Received: 7 July 2023 | Revised: 3 May 2024 | Accepted: 7 May 2024

DOI: 10.1111/jbi.14961

Journal of Biogeography | WILEY

RESEARCH ARTICLE

Multiple Pleistocene refugia for Arctic Bell-Heather revealed with genomic analyses of modern and historic plants

Cassandra Elphinstone[1,2] | Fernando Hernández[1] | Marco Todesco[1] | Jean-Sébastien Légaré[1] | Winnie Cheung[1] | Paul C. Sokoloff[3] | Annika Hofgaard[4] | Casper T. Christiansen[5] | Esther R. Frei[6,7] | Esther Lévesque[8] | Gergana N. Daskalova[9] | Haydn J. D. Thomas[9] | Isla H. Myers-Smith[9] | Jacob A. Harris | Jeffery M. Saarela[3] | Jeremy L. May[10] | Joachim Obst | Julia Boike[11,12] | Karin Clark[13] | Katie MacIntosh | Katlyn R. Betway-May[14] | Liam Case | Mats P. Björkman[15,16] | Michael L. Moody[17] | Niels Martin Schmidt[18] | Per Molgaard[19] | Robert G. Björk[15,16] | Robert D. Hollister[14] | Roger D. Bull[3] | Sofie Agger[2] | Vincent Maire[8] | Greg H. R. Henry[2] | Loren H. Rieseberg[1]

Correspondence
Cassandra Elphinstone, Department of Botany, University of British Columbia, Unceded xʷməθkʷəy̓əm (Musqueam) Territory, Vancouver, BC V6T 1Z4, Canada.
Email: cassandra.elphinstone@gmail.com

Funding information
Natural Sciences and Engineering Research Council of Canada; GenomeBC, Grant/Award Number: SIP009; US National Science Foundation, Grant/Award Number: DEB-1637459; ArcticNet; US National Science Foundation, Grant/Award Number: DEB-1504224

Abstract

Aim: Arctic plants survived the Pleistocene glaciations in unglaciated refugia. The number, ages, and locations of these refugia are often unclear. We use high-resolution genomic data from present-day and Little-Ice-Age populations of Arctic Bell-Heather to re-evaluate the biogeography of this species and determine whether it had multiple independent refugia or a single refugium in Beringia.

Location: Circumpolar Arctic and Coastal British Columbia (BC) alpine.

Taxon: *Cassiope tetragona* L., subspecies *saximontana* and *tetragona*, outgroup *C. mertensiana* (Ericaceae).

Methods: We built genotyping-by-sequencing (GBS) libraries using *Cassiope tetragona* tissue from 36 Arctic locations, including two ~250- to 500-year-old populations collected under glacial ice on Ellesmere Island, Canada. We assembled a de novo GBS reference to call variants. Population structure, genetic diversity and demography were inferred from PCA, ADMIXTURE, fastsimcoal2, SplitsTree, and several population genomics statistics.

Results: Population structure analyses identified 4–5 clusters that align with geographic locations. Nucleotide diversity was highest in Beringia and decreased eastwards across Canada. Demographic coalescent analyses dated the following splits with Alaska: BC subspecies *saximontana* (5 mya), Russia (~1.4 mya), Europe (>200–600 kya),

Greg H. R. Henry and Loren H. Rieseberg are joint senior authors.

For affiliations refer to page 1970.

and Greenland (~60 kya). Northern Canada populations appear to have formed during the current interglacial (7–9 kya). Admixture analyses show genetic variants from Alaska appear more frequently in present-day than historic plants on Ellesmere Island. **Conclusions:** Population and demographic analyses support BC, Alaska, Russia, Europe and Greenland as all having had independent Pleistocene refugia. Northern Canadian populations appear to be founded during the current interglacial with genetic contributions from Alaska, Europe and Greenland. We found evidence, on Ellesmere Island, for continued recent gene flow in the last 250–500 years. These results suggest that a re-analysis of other Arctic species with shallow population structure using higher resolution genomic markers and demographic analyses may help reveal deeper structure and other circumpolar glacial refugia.

