combined samples from multiple years for the Alert, Svalbard, Utqiagvik, and Mitivik Island populations.

*Testing for Alaskan island population neutral divergence*

The island populations of Alaska (Attu, Adak, and Pribilof islands) are geographically clustered (Figure 1), allowing dispersal among the islands possibly resulting in a single metapopulation. We thus tested these three populations for neutral population divergence. Based on the results of Fisher's test, we combined Attu and Adak Island samples from 1999 forming the population 'Aleutian Islands' since there were no significant differences in neutral allele frequencies ($p = 0.14$). We retained Pribilof Islands individuals as a separate population for further analyses as it had a significantly different ($p< 0.00001$) neutral allele frequency distribution from the Aleutian Islands samples. These two Alaskan island populations combined with the other four populations, resulted in a total of six populations for downstream analyses (Table 2).

*Population genetic divergence*

We assessed population differentiation using both neutral microsatellite and functional SNP markers using pairwise Fisher's exact test of allele frequency variation in the genepop package (Rousset, 2008) in R. We also estimated pairwise $F_{ST}$ for both marker types using GENODIVE (version 3.0) (Miermans, 2020). We corrected all$p$ -values for multiple comparisons using the sequential Bonferroni procedure (Rice, 1989) where necessary.

*Neighbour-joining cluster analyses*

To visually assess the pattern of population genetic divergence for the two marker types (microsatellite and SNP loci), we performed unrooted neighbour-joining (NJ) cluster analyses with Cavalli-Sforza and Edward's (1967) chord distance ($D_c$ ) using the 'ape' package (Paradis & Schliep, 2019) in R. Chord distance was used as it is expected to provide better tree topology estimation for closely related populations, although it may compromise branch length estimation (Angers & Bernatchez, 1998). The percent support for branches was estimated using bootstrapping, with replacement, among loci using 10,000 permutations in the 'poppr' package (Kamvar et al., 2014) in R.

*Selection signatures at candidate loci*

To detect a signature of selection at functional SNP loci, it is important to separate the effects of genetic drift from selection. For this purpose, we used the microsatellite markers to estimate the neutral allelic distribution; it is expected that both functional SNP loci and microsatellites undergo change due to genetic drift and gene flow, but only SNP loci are expected to be under selection due to potential local habitat-specific environmental conditions.

*Global selection at candidate loci*

To assess whether candidate loci were under divergent selection across the six populations, we compared global estimates of Hedrick's$G'_{ST}$ (Hedrick, 2005), calculated using the 'diveRsity' package (Keenan et al., 2013) in R, between neutral microsatellite and functional SNP markers. Hedrick's$G'_{ST}$ is suitable for comparing genetic divergence measures among different marker types since it standardizes for heterozygosity (Hedrick, 2005). To assess an overall signature of selection at SNP loci relative to microsatellite markers across all populations, we first developed a 'neutral range' mean$G'_{ST}$ with 99% confidence intervals (CI) for the microsatellite marker $G'_{ST}$ values using the 'diffCalc' function of the R package 'diveRsity'. Specifically, we used bias-corrected bootstrapping across microsatellite loci to estimate the neutral CI range, representing neutrality expectation (presumably due to genetic drift) based on the $G'_{ST}$ distribution of microsatellite markers. Next, we determined whether the$G'_{ST}$ values for individual SNP loci fell outside of the neutral ranges, as such loci are likely to be under selection. Since the calculated neutral range for $G'_{ST}$ did not include zero, we were able to identify SNP genes under stabilizing (lower than neutral expected) and divergent (higher than neutral expected) selection.

*Pairwise selection at SNP loci*

While it is possible for individual functional SNP markers to show a global selection signature, others may only show signatures of divergent or stabilizing selection at the population level due to specific differences in

local conditions among individual populations. To assess genetic divergence patterns among pairs of snow bunting populations, we calculated pairwise estimates of $G'_{ST}$ using both microsatellite and SNP genotype data and compared the SNP loci pairwise $G'_{ST}$ values with the presumed-neutral microsatellite loci range (created using 'diffCalc' function's bias-corrected bootstrapping loci approach as explained above) at the 99.9% CI to detect signatures of divergent and stabilizing selection. We used higher CI (99.9% versus 99% neutral CI used in global comparison) to avoid detection of false positives for pairwise comparisons since we are assessing 101 SNPs (useable SNPs post-filtration step; see "Bioinformatics" section above for details) and fifteen pairwise population comparisons. Corrections for multiple comparisons were not necessary as neutral range was individually developed for each comparison. We first combined all the results from the pairwise comparisons to investigate overall levels of genetic drift and selection, and also conducted a Chi-squared test to assess whether the pattern of selection signatures differed across the seven gene function categories. However, for some population pairs it was not possible to identify SNPs under stabilizing selection since the neutral $G'_{ST}$ range for that pairwise comparison included zero. As such, we have reported the SNP loci showing likely signals of divergent selection for all fifteen pairwise comparisons, but stabilizing selection for only nine of fifteen comparisons. For the six comparisons which had neutral ranges that included zero, the SNP loci with $G'_{ST}$ values less than expected neutral range (i.e., negative $G'_{ST}$ values) were identified as "undetermined".

To gain further insight into specific genes that showed evidence for divergent selection, we explored the function of selected SNP loci with $G'_{ST}$ values that had no "undetermined" classifications across any of the fifteen pairwise comparisons. Therefore, each SNP locus in this subset was identified as either under neutral processes, stabilizing selection or divergent selection for all fifteen pairwise comparisons. This approach allowed us to assess the selection status of divergent SNP loci across all other population pairs – this allows the comparison of the role of these functional markers across all other population comparison(s) to highlight specific differences, allowing us to identify specific genes contributing to population divergence and local adaptation.

*Selection signature and variant type*

To characterize the role of SNP variant type (i.e., missense, synonymous, downstream or upstream), we determined the proportion of SNPs that showed signatures of neutral processes or selection at the global and pairwise levels (with combined data across all fifteen comparisons) within each variant type. Given our functional SNPs were derived from transcribed sequences, we would expect selection to be more common among missense variants, as they would result in a different amino acid sequence in the protein.

**Results**

Microsatellite vs. SNP marker characteristics

We developed nine microsatellite markers (Supplementary Material S5) and applied them across all individuals to assess reproductive isolation and establish "neutral" control data for functional SNP locus divergence. We also developed 117 functional SNP loci (Supplementary Material S4) from a *de-novo* transcriptome for snow buntings which were expected to show local selection effects among breeding populations based on their putative gene function. The microsatellite panel was more polymorphic than the SNP panel: observed heterozygosity for microsatellite markers was higher (0.345-0.708) than for SNP locus markers (0.098-0.111) (Supplementary Material S6). We were able to successfully extract DNA for all 221 samples across six populations for microsatellite and SNP marker genotyping.

RNA sequencing and SNP marker development

RNA sequencing produced more than 720 million pair-end reads from 16 birds (Supplementary Material S1), 14 of which were used to *de-novo* assemble the transcriptome containing 866.3Mb assembled into 534,815 trinity 'genes,' and we used 373Mb of sequence data to assemble the novel transcriptome. The resulting transcriptome was used to characterize 11,378 SNP variants, approximately one variant per 32,782 bp of reference transcriptome. We first removed variants in transcripts with no valid start codon from the identified

SNPs, as such variants are likely from incomplete or non-coding transcripts. This resulted in 9,756 useable sequence variants (see Supplementary Material S2 for detailed summary statistics for SNP characterization). After optimization of multiplex groups, we retained 117 SNP loci (out of 192) to be genotyped in five multiplex groups (Supplementary Material S4).

Microsatellite and SNP marker genotyping

All individuals were successfully genotyped at all microsatellite loci. For SNP genotyping, 101 out of 117 loci were genotyped in at least 70% of the individuals (our threshold for inclusion in the analyses), and 219 of the 221 individual birds were successfully genotyped at $> 90\%$ of these 101 SNP loci and were retained for analyses. Thus, downstream analyses for the SNP loci were conducted using 101 SNP loci genotyped for 219 individuals. The final 101 SNPs consisted of 52 downstream, 11 upstream, 28 missense, and 10 synonymous variants.