KEYWORDS
climate change, empirical population genetics, gene flow, genotyping-by-sequencing, glaciations, historic DNA, ITEX, range expansion, refugia, tundra plants

1 | INTRODUCTION

The Arctic flora has a complex biogeographic history due to repeated glacial cycles causing extreme range expansions and contractions throughout the Pleistocene (Abbott & Brochmann, 2003; Dyke & Prest, 1987; Hultén, 1937). While complex biogeographic histories can be difficult to deduce from fossil data, genomic data offer a means to infer refugia, their relative ages, potential colonization routes and estimate gene flow (Wang et al., 2021). Genomics studies have great potential in the Arctic to determine past population demography and predict future climate adaptation (Wullschleger et al., 2015). Few studies have tested the ages of potential Arctic plant refugia using high-resolution genomic data (Ikeda et al., 2017), and there is still much uncertainty as to where and when these refugia formed for various species. By combining broad geographical sampling, high-resolution genomic markers and historical samples, we can begin to infer the age of Arctic refugia and observe recent gene flow in Arctic plant populations.

Beringia (the region between eastern Siberia and western Yukon) has been proposed as an unglaciated refuge and the main origin of Arctic plant species (Brochmann & Brysting, 2008; Hultén, 1937). Genetic diversity is generally higher than the rest of the Arctic, suggesting Beringia lineages are older and have had more time to diversify; however, details for individual species remain uncertain (Alsos et al., 2022; Abbott & Brochmann, 2003; Eidesen et al., 2013). Species distribution models suggest other Pleistocene-aged refugia for Arctic plants in the mountains in Northwestern USA, Western Europe and South-Central Asia (Pellissier et al., 2016). Nunataks, unglaciated rocky outcrops in icefields, have also been proposed as potential cryptic refugia where Arctic/alpine species may have survived in small, isolated populations during the glaciations (Beatty & Provan, 2010; Westergaard et al., 2011). Fossils and pollen records support that many High Arctic plant species had circumpolar distributions at various times since the polar tundra biome initially formed about 2–3 mya (millions of years ago) (Brochmann & Brysting, 2008). However, most previous molecular datasets on Arctic plants used highly conserved chloroplast DNA sequences or a small number of amplified fragment length polymorphisms (AFLPs), so fine-scale resolution of genomic clusters and estimation of demographic history was not possible.

Since 1950–1960, following the Little Ice Age (Little-Ice-Age; 1450–1850 CE), polar-frozen-based glaciers in the High Arctic have been retreating, uncovering historic plants essentially intact (Bergsma et al., 1984; Jones & Henry, 2003). At Alexandra Fiord, Ellesmere Island, in the last 60 years, the ice has retreated approximately 300 m from its maximum extent (O'Kane, 2018). *Cassiope tetragona* is one of the best-preserved species to emerge from under the ice, and thus, we have used these historic populations to extend our genetic sampling backwards in time.

Cassiope tetragona (L.) D. Don is a diploid ($2n = 26$), perennial, ecologically well-studied (Havstrom et al., 1995; Mallik et al., 2011; Molau, 1997; Rayback & Henry, 2006) Arctic/alpine species. Two subspecies of *C. tetragona* have been described: ssp. *tetragona* (*L.*) *D. Don*, with a circumpolar distribution; and ssp. *saximontana* (*Small*) *A. E. Porsild*, found in Western Canada. *Cassiope tetragona* is insect pollinated (Alsos et al., 2013) but may also self. Small seeds enable wind dispersal (Eidesen et al., 2007), facilitating widespread colonization. Under experimental conditions, *C. tetragona* has low seed germination (7.8%) (Alsos et al., 2013), possibly hampering post-dispersal establishment. Previous RADseq data on the *Cassiope* genus showed evidence that the genus originated in the north (Siberia) and later spread southwards (through the Himalayan-Hengduan Mountains) (Hou et al., 2016). *Cassiope tetragona* is believed to have split from its sister species *C. mertensiana* (Bong.) G. Don about 18.6 mya (CI: 14.8–37.2 mya; Kumar et al., 2017).

Previous phylogenetic analyses of the Arctic populations of *C. tetragona* suggested that the species may have expanded its range eastward out of a Beringian refuge across to Europe as recently as the last interglacial period (<11 thousand years ago (kya)) (Eidesen et al., 2007). They also suggested a westward expansion to Siberia in the mid-to-late Pleistocene. The presence of additional refugia could

not be ruled out due to limited genotypic information (265 AFLP markers in 56 populations) that may have obscured more complex patterns of genetic diversity.