Population genetic analyses

*Population genetic divergence*

The microsatellite marker global $F_{ST}$ value was 0.031 across all populations. The pairwise $F_{ST}$values across all fifteen comparisons ranged from -0.0001 to 0.100 (Table 3). The Fisher's exact test for microsatellite allele frequency distributions showed a highly significant population effect ($p$ [?] 0.001) in all but one population pairwise comparison (Utqiagvik/Svalbard; uncorrected $p = 0.011$), although that comparison was significant prior to Bonferroni correction (Table 3). The microsatellite pairwise $F_{ST}$ values also showed highly significant population differentiation in 13/15 population pairs ($F_{ST}$ : 0.009-0.100, $p$ [?] 0.012) comparisons (Table 3). The population pairs Alert/Mitivik Island and Utqiagvik/Svalbard ($F_{ST}$ : -0.0001 for both pairs, uncorrected $p$ -values of 0.564 and 0.464, respectively) did not show significant population differentiation before or after the sequential Bonferroni correction (Table 3). Combined results from the Fisher's exact test and $F_{ST}$ estimation at neutral markers provide evidence of significant divergence between all population pairs, with the exception of Utqiagvik/Svalbard and Alert/Mitivik Island population pairs which exhibited weak isolation.

The SNP marker global $F_{ST}$ value was 0.022 across all populations. The SNP marker pairwise $F_{ST}$values across all fifteen comparisons ranged from 0.004 to 0.053 (Table 4). The Fisher's exact test for functional SNP marker allele frequency distribution showed highly significant population differentiation in 9/15 population pairs ($p$ [?] 0.009), with non-significant differentiation for: Aleutian Island/Pribilof Island, Alert/Utqiagvik, Alert/Mitivik Island, Alert/Svalbard, Utqiagvik/Mitivik Island, and Mitivik Island/Svalbard (uncorrected $p$ -values: 0.56-0.88; Table 4). The SNP marker pairwise $F_{ST}$ estimates matched the Fisher's exact test results as the same population pairs (listed above) showed significant divergence ($F_{ST}$ : 0.024-0.053, $p$ [?] 0.003; Table 4) and the remaining pairs showed non-significant divergence ($F_{ST}$ : 0.004-0.013,$p$ [?] 0.039; Table 4). Broadly, the combined SNP Fisher's exact test and $F_{ST}$ estimation did not show significant genetic differentiation between the non-migratory populations (Aleutian and Pribilof islands), and among a majority (exception: Utqiagvik/Svalbard comparison) of the migratory populations (Alert, Utqiagvik, Mitivik Island, and Svalbard). However, all migratory - non-migratory population comparisons did show significant differentiation with the SNP marker data. The Utqiagvik/Svalbard and Alert/Mitivik Island population pairs had significant levels of SNP marker differentiation, although they were not significantly divergent based on neutral (microsatellite) marker data. Overall, our analyses show substantial genetic divergence among our six sampled populations.

*Neighbour-joining cluster diagrams*

The neighbour-joining (NJ) cluster diagrams (Figure 2) based on microsatellite and SNP genotypes show similar overall patterns of divergence. The microsatellite marker data show three strongly supported clusters (100% branch support): Utqiagvik & Svalbard, Alert & Mitivik Island, and Aleutian & Pribilof islands; although Aleutian and Pribilof islands also show strongly supported divergence (100% branch support) from each other (Figure 2a). On the other hand, the SNP genotype data do not show as strong support for

population clustering among the six populations. However, the Aleutian & Pribilof islands, and Utqiagvik & Alert population pairs show strong patterns of divergence (99% and 100% branch support) between these populations and the others based on SNP marker data (Figure 2b).

*Selection signatures at SNP loci*

*Global selection at SNP loci*

The global $G'_{ST}$ values for 9 microsatellite and 101 SNP loci across the six populations (221 individuals) were 0.203 and 0.0393, respectively. Per-locus $G'_{ST}$ values ranged from 0.0454 to 0.535 for the microsatellite markers, and from -0.0789 to 0.267 for the SNP markers. Global differentiation patterns showed 94 out of 101 SNP loci to be consistent with stabilizing selection, as their level of divergence was lower than the neutral expectation (Figure 3; Supplementary Material S7). The seven remaining SNP loci showed divergence levels consistent with neutral processes (Figure 3; Supplementary Material S7). The SNP loci showing a global neutral pattern of divergence belonged to four gene function categories: immune response (1 SNP); lipid metabolism (2 SNPs); nervous system development (1 SNP) and reproduction (3 SNPs). We did not detect any SNP loci showing a population divergence pattern consistent with divergent selection across all six populations (i.e., global $G'_{ST}$ values), possibly due to inconsistent patterns of selection acting among the populations, making a pairwise analysis important to assess divergence patterns.

*Pairwise selection at SNP loci*

The overall pairwise $G'_{ST}$ values ranged from 0.007 to 0.4508 for the microsatellite markers, and from 0.0076 to 0.0655 for the SNP locus markers among the fifteen comparisons, depending on the population compared (Supplementary Material S8).

Controlling for the neutral processes (99% CI using microsatellite genotype data) for all fifteen pairwise population comparisons of the 101 SNP loci (1286 possible $G'_{ST}$ values), we mostly observed signatures of stabilizing selection (51%) and genetic drift (38%), followed by undetermined (7%) and divergent selection (4%). We observed roughly equivalent patterns of divergence among the gene function categories (Supplementary Material S9). The distribution of the types of selection did not differ significantly among different functional categories ($\chi^2 = 20.33$, $p = 0.32$).

We detected signatures of stabilizing selection across all SNP loci in all but six population comparisons (Figure 4), where the neutral $G'_{ST}$ ranges included zero (Alert/Utqiagvik, Alert/Mitivik Island, Alert/Svalbard, Utqiagvik/Mitivik Island, Utqiagvik/Svalbard, and Mitivik Island/Svalbard). For these six comparisons, the SNP loci that had $G'_{ST}$ values less than the neutral expected range (i.e., negative $G'_{ST}$ values) were identified as "undetermined". Therefore, we are likely underestimating overall stabilizing selection effects. Overall, we observed that 67.5% - 96.3% of the SNP loci are under stabilizing selection among the nine pairwise population comparisons where we were able to test for signatures of stabilizing selection (Figure 4, Supplementary Material S8).

We did not observe any signatures of divergent selection in six (Aleutian Islands/Alert, Aleutian Islands/Utqiagvik, Aleutian Islands/Mitivik Island, Aleutian Islands/Pribilof Islands, Aleutian Islands/Svalbard, and Alert/Utqiagvik) out of fifteen pairwise population comparisons (Figure 4). For the remaining nine population pairs, we observed that 1.09%-23.1% of the SNP loci exhibited divergent selection (Figure 4, Supplementary Material S8). The Utqiagvik/Svalbard population comparison showed the most divergence (23.1%) based on our selected functional locus SNPs, followed by Pribilof Islands/Svalbard and Alert/Mitivik Island population comparisons which exhibited directional divergence at 7.23%, and 6.60% of tested SNP loci, respectively (Supplementary Material S8).

To compare selection signatures across all pairwise population comparisons, SNP marker data would have to be 100% available (i.e., not classified as "undetermined" in any comparison). Thus, only a minority (11/101) of SNP marker loci could be broadly compared across all pairwise comparisons (Figure 5; Table 5). Based on those 11 SNP marker loci, high levels of stabilizing selection signatures were generally observed when the non-migratory populations (Aleutian & Pribilof islands) were compared to each other or with

migratory populations (Alert, Utqiagvik, Svalbard, and Mitivik Island), whereas comparisons among the migratory populations showed mixed signatures of neutral processes and divergent selection depending on the SNP locus (Table 5). Of these 11 SNP loci, 7 showed evidence of divergent selection in at least one population comparison, while the Utqiagvik/Svalbard population pair comparison showed 6 of these 11 loci under divergent selection (Table 5). Broadly, the divergent genes from pairwise comparisons in this subset were associated with cellular housekeeping, lipid metabolism, nervous system development, reproduction, and stress (Table 5, Supplementary Material S10).