We re-investigate the biogeography of *C. tetragona* from 36 extant pan-Arctic populations, two historic (250- to 500-year-old) populations sampled from under glacial ice in the Canadian High Arctic and an outgroup (*C. mertensiana*) using up to 26,350 single nucleotide polymorphisms (SNPs). Our goal is to determine whether the current circumpolar populations of *C. tetragona* derived from recent expansion out of a Beringian refuge (Eidesen et al., 2007) or whether *C. tetragona* populations could have persisted in several other glacial refugia during the Pleistocene glaciations. Specifically, we estimate the ages of genetic clusters to see if they are younger or older than the Last Glacial Maximum (LGM) (ending 11 kya). We also investigate if and how Little-Ice-Age populations differ from present-day populations.

2 | MATERIALS AND METHODS

2.1 | Sample collection and DNA extraction

During summer 2017, researchers primarily associated with the International Tundra Experiment (ITEX), a collaborative group of researchers studying tundra ecosystems (Henry et al., 2022), collected *C. tetragona* leaf tissue on silica from 36 pan-Arctic and alpine locations (Table S3, Appendix S1 – Field sampling) covering all countries in *C. tetragona*'s range. Two historic populations (Alexandra Fiord and Sverdrup Pass, Ellesmere Island, Nunavut, Canada) were retrieved from retreating glacial ice. We observed that the frozen plants appeared to have been caught in a summer snowstorm or early onset of winter since many of the plants we found were frozen in flower (not having set seeds yet). Two populations of *C. mertensiana*, from the southwestern British Columbia (BC) alpine, served as an outgroup (MER). Population codes were assigned according to the geographic location of the sampling site (Table S3).

DNA was extracted from 20 mg of dried leaf tissue from each of the 5 to 12 individuals per site using a 3% CTAB protocol (Zeng et al., 2002) that was modified for high phenol content and acidity in *C. tetragona* leaves (Appendix S1 – Modified 3% CTAB). The modified CTAB protocol was also used to extract DNA from 24 historic plants. DNA was obtained from 387 samples, which were subsequently prepared for sequencing.

2.2 | Sequencing libraries

Genotyping-by-sequencing (GBS) was selected as a cost-effective way to obtain enough SNPs to differentiate populations in recently geographically expanded species such as *C. tetragona*. GBS was conducted using New England Biolabs PstI Hi-Fidelity and MspI restriction enzymes following a modified version of the Elshire et al. (2011) protocol (Appendix S1 Genotyping-by-sequencing library preparation). After a polymerase chain reaction (PCR) enrichment step, pooling, size selection of 300–500 base pair (bp) fragments and enzymatic depletion of repetitive sequences (Moyers et al., 2017), enriched libraries for 371 samples were successfully prepared, including 10 of 12 historic samples from Alexandra Fiord and all 12 samples from Sverdrup Pass (16 of the 387 samples failed to amplify). Libraries were sequenced at the Centre d'Expertise et de Services Génome Québec on two lanes of the Illumina HiSeq4000, generating 571 million 150 bp paired end reads.

2.3 | Data processing

De-multiplexing was completed using a Perl script from Owens et al. (2016). Other data processing steps were run in dDocent 2.9.4 (Puritz et al., 2014). The ends of reads with a quality score <20 or an average quality score that was <10 in a sliding window of 5 bp were trimmed with Trimmomatic (Bolger et al., 2014). A reference assembly was not available for *C. tetragona*. We used GBS reads from 55 *C. tetragona* individuals (from ATQ, BARD, LAJ, PET, SAM and YED; see Table S3) to build a de novo GBS assembly for paired end reads. The assembly was built with a c-parameter (% similarity to cluster) of 0.95, within individual coverage of 3, and between individual coverage of 5 in CD-HIT (LaCava et al., 2020) and Rainbow (Chong et al., 2012). *Cassiope mertensiana* individuals were excluded from the assembly to prevent duplicating contigs for divergent regions.