## Discussion

Heterogeneous environmental conditions across time and space can drive adaptive population divergence among even populations that are only partially reproductively isolated (Hereford, 2009). Here we assessed population structure and functional divergence among six geographically isolated breeding populations of Arctic-breeding snow buntings. Both our neutral (microsatellite DNA loci) and functional (coding-gene SNPs) genetic marker data show substantial population divergence among all populations, likely indicative of some level of reproductive isolation. Furthermore, we demonstrated that the observed population differentiation patterns at selected known-function SNPs likely resulted primarily from stabilizing, but also divergent, selection at the candidate loci. The global divergence analyses showed strong evidence of stabilizing selection across all populations, which is not surprising given the expected canalization of the vital functional gene loci chosen for this study. At the pairwise population comparison level, our functional marker results show signatures of both neutral drift and selection, with functional divergent selection observed at some SNP loci. Such selection effects likely reflect local adaptation of different snow bunting populations to their local environments (e.g., both wintering and migratory route selection pressures).

Although both of our marker types yielded broad spatial divergence patterns separating resident and migratory populations, finer genetic structure differed based on the marker type. A prime example is the lack of divergence between the Alert & Mitivik Island and the Utqiagvik & Svalbard populations, indicative of gene flow between both population pairs. The Alert & Mitivik Island population pair exhibited similar microsatellite and SNP divergence patterns; however, the Utqiagvik & Svalbard population pair curiously exhibited significant SNP divergence, but no divergence with the microsatellite markers. Such a pattern is consistent with a strong selection signature at the SNP loci, despite gene flow. Our observation of gene flow between the Alert and Mitivik Island populations is a new finding, but supports previous work in this species using stable hydrogen isotope analysis and light-level geolocator tracking that suggested two parallel migratory systems for the high and low Eastern Canadian Arctic with Hudson Bay as a migratory divide (Macdonald et al., 2012). Thus, it is possible that the Alert population follows the same migratory route as the Mitivik Island population (i.e., to the West of Hudson Bay, NU, Canada; Macdonald et al., 2012). Additionally, because the Mitivik Island population has been shown to winter in the Canadian provinces of Saskatchewan and Alberta (Macdonald et al., 2012), it is further possible the individuals in these populations winter together, or even mix during spring migration to the breeding grounds. On the other hand, the presence of potential gene flow between the Utqiagvik and Svalbard populations is surprising, given the geographic distance between the two sample sites. Although we do not currently know where birds from the Utqiagvik breeding population overwinter, recent tracking of the Svalbard buntings using light-level geolocators (GLS) indicate they overwinter in the Asian Western Siberian Steppe where they utilize the high abundance of grain croplands and face very little interspecific competition (Snell et al., 2018). This could also be true for individuals breeding at Utqiagvik, providing a potential mechanism for gene flow between the two populations. If true, Svalbard birds would be migrating west in the Fall, and Utqiagvik birds east in the fall, to share wintering grounds in the Asian Western Siberian Steppes. Such cross-hemisphere migration has been demonstrated in a similarly-sized songbird, the northern wheatear (*Oenanthe oenanthe* ), using GLS (Bairlein et al., 2012). Nevertheless, a detailed migration study is needed for Utqiagvik snow buntings to empirically test the possibility of a shared wintering ground.

While fairly spatially distant snow bunting populations showed genetic connectivity, we surprisingly found significant differentiation between the two non-migratory populations in Alaska (Aleutian and Pribilof is-

lands), based on microsatellite data. These populations exhibited substantial divergence despite being geographically close (Figure 1). Migratory life history is a critical component of genetic population structure; high dispersal rates result in genetically homogeneous populations, whereas restricted dispersal allows for development of genetically differentiated populations (Milgroom, 2015) due to elevated isolation and drift (Arguedas & Parker, 2000; Winker et al., 2000). Our results support this pattern in snow buntings, with both microsatellite and SNP data clustering resident and migratory populations separately. Overall, in addition to identifying significant global population differentiation, the genetic markers used in this study add to our knowledge of migratory connectivity patterns among breeding snow bunting populations. More importantly, our results shed light on the vulnerability of common wintering grounds for some populations should these sites face human-induced stressors such as habitat degradation.

Although local adaptation is predicted in populations experiencing divergent local selection pressures, it is rarely directly demonstrated empirically since it requires common-garden or reciprocal transplant experiments (Kawecki & Ebert, 2004) which are not practical for many wild populations (Blanquart et al., 2013). Avian species have been shown to exhibit population-level patterns of variation in timing of migration and brood initiation (Gu et al., 2021; Wanamaker et al., 2020), body size and mass (Blondel et al., 2006), song (Badyaev et al., 2008), personality (Mouchet et al., 2021), and plumage (Antoniazza et al., 2010) that have been identified as possible locally adapted traits. Although those studies provide strong indirect evidence of local adaptation, they may reflect phenotypic plasticity or even flexibility. In this study, we used SNP gene loci expected to be under selection based on their putative gene function, and as such likely to reflect environmental and ecological differences driving genetic variation among populations (Wellband et al., 2018). Although more than a quarter (28/101) of our SNP markers are missense variants, all were in very strong linkage disequilibrium with the target known-function genes, making our study unique from other SNP-based studies in birds that used random SNPs located in both coding and non-coding regions of the genome (e.g., Tiffin & Ross-Ibarra, 2014; Pardo-Diez et al., 2015). Our *a priori* choice of candidate gene function improves the likelihood of detecting functional patterns of population differentiation consistent with local adaptation in breeding snow bunting populations. Local adaptation has implications for management and conservation aimed at preserving local genetic diversity, especially as Arctic-migratory species continue to face strong effects of climate change and other anthropogenic stressors worldwide.

Generally, locally adapted populations are predicted to exhibit significantly higher (for divergent selection) or lower (for stabilizing selection) genetic differentiation than expected under neutral processes (Schlötterer, 2002; Hoban et al., 2016). Consistent with this idea, a high proportion of our candidate loci exhibited evidence of being under selection at candidate functional loci. Only a handful of previous studies have assessed patterns of divergence at both coding (i.e., functional) and non-coding (i.e., presumed neutral) marker loci, to interpret selection patterns in migratory bird species. Furthermore, the majority of those studies used randomly selected genome-wide SNPs and they inferred divergent selection at functional loci based on presumed linkage disequilibrium. For example, Zhan et al. (2015) used a targeted approach comparing thirteen wild populations of saker falcon (*Falco cherrug* ) across Eurasia using SNP data and inferred that the *MHC* genes were under directional selection ($F_{ST} > 0.5$), with the remaining candidate SNPs showing signatures of stabilizing selection or drift. Although SNP-based selection studies are becoming more common in migratory bird species (e.g., Ruegg et al., 2014; Bay et al., 2021; Larison et al., 2021; Ruegg et al., 2021), there have only been two such studies on Arctic-breeding migratory birds, both of which employed a random SNP approach and reported no or low levels of selection. For example, Colston-Nepali et al. (2020) used RAD-seq to genotype six breeding colonies of northern fulmar (*Fulmarus glacialis* ) at 6,614 genome-wide SNPs; however, no outlier loci were identified. Similarly, Tigano et al. (2017) used 2220 genome-wide SNPs across five colonies of Arctic-breeding thick-billed murres and found ˜6% outlier SNPs (only 28% of which showed divergent selection). However, random SNP surveys do not have *a priori* SNP gene function and often it is difficult to assign function. For example, Tigano et al. (2017) found only 6 of their 111 identified outlier SNP loci could be assigned a putative function (GO term). In contrast, we detected strong signatures of stabilizing selection at known-function SNP loci, with some showing evidence of divergent selection in pairwise population comparisons. The high level of stabilizing selection likely results from canalization of the

genes associated with our SNP loci as they were selected to reflect critical organismal and cellular functions.

Our candidate SNP loci that exhibited consistent patterns of divergent selection may reflect local adaptation. For example, ACVR2A (divergent between Utqiagvik and Mitivik, and between Utqiagvik and Svalbard population pairs) codes for a receptor that is involved in the induction of adipogenesis and growth (Donaldson et al., 1992). It has been shown that fat reserves aid in thermogenesis, cold endurance (Vézina et al., 2012; Montgomerie & Lyon, 2020) and modulate the adrenocortical response to environmental stress (Wingfield et al., 2004). Those functions not only facilitate successful breeding in Arctic conditions, but also help snow buntings survive challenging conditions (i.e., scarce food resources and cold temperatures) on arrival as they prepare for breeding (Le Pogam et al., 2021). PTPRZ1 (divergent between Pribilof Islands and Svalbard, and between Utqiagvik and Svalbard population pairs) is mainly involved in development of myelinating oligodendrocytes and is thought to play a role in the establishment of contextual memory and learning (The UniProt Consortium, 2015). The role of spatial memory and learning has been explored in passerines for behaviours associated with food hoarding (Hitchcock & Sherry, 1990; Brodin, 1994; Healy & Krebs, 1996; Smulders & DeVoogd, 2000) and vocal communication (Nottebohm, 1999; Zeigler & Marler, 2004). Our candidate gene loci implicated in divergent selection warrant further examination of allelic variation at the loci under selection in snow buntings and possibly other migratory avian species.

Arctic-breeding migratory bird species utilize diverse breeding and over-wintering habitats, resulting in a substantial variation in experienced local abiotic factors such as temperature, wind, precipitation, and snow cover (among others; Martin & Wiebe, 2004; Wingfield et al., 2004). Variation likely drives selection pressures on Arctic-breeding birds that have short breeding times and high energetic demands (Le Pogam et al., 2021), which may contribute to local adaptation (Macdonald et al., 2012; Tigano et al., 2017; Snell et al., 2018). To our knowledge, this is the first study to investigate global population structure and genetic divergence consistent with local adaptation in a circum-polar arctic-breeding songbird. Consistent with our predictions, we observed strong evidence of genetic isolation coupled with some SNP loci showing divergent selection signatures. However, we observed stabilizing selection signatures across most SNP loci, and while high levels of stabilizing selection are reasonable given the nature of our candidate genes, we did not expect such a dominant role for apparent stabilizing selection among our SNP loci. Identifying population genetic structure at a pan-Arctic scale in snow buntings is especially important for their conservation as they face severe effects of climate change in their breeding areas coupled with other anthropogenic stressors in their overwintering areas (Walker et al., 2015). Our analyses of population divergence using both neutral and functional markers provide conservation-related data particularly valuable for species such as the snow bunting, which, because of its migratory life history, experiences diverse management jurisdictions, habitat degradation, and survival challenges.

## Acknowledgements

We would like to thank Environment and Climate Change Canada (ECCC), the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chairs Program, the Nunavut Wildlife Management Board, the Northern Scientific Training Program, the Polar Continental Shelf Project, POLAR Canada, the Canadian Network of Centres of Excellence (ArcticNet), the Government of Ontario, the University of Windsor and the Arctic Institute of North America for logistical support, research and/or personal funding. We would also like to thank the snow bunting field teams at Alert, Utqiagvik, Mitivik Island, Svalbard, Aleutian Islands, and Pribilof Islands for their help in sample collection. We thank Stacey McIntyre and Calvin Kellendonk for their valuable assistance during early stages of this project.

## References

Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* , *44* , 367-388.

Angers, B., & Bernatchez, L. (1998). Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis* , Salmonidae) populations from microsatellites. *Molecular Biology and Evolution, 15* (2), 143-159.

Antoniazza, S., Burri, R., Fumagalli, L., Goudet, J., & Roulin, A. (2010). Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (*Tyto alba* ). *Evolution: International Journal of Organic Evolution* , *64* (7), 1944-1954.

Arguedas, N., & Parker, P. G. (2000). Seasonal migration and genetic population structure in house wrens. *The Condor* , *102* (3), 517-528.

Atkins, K. E., & Travis, J. M. J. (2010). Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology* , *266* (3), 449-457.

Badyaev, A. V., Young, R. L., Oh, K. P., & Addison, C. (2008). Evolution on a local scale: developmental, functional, and genetic bases of divergence in bill form and associated changes in song structure between adjacent habitats. *Evolution: International Journal of Organic Evolution* , *62* (8), 1951-1964.

Bairlein, F., Norris, D. R., Nagel, R., Bulte, M., Voigt, C. C., Fox, J. W., Hussell, D. J. T., & Schmaljohann, H. (2012). Cross-hemisphere migration of a 25 g songbird. *Biology Letters* , *8* (4), 505-507.

Barter, R. L., & Yu, B. (2018). Superheat: An R package for creating beautiful and extendable heatmaps for visualizing complex data.*Journal of Computational and Graphical Statistics, 27* (4), 910-922.

Bay, R. A., Karp, D. S., Saracco, J. F., Anderegg, W. R., Frishkoff, L. O., Wiedenfeld, D., ... & Ruegg, K. (2021). Genetic variation reveals individual-level climate tracking across the annual cycle of a migratory bird. *Ecology Letters* , *24* (4), 819-828.

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology* , *13* (4), 969-980.

Bensch, S., Price, T., & Kohn, J. (1997). Isolation and characterization of microsatellite loci in a *Phylloscopus*warbler. *Molecular Ecology* , *6* (1), 91-92.

Besemer, J., Lomsadze, A., & Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Research* , *29* (12), 2607-2618.

Blanquart, F., Kaltz, O. Nuismer, S. L., & Gandon S. (2013). A practical guide to measuring local adaptation. *Ecology Letters* ,*16* (9), 1195-1205.

Blondel, J., Thomas, D. W., Charmantier, A., Perret, P., Bourgault, P., & Lambrechts, M. M. (2006). A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics.*Bioscience* , *56* (8), 661-673.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* , *30* (15), 2114-2120.

Bonin, C. A. (2021). Genetic Consequences of Dispersal, Philopatry and Reproductive Behaviors. In *Ethology and Behavioral Ecology of Otariids and the Odobenid* (pp. 223-241). Springer, Cham.

Both, C., Bouwhuis, S., Lessells, C. M., & Visser, M. E. (2006). Climate change and population declines in a long-distance migratory bird. *Nature* , *441* (7089), 81-83.

Brodin, A. (1994). The disappearance of caches that have been stored by naturally foraging willow tits. *Animal Behaviour* , *47* (3), 730-732.

Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., ... & Bateman, J. (2017). A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Reports* , *18* (3), 762-776.

Butcher, G. S., & Niven, D. K. (2007). Combining data from the Christmas Bird Count and the Breeding Bird Survey to determine the continental status and trends of North America birds.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* , *6* (2), 80-92.

Colston-Nepali, L., Provencher, J. F., Mallory, M. L., Franckowiak, R. P., Sun, Z., Robertson, G. J., & Friesen, V. L. (2020). Using genomic tools to inform management of the Atlantic northern fulmar. *Conservation Genetics* , *21* (6), 1037-1050.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics* , *27* (15), 2156-2158.

Dawson, D. A., Horsburgh, G. J., Kupper, C., Stewart, I. R., Ball, A. D., Durrant, K. L., ... & Krupa, A. P. (2010). New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility–as demonstrated for birds. *Molecular Ecology Resources* , *10* (3), 475-494.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., ... & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* , *43* (5), 491-498.

Donaldson, C. J., Mathews, L. S., & Vale, W. W. (1992). Molecular cloning and binding properties of the human type II activin receptor. *Biochemical and Biophysical Research Communications* , *184* (1), 310-316.

Falconer, C. M., Mallory, M. L., & Nol, E. (2008). Breeding biology and provisioning of nestling snow buntings in the Canadian High Arctic. *Polar Biology* , *31* (4), 483-489.

Fossoy, F., Stokke, B. G., Kasi, T. K., Dyrset, K., Espmark, Y., Hoset, K. S., ... & Moksnes, A. (2014). Reproductive success is strongly related to local and regional climate in the Arctic snow bunting (*Plectrophenax nivalis* ). *Polar Biology* , *38* (3), 393-400.

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907* .

Gibbs, H. L., Tabak, L. M., & Hobson, K. (1999). Characterization of microsatellite DNA loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus* ). *Molecular Ecology* , *8* , 1551-1551.

Gordon, A., & Hannon, G. J. (2010). Fastx-toolkit. FASTQ/A short-reads preprocessing tools (unpublished), http://hannonlab.cshl.edu/fastx_toolkit.

Gu, Z., Pan, S., Lin, Z., Hu, L., Dai, X., Chang, J., ... & Zhan, X. (2021). Climate-driven flyway changes and memory-based long-distance migration. *Nature* , *591* (7849), 259-264.

Guindre-Parker, S., Gilchrist, H. G., Baldo, S., Doucet, S. M., & Love, O. P. (2013a). Multiple achromatic plumage ornaments signal to multiple receivers. *Behavioral Ecology* , *24* (3), 672-682.