Reads were mapped to the de novo GBS assembly using BWA-MEM (Li & Durbin, 2009). *Cassiope tetragona* ssp. *tetragona* were mapped with a match score of 1, a mismatch score of 4 and a gap penalty of 6 (the conservative default values). Our outgroup samples from BC, *C. mertensiana* and *C. tetragona* ssp. *saximontana* were mapped with more relaxed parameters of match score of 1, a mismatch score of 3 and a gap penalty of 5. Variant calling was done on all individuals via FreeBayes (Garrison & Marth, 2012) using the dDocent 2.9.4 default parameters.

High levels of DNA degradation in historic DNA samples can often lead to high levels of bacterial contamination in sequencing libraries. For this reason, historic plants were excluded from the de novo GBS assembly. Bacterial vs eukaryotic reads were identified by BLASTing each read to the NCBI database and calculating the percentage of total reads that mapped to Bacteria versus Eukaryotes (Table S5). Bacterial reads were removed by mapping the historic reads to the de novo reference genome.

2.4 | SNP filtering

We filtered all variants in the raw variant call file (vcf) using vcftools 0.1.16 and plink 1.9b_6.21-x86_64 (Danecek et al., 2011). For filtering, we keep only biallelic SNPs with <60% of heterozygote individuals, quality score >30, present in at least 90% of the samples and with a minor allele frequency (MAF) > 0.01. For all analyses except the population statistics, SplitsTree and ABBA-BABA tests, linkage

disequilibrium (LD) was reduced by keeping 1 SNP/300 bp (length of the longest contig). We also removed individuals missing >10% of the variants found across all sites and the outgroup individuals (with poor mapping) from all analyses except the demographic modelling. For demographic models, we projected data to 10 individuals per group and retained the outgroup and any individuals missing up to 30% of SNPs.

2.5 | Population structure

Population structure is influenced by the demographic history of populations. Multiple significant clusters suggest the possibility of multiple independent refugia. To visualize variation and clustering among individuals and sites, principal component analysis (PCA) was performed in R 4.2.2 using the packages SNPRelate (Zheng et al., 2020) and tidyverse (Wickham et al., 2019). ADMIXTURE 1.3.0 (Alexander & Lange, 2011) was used to assign potential genome ancestry to SNPs using a model-based approach with the assumption that ancestral populations were in Hardy–Weinberg Equilibrium (HWE). The termination criterion was set to when the log likelihood changes between iterations fell below 10^{-10}. The tenfold cross-validation (CV) error was recorded for each value of K (1 to 14) tested. The last large drop in CV error was used to select the K value best representative of the number of clusters in the data. All pairwise F_{ST} values between the identified clusters were calculated. SplitsTree (Huson & Bryant, 2006; Huson, 1998) and TreeMix (Pickrell and Pritchard, 2012) were additionally run to confirm population structure (Appendix S1 – SplitsTree, TreeMix).

2.6 | Population statistics

Using population statistics, such as nucleotide diversity, which are known to be associated with population age and effective population size, we can make predictions regarding which locations may have been refugia based on their potential sizes/ages. Each sampling location was grouped together as a "population." Nucleotide diversity (π) was calculated using the vcftools 0.1.16 – site-π option and across 150 bp windows using the – window-π option. These π values for each SNP or window were summed to obtain a single sites-π and windows-π summed measure for each location. Descriptions of other common population statistics can be found in Appendix S2 – Population statistics.

To investigate whether deglaciation time was associated with nucleotide diversity, values of π for the Arctic populations (ssp. *tetragona*) were compared with deglaciation dates. Deglaciation dates were obtained from Dalton et al. (2020) for North America and Hughes et al. (2016); Stroeven et al. (2016)) for Europe using QGIS 2.18.10 (http://qgis.org).

F_{ST} for all pairs of sampled populations was calculated in vcftools with the –weir-fst-pop option. Three Mantel tests were run to explore the influence of geographic distance on relatedness. All tests used the mantel.rtest in ade4 (Thioulouse et al., 2018), R 3.5.0 with 9999 replicates. Populations with dominant admixture ancestry associated with the 'Alaska' cluster, 'Europe' cluster and 'Greenland' cluster were split into three groups, respectively, and a Mantel test was run on each group.