Guindre-Parker, S., Gilchrist, H. G., Baldo, S., & Love, O. P. (2013b). Alula size signals male condition and predicts reproductive performance in an Arctic-breeding passerine. *Journal of Avian Biology* , *44* (3), 209-215.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... & MacManes, M. D. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* , *8* (8), 1494-1512.

Healy, S. D., & Krebs, J. R. (1996). Food storing and the hippocampus in Paridae. *Brain, Behavior and Evolution* , *47* (4), 195-199.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* , *59* (8), 1633-1638.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist* , *173* (5), 579-588.

Hitchcock, C. L., & Sherry, D. F. (1990). Long-term memory for cache sites in the black-capped chickadee. *Animal Behaviour* , *40* (4), 701-712.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., ... & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist* , *188* (4), 379-397.

Hoset, K. S., Espmark, Y., Fossoy, F., Stokke, B. G., Jensen, H., Wedege, M. I., & Moksnes, A. (2014). Extra-pair paternity in relation to regional and local climate in an Arctic-breeding passerine. *Polar Biology* , *37* (1), 89-97.

Job, J. R., Kohler, S. L., & Gill, S. A. (2015). Song adjustments by an open habitat bird to anthropogenic noise, urban structure, and vegetation. *Behavioural Ecology* , *27* (6), 1734-1744.

Jonzen, N., Linden, A., Ergon, T., Knudsen, E., Vik, J. O., Rubolini, D., ... & Stenseth, N. C. (2006). Rapid advance of spring arrival dates in long-distance migratory birds. *Science* , *312* (5782), 1959-1961.

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

Kawecki, T. J., & Ebert D. (2004). Conceptual issues in local adaptation. *Ecology Letters* , *7* (12), 1225-1241.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodohl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* , *4* (8), 782-788.

Kirk, H., & Freeland, J. R. (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. *International Journal of Molecular Sciences* , *12* (6), 3966-3988.

Kopylova, E., Noe, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* , *28* (24), 3211-3217.

Kuhn, K., Schwenk, K., Both, C., Canal, D., Johansson, U. S., van der Mije, S., ... & Packert, M. (2013). Differentiation in neutral genes and a candidate gene in the pied flycatcher: using biological archives to track global climate change. *Ecology and Evolution* , *3* (14), 4799-4814.

Lancaster, L. T., Fuller, Z. L., Berger, D., Barbour, M. A., Jentoft, S., & Wellenreuther, M. (2022). Understanding climate change response in the age of genomics. *Journal of Animal Ecology* , *91* (6), 1056-1063.

Larison, B., Lindsay, A. R., Bossu, C., Sorenson, M. D., Kaplan, J. D., Evers, D. C., ... & Ruegg, K. (2021). Leveraging genomics to understand threats to migratory birds. *Evolutionary Applications* ,*14* (6), 1646-1658.

Lehtonen, P. K., Laaksonen, T., Artemyev, A. V., Belskii, E., Berg, P. R., Both, C., ... & Primmer, C. R. (2012). Candidate genes for colour and vision exhibit signals of selection across the pied flycatcher (*Ficedula hypoleuca* ) breeding range. *Heredity* , *108* (4), 431-440.

Le Pogam, A., O'Connor, R. S., Love, O. P., Petit, M., Regimbald, L., & Vezina, F. (2021). Coping with the worst of both worlds: Phenotypic adjustments for cold acclimatization benefit northward migration and arrival in the cold in an Arctic-breeding songbird. *Functional Ecology, 35* (6), 1240-1254.

Levin, L., Bar-Yaacov, D., Bouskila, A., Chorev, M., Carmel, L., & Mishmar, D. (2015). LEMONS–a tool for the identification of splice junctions in transcriptomes of organisms lacking reference genomes. *PloS one* , *10* (11), e0143329.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* , *25* (14), 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* , *25* (16), 2078-2079.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* , *4* (12), 981-994.

Lyngs, P. (2003). *Migration and Winter Ranges of Birds in Greenland* . Copenhagen: Danish Ornithological Society.

Macdonald, C. A., Fraser, K. C., Gilchrist, H. G., Kyser, T. K., Fox, J. W., & Love, O. P. (2012). Strong migratory connectivity in a declining Arctic passerine. *Animal Migration* , *1* (1), 23-30.

Martin, K., & Wiebe, K. L. (2004). Coping mechanisms of alpine and arctic breeding birds: extreme weather and limitations to reproductive resilience. *Integrative and Comparative Biology* , *44* (2), 177-185.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal* , *17* (1), 10-12.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* , *20* (9), 1297-1303.

McKinnon, E. A., Macdonald, C. M., Gilchrist, H. G., & Love, O. P. (2016). Spring and fall migration phenology of an Arctic-breeding passerine. *Journal of Ornithology* , *157* (3), 681-693.

Meirmans, P. G. (2020). genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources* , *20* (4), 1126-1131.

Milgroom, M. G. (2015). Migration and population structure.*Population biology of plant pathogens: Genetics, Ecology, and Evolution* , 119-146.

Montgomerie, R. and B. Lyon (2020). Snow Bunting (*Plectrophenax nivalis* ), version 1.0. In Birds of the World (S. M. Billerman, B. K. Keeney, P. G. Rodewald, and T. S. Schulenberg, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. *https://doi-org.ledproxy2.uwindsor.ca/10.2173/bow.snobun.01*

Mouchet, A., Cole, E. F., Matthysen, E., Nicolaus, M., Quinn, J. L., Roth, A. M., ... & Dingemanse, N. J. (2021). Heterogeneous selection on exploration behavior within and among West European populations of a passerine bird. *Proceedings of the National Academy of Sciences* ,*118* (28), 1-6.

Nottebohm, F. (1999). The anatomy and timing of vocal learning in birds. *The Design of Animal Communication* (pp. 63-110). Massachusetts Institute of Technology.

O'Connor, R. S., Le Pogam, A., Young, K. G., Robitaille, F., Choy, E. S., Love, O. P., ... & Vezina, F. (2021). Limited heat tolerance in an Arctic passerine: Thermoregulatory implications for cold-specialized birds in a rapidly warming world. *Ecology and Evolution* , *11* (4), 1609-1619.

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K., & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics* , *26* (4), 177-187.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* , *35* (3), 526-528

Pardo-Diaz, C., Salazar, C., & Jiggins, C. D. (2015). Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution* , *6* (4), 445-464.

Petren, K. (1998). Microsatellite primers from Geospiza fortis and cross-species amplification in Darwin's finches. *Molecular Ecology* , *7* (12), 1782-1784.

Piertney, S. B., Marquiss, M., & Summers, R. (1998). Characterization of tetranucleotide microsatellite markers in the Scottish crossbill (*Loxia scotica* ). *Molecular Ecology* , *7* (9), 1261-1263.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* , *43* (1), 223-225.

Rolland, J., Jiguet, F., Jonsson, K. A., Condamine, F. L., & Morlon, H. (2014). Settling down of seasonal migrants promotes bird diversification. *Proceedings of the Royal Society B: Biological Sciences* , *281* (1784), 20140473.

Romero, L. M., Soma, K. K., & Wingfield, J. C. (1998). Changes in pituitary and adrenal sensitivities allow the snow bunting (*Plectrophenax nivalis* ), an Arctic-breeding song bird, to modulate corticosterone release seasonally. *Journal of Comparative Physiology B* , *168* (5), 353-358.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* , *8* (1), 103-106.

Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., ... & Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology* , *23* (23), 5726-5739.

Ruegg, K., Anderson, E. C., Somveille, M., Bay, R. A., Whitfield, M., Paxton, E. H., & Smith, T. B. (2021). Linking climate niches across seasons to assess population vulnerability in a migratory bird. *Global Change Biology* , *27* (5), 3519-3531.

Schlotterer, C. (2002). Towards a molecular characterization of adaptation in local populations. *Current Opinion in Genetics & Development* , *12* (6), 683-687.

Sefc, K. M., Payne, R. B., & Sorenson, M. D. (2001). Characterization of microsatellite loci in village indigobirds *Vidua chalybeata*and cross-species amplification in estrildid and ploceid finches. *Molecular Ecology Notes* , *1* (4), 252-254.

Sekercioglu, C. H. (2007). Conservation ecology: area trumps mobility in fragment bird extinctions. *Current Biology* , *17* (8), R283-R286.

Smulders, T. V., & DeVoogd, T. J. (2000). Expression of immediate early genes in the hippocampal formation of the black-capped chickadee (*Poecile atricapillus* ) during a food-hoarding task. *Behavioural Brain Research* , *114* (1-2), 39-49.

Snell, K. R., Stokke, B. G., Moksnes, A., Thorup, K., & Fossoy, F. (2018). From Svalbard to Siberia: Passerines breeding in the High Arctic also endure the extreme cold of the Western Steppe. *PloS one* , *13* (9), e0202114.

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution* , *29* (12), 673-680.

Tigano, A., Shultz, A. J., Edwards, S. V., Robertson, G. J., & Friesen, V. L. (2017). Outlier analyses to test for local adaptation to breeding grounds in a migratory arctic seabird. *Ecology and Evolution* , *7* (7), 2370-2381.

UniProt Consortium. (2015). UniProt: a hub for protein information. *Nucleic Acids Research* , *43* (D1), D204-D212.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research* , *40* (15), e115-e115.

Valladares, F., Matesanz, S., Guihaumon, F., Araujo, M. D., Balaguer, L., Benito-Garzon, M., ... & Zavala, M. A. (2014). The effects of phenotypic plasticity and local adatation on forecasts of species range shifts under climate change. *Ecology Letters* , 17(11), 1351-1364.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... & Banks, E. (2013). From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics* , *43* (1), 11-10.

Vezina, F., Williams, T. D., Piersma, T., & Guy Morrison, R. I. (2012). Phenotypic compromises in a long-distance migrant during the transition from migration to reproduction in the High Arctic. *Functional Ecology* , *26* (2), 500-512.

Visser, M. E., Gienapp, P., Husby, A., Morrisey, M., de la Hera, I., Pulido, F., & Both, C. (2015). Effects of spring temperatures on the strength of selection on timing of reproduction in a long-distance migratory bird. *PLoS Biology* , *13* (4), e1002120.

Vo, A. T., & Jedlicka, J. A. (2014). Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for faecal and swab samples. *Molecular Ecology Resources* , *14* (6), 1183-1197.

Walker, B. G., Meddle, S. L., Romero, L. M., Landys, M. M., Reneerkens, J., & Wingfield, J. C. (2015). Breeding on the extreme edge: modulation of the adrenocortical response to acute stress in two High Arctic passerines. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* , *323* (4), 266-275.

Wanamaker, S. M., Singh, D., Byrd, A. J., Smiley, T. M., & Ketterson, E. D. (2020). Local adaptation from afar: migratory bird populations diverge in the initiation of reproductive timing while wintering in sympatry. *Biology Letters* , *16* (10), 20200493.

Warner, N. A., Sagerup, K., Kristoffersen, S., Herzke, D., Gabrielsen, G. W., & Jenssen, B. M. (2019). Snow buntings (*Plectrophenax nivalis* ) as bio-indicators for exposure differences to legacy and emerging persistent organic pollutants from the Arctic terrestrial environment on Svalbard. *Science of The Total Environment* , *667* , 638-647.

Wellband, K. W., Pettitt-Wade, H., Fisk, A. T., & Heath, D. D. (2018). Standing genetic diversity and selection at functional gene loci are associated with differential invasion success in two non-native fish species. *Molecular Ecology* , *27* (7), 1572-1585.

Wingfield, J. C., Owen-Ashley, N. T., Benowitz-Fredericks, Z. M., Lynn, S., Hahn, T. P., Wada, H., ... & Romero, L. M. (2004). Arctic spring: the arrival biology of migrant birds. *Acta Zoologica Sinica* ,*50* (6), 948-960.

Winker, K., Graves, G. R., & Braun, M. J. (2000). Genetic differentiation among populations of a migratory songbird:*Limnothlypis swainsonii* . *Journal of Avian Biology* , *31* (3), 319-328.

Wonke, G., Wallschlager, D., Moll, K., & Tiedemann, R. (2007). Ten new microsatellite loci for the yellowhammer (*Emberiza citrinella* ) and their cross-species applicability among related taxa. *Molecular Ecology Notes* , *7* (6), 1278-1280.

Yeaman, S., & Otto, S. (2011). Establishment and maintenance of adaptive genetic divergence under migration, selection and drift.*Evolution* , 65(7), 2123-2129.

Zeigler, H., & Marler, P. E. (2004). Behavioral neurobiology of birdsong. New York Academy of Sciences.

Zhan, X., Dixon, A., Batbayar, N., Bragin, E., Ayas, Z., Deutschova, L., ... & Bruford, M. W. (2015). Exonic versus intronic SNPs: contrasting roles in revealing the population genetic differentiation of a widespread bird species. *Heredity* , *114* (1), 1-9.

Zimmerman, S. J., Aldridge, C. L., Oh, K. P., Cornman, R. S., & Oyler-McCance, S. J. (2019). Signatures of adaptive divergence among populations of an avian species of conservation concern. *Evolutionary Applications* , *12* (8), 1661-1677.

**Data Accessibility**

All genetic data associated with this project will be deposited in the SRA or on Dryad, and their associated accession numbers will be listed here upon article acceptance.

**Author Contributions**

KKP, OPL and DDH conceived this work. KKP wrote the manuscript and collected all the data. KWW assisted with de-novo transcriptome assembly and SNP characterization. OPL, ALP, FV, NTA, FF, KW, MMR, HGG, AM, BGS and JW were involved in sample design and/or sample collection for the population genetic study. KKP, OPL, ALP and FV were involved in RNASeq sample design and/or collection. KKP conducted the data analyses and wrote the first draft of the manuscript and all authors contributed to the revisions of the manuscript.

**Tables**

**Table 1:** Location and life history trait data for six snow bunting (Plectrophenax nivalis) populations used in this study.

| | Alert, NU, Canada (A) | Utqiagvik (Barrow), AK, USA (U) | Mitivik (East Bay) Island, NU, Canada (M) | Svalbard, Norway (S) | Aleutian Islands, AK, USA (AI) | Pribilof Islands, AK, USA (PI) |
|---|---|---|---|---|---|---|
| Subspecies [1] | *Plectrophenax nivalis nivalis* | *Plectrophenax nivalis nivalis* | *Plectrophenax nivalis nivalis* | *Plectrophenax nivalis nivalis* | *Plectrophenax nivalis townsendi* | *Plectrophenax nivalis townsendi* |
| Migratory/ Resident | Migratory [1] | Migratory [1] | Migratory [2] | Migratory [3] | Resident **[1]** | Resident **[1]** |
| Migration Distance | Currently unknown | Currently unknown | Fall: ˜2660 ± 59 km; Spring: 2147 ± 69 km [4] | Fall: >1000 km [3] | N/A | N/A |
| Nesting Location | Rocky cavities [5] | Cavities in various human-made objects or nest boxes [6] | Rocky nesting cavities in Arctic tundra [7] | Cavities in various human-made objects or nest boxes [8] | Rocky cavities on the ground **[1]** | Rocky cavities on the ground **[1]** |
| Clutch size | 5-6 eggs [5] | 3-8 eggs [9] | 5-7 eggs [10] | 5-7 eggs [8,11] | Currently unknown | Currently unknown |
| # of broods per year | 1 [6] | 1, but can be 2 if weather conditions are favourable [9] | 1 [7] | 1, but can be 2 if weather conditions are favourable [8] | Currently unknown | Currently unknown |
| Wintering location | Currently unknown | Currently unknown | Manitoba, Saskatchewan and Alberta, Canada [4] | Siberian steppe [3] | N/A | N/A |
| Breeding season | May-July [12] | May- July [6] | Late May-Aug [7,10] | May-July [8] | May-Sept **[1]** | May-Sept **[1]** |

[1]Montgomerie & Lyon, 2020;[2]Macdonald et al., 2012;[3]Snell et al., 2018;[4]McKinnon et al., 2016;[5]Vézina, pers. comm.;[6]Romero et al., 1998;[7]Guindre-Parker et al., 2013a;[8]Fossøy et al., 2014;[9]Ashley, pers.

comm.;[10]Guindre-Parker et al., 2013b;[11]Warner et al., 2019;[12]O'Connor et al., 2021

**Table 2:** Summary statistics for snow bunting (*Plectrophenax nivalis* ) samples used for DNA extraction for the breeding population genetics study. These 221 samples were collected from the snow bunting populations during their breeding season (May-September).

| Population | Location | Type of Sample | DNA Extrac |
|---|---|---|---|
| Alert, NU, Canada (A) | 82.30°N, 62.20°W | Dry blood spot on a filter paper | SPRI Beads |
| Svalbard, Norway (S) | 78.13°N, 15.38°E | Packed red blood cells (RBC) in ethanol | SPRI Beads |
| Utqiagvik (Barrow), AK, USA (U) | 71.10°N, 156.40°W | Frozen RBC<br>Whole blood and frozen RBC | QIAamp DN |
| Mitivik (East Bay) Island, NU, Canada (M) | 64.01°N, 81.47°W | Dry blood spot on a filter paper | SPRI Beads |
| Aleutian Islands, AK, USA (AI) | 51.89°N, 176.64°W<br>52.89°N, 173.11°W | Muscle tissue preserved in ethanol | SPRI Beads |
| Pribilof Islands, AK, USA (PI) | 57.14°N, 170.23°W | Muscle tissue preserved in ethanol | SPRI Beads |

RBC: red blood cells, SPRI: solid phase reversible immobilization

Permits: Alert sample collection approved by Animal Care Committee at Université du Québec à Rimouski (CPA-61-15-163 and CPA-68-17-186), Utqiagvik sample collection approved by the Institutional Animal Care and Use Committee at Western Kentucky University (A3558-01), Mitivik Island sample collection approved by the Animal Care Committee at University of Windsor (AUPP-09-14), Svalbard sample collection approved by Norwegian Food Safety Authority (2014/38064-2 (FOTS ID 4701)), and Aleutian and Pribilof Islands sample collection approved by appropriate animal care permits at University of Alaska Museum.

**Table 3:** Microsatellite marker pairwise$F_{ST}$ values (below diagonal) and $p$ values for Fisher's exact test of population differentiation (above diagonal) for six snow bunting (*Plectrophenax nivalis* ) breeding populations. Bold indicates statistically significant differences after sequential Bonferroni correction at 5% level. See Table 2 for descriptions of population codes.

| | AI | PI | A | U | M | S |
|---|---|---|---|---|---|---|
| AI | – | **<0.0001** | **<0.0001** | **<0.0001** | **<0.0001** | **<0.0001** |
| PI | **0.091** | – | **<0.0001** | **0.0011** | **<0.0001** | **<0.0001** |
| A | **0.100** | **0.036** | – | **<0.0001** | **<0.0001** | **0.0001** |
| U | **0.094** | **0.035** | **0.011** | – | **<0.0001** | 0.0111 |
| M | **0.095** | **0.039** | -0.0001 | **0.012** | – | **<0.0001** |
| S | **0.081** | **0.028** | **0.012** | -0.0001 | **0.009** | – |

**Table 4:** Single Nucleotide Polymorphism (SNP) loci pairwise$F_{ST}$ values (below diagonal) and $p$ values for Fisher's exact test of population differentiation (above diagonal) for six snow bunting (*Plectrophenax nivalis* ) breeding populations. Bold indicates statistically significant differences after sequential Bonferroni correction at 5% level. See Table 2 for descriptions of population codes.

| | AI | PI | A | U | M | S |
|---|---|---|---|---|---|---|
| AI | – | 0.009 | **0.0009** | **<0.0001** | **<0.0001** | **<0.0001** |
| PI | 0.021 | – | **<0.0001** | **0.0008** | **<0.0001** | **<0.0001** |
| A | **0.042** | **0.051** | – | 0.7705 | 0.8786 | 0.6237 |

| | AI | PI | A | U | M | S |
|---|---|---|---|---|---|---|
| U | **0.035** | **0.039** | 0.004 | – | 0.5645 | **0.0008** |
| M | **0.042** | **0.047** | 0.008 | 0.012 | – | 0.5531 |
| S | **0.053** | **0.044** | 0.013 | **0.024** | 0.005 | – |

**Table 5:** Selection effects on a subset of functional Single Nucleotide Polymorphism (SNP) loci among six snow bunting (*Plectrophenax nivalis* ) breeding populations assayed at the global and pairwise levels. The 11 SNP loci were chosen based on having selection status data for all possible pairwise comparisons (See text for more detail). For each SNP locus the associated gene, type of variant, specific SNP, amino acid substitution and gene function category are given. The gene function categories include: Energetics (E), Cellular Housekeeping (H), Lipid Metabolism (LM), Nervous System Development (NS), Reproduction (R), and Stress (S). For each SNP locus, divergent selection (black), stabilizing selection (green), or genetic drift (blue) is shown based on 99% and 99.9% neutral marker confidence intervals for global and pairwise comparisons, respectively. See Table 2 for descriptions of population codes.

| **Primer Name** | **Gene Description** | **Type of Variant** | **Nucleotide Variant** | **Amino Acid Variant** | **Category** | **Global Comparison** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** | **Population Compared** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | **AI/PI** | **AI/A** | **AI/U** | **AI/M** | **AI/S** | **PI/A** | **PI/U** | **PI/M** | **PI/S** | **A/U** | **A/M** | |
| **SNP_10** | Serine/threonine-protein kinase LATS2 | Missense | G/T | Ser/Asn | H | | | | | | | | | | | | | |
| **SNP_13** | DNA repair protein complementing XP-C cells | Missense | G/A | Arg/Lys | H | | | | | | | | | | | | | |
| **SNP_100** | Corticotropin-releasing factor receptor 1 | Upstream | G/A | - | H | | | | | | | | | | | | | |
| **SNP_156** | Hexosaminidase D | Downstream | A/G | - | LM | | | | | | | | | | | | | |

| Primer Name | Gene Description | Type of Variant | Nucleotide Variant | Amino Acid Variant | Global Comparison | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared | Population Compared |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SNP_-41** | Activin receptor type-2A | Missense | A/G | Ser/Pro | LM | | | | | | | | | | | | |
| **SNP_-105** | Ankyrin repeat and LEM domain-containing protein 2 | Downstream | G/A | - | NS | | | | | | | | | | | | |
| **SNP_-56** | Activated CDC42 kinase 1 | Missense | G/A | Val/Met | NS | | | | | | | | | | | | |
| **SNP_-175** | Protocadherin gamma-C5 | Synonymous | G/A | Pro/Pro | NS | | | | | | | | | | | | |
| **SNP_-24** | BTB/POZ domain-containing protein KCTD17 | Missense | T/A | Cys/Ser | R | | | | | | | | | | | | |
| **SNP_-60** | Receptor type tyrosine-protein phosphatase zeta | Missense | A/C | His/Pro | R | | | | | | | | | | | | |
| **SNP_-140** | Transcription regulator protein BACH2 | Downstream | T/G | - | S | | | | | | | | | | | | |

## Figures

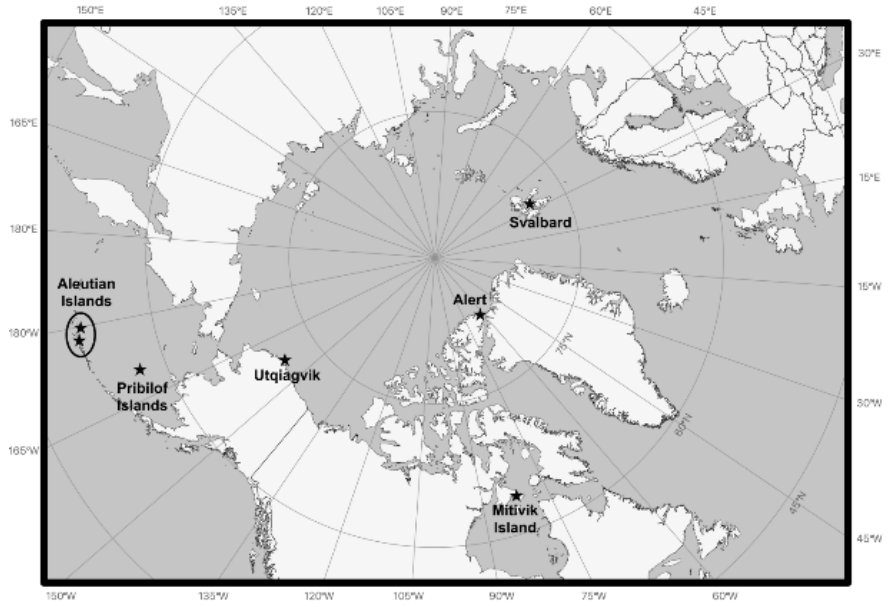


**Figure 1:** Map showing the snow bunting (*Plectrophenax nivalis* ) sampling sites (as indicated by stars) for the breeding population genetics study. Map created using the Free and Open Source QGIS. See Table 1 for descriptions of sample locations.

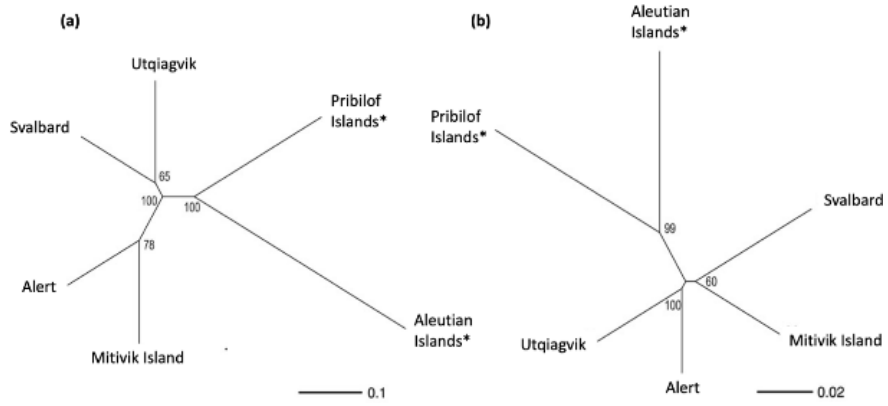


**Figure 2:** Unrooted neighbor-joining cluster analysis diagrams of snow bunting (*Plectrophenax nivalis* ) breeding populations based on Cavalli-Sforza and Edwards' (1967) chord distance for microsatellite (Panel a) and Single Nucleotide Polymorphism (SNP) (Panel b) markers. The data were bootstrapped over loci with replacement, using 10000 permutations; numbers at branch sites represent the bootstrap support (%) of the branch (support less than 50% is not shown). Asterisks represent non-migratory (resident) populations; others are migratory populations.

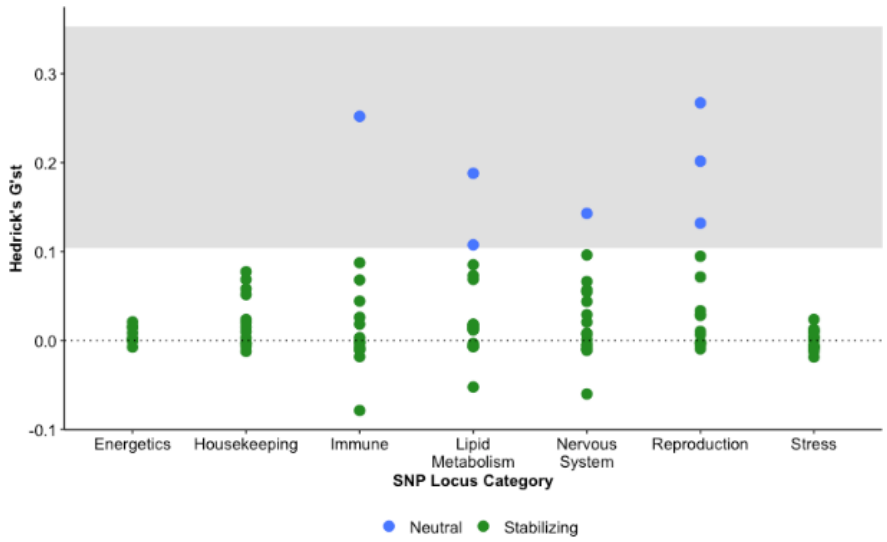


**Figure 3:** Distribution of global Hedrick's $G'_{ST}$values across the six sampled snow bunting (Plectrophenax nivalis) breeding populations for each of the 101 functional Single Nucleotide Polymorphism (SNP) loci used in the study. The SNP marker genes were selected from seven broad putative gene function categories. The 99% confidence interval range for neutral divergence (i.e., based on microsatellite marker data $G'_{ST}$ values) is shown in grey.

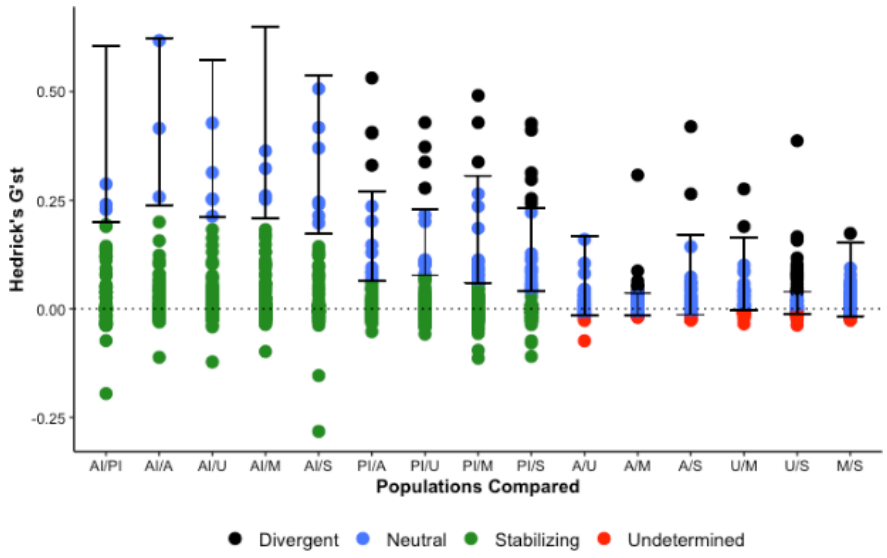


**Figure 4:** Distribution of pairwise Hedrick's $G'_{ST}$ values for the sampled snow bunting (*Plectrophenax nivalis* ) breeding populations based on the 101 functional Single Nucleotide Polymorphism (SNP) loci. Pairwise comparisons identified as under neutral processes (blue dots) fall within the 99.9% microsatellite marker confidence interval range (shown as error bars for each population comparison). Divergent (black dots) and stabilizing (green dots) selection were determined using the same neutral CI. It was not possible to determine selection status ("Undetermined'; red dots) due to the neutral microsatellite range including zero. See Table 2 for descriptions of population codes.

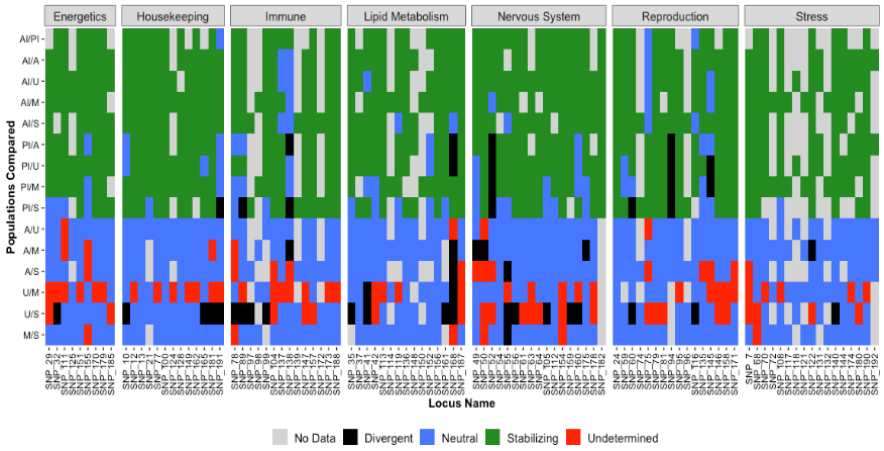


**Figure 5:** Summary map of Single Nucleotide Polymorphism (SNP) marker selection status for all pairwise comparisons of six snow bunting (*Plectrophenax nivalis* ) breeding populations. All SNPs belong to one of the gene function categories as shown at the top of the figure. All selection status results are based on pairwise Hedrick's $G'_{ST}$ value comparisons with the 99.9% neutral marker range. For some pairwise comparisons, we could not estimate Hedrick's $G'_{ST}$ values (grey squares; "No data"), likely due to insufficient sequence reads. The red squares (Undetermined) are pairwise comparisons where the neutral range included zero, making stabilizing selection impossible to detect. See Table 2 for descriptions of population codes.