2.7 | Demographic models (site frequency spectrum)

To estimate the relative ages of the various admixture clusters, we ran coalescent demographic analyses on the folded site frequency spectra using fastsimcoal2 (Excoffier et al., 2021). Specifically, we aimed to estimate divergence times of the major genetic groups to determine whether they were old enough to have been glacial refugia. Each model was independently calibrated to test whether these groups consistently diverged before or after the LGM. To calibrate divergence times, in the first two models (A and B), we used the split between *C. tetragona* and *C. mertensiana* (18.6 mya) (Kumar et al., 2017) as a fixed parameter, while in models C and D, we assumed that populations in Northern Canada (NWT and Nunavut) were formed since the last major glaciation (7–9 kya, respectively) and used those dates as fixed parameters. Divergence times were expressed in generations. *Cassiope tetragona*'s generation time is not known but observationally it does not flower for at least the first 5 years of its life and can live to be up to 70–150 years old (Rayback & Henry, 2006; Weijers et al., 2017). To confirm that varying generation times did not affect our conclusions, we ran a model varying the generation times, assuming 10, 50 and 100 years, and compared the results. We also ran each model with and without an instantaneous bottleneck during each new population founding event.

Using fastsimcoal2, we were computationally unable to run more than five clusters within one model, hence the split of models A and B and C and D. The four main distinct models (A–D, Figure 2 and Appendix S1-Fastsimcoal2) each include five clusters: ModelA: *C. mertensiana*, ssp. *saximontana* and ssp. *tetragona* from Russia, Alaska and Europe; ModelB: *C. mertensiana*, and ssp. *tetragona from* Russia, Alaska, Europe and Greenland; ModelC: only ssp. *tetragona* from Russia, Alaska, Europe, NWT and Nunavut; and ModelD: only ssp. *tetragona* from Alaska, Europe, Greenland, NWT and Nunavut. The order in which each cluster split off from Alaska/Beringia was not defined by the model initially, allowing for flexibility in when each refugia was founded. For the first two models, we selected individuals with >95% of membership to their group in ADMIXTURE, excluding hybrids. Each group was projected to ten individuals (-proj 20 in easySFS) (Gutenkunst et al., 2009), except for *C. mertensiana*, which only had seven non-hybrid individuals (-proj 14). ModelC and ModelD included more recently formed Canadian hybrid populations, so to account for the admixed ancestry of these populations, we included in the models a bidirectional gene flow matrix. Each model was run 50 times. Within each run, 100,000 simulations were done, each with 48 optimization iterations. We did not compare models A through D. We ran all four models separately and, for each model, we compared the 50 runs including different combinations of

parameters. The best run of each model was selected based on the maximum estimated likelihood and used for extracting the values for each parameter. Site frequency spectra were plotted for the raw data in python 3.10 with dadi (Gutenkunst et al., 2009).

2.8 | Historic samples

To determine the ages of the historic tissue from under the glacial ice, leaf/stem tissue from five of the historic samples was C-14 dated using Accelerator Mass Spectrometry (AMS) at the André E. Lalonde AMS Laboratory, University of Ottawa.

To test recent admixture between two major genetic clusters, present-day and historic populations at Alexandra Fiord were used. We employed ABBA-BABA tests in Dsuite (Durand et al., 2011; Green et al., 2010; Malinsky et al., 2021) across four populations: P1, P2, P3 and the outgroup (O), related by the phylogeny (((P1, P2), P3), O). SNPs for this analysis were filtered to have as many variants as possible to compare between individuals from the outgroup (seven samples of *C. mertensiana* without admixture), AlexOld and AlexNew samples. SNPs were filtered out if they: had a quality score <30, were indels or non-biallelic SNPs and called in less than 95% of individuals (from the outgroup, AlexNew and AlexOld). The allele carried by the outgroup was designated as the ancestral allele (A), and the derived allele was designated as B. Under the null hypothesis, which assumes no gene flow after P1 and P2 split, the ABBA (B shared by P2 and P3) and BABA (B shared by P1 and P3) patterns are expected to occur with equal frequency due to incomplete lineage sorting. A significant increase in ABBA or BABA is consistent with introgression between P3 and either P1 (ABBA < BABA, negative *D* values) or P2 (ABBA > BABA, positive *D* values). In each scenario, P1 and P2 were represented by the historic (AlexOld) and present-day (AlexNew) samples, respectively, and as P3, we used multiple populations from Alaska (ATQ, BARD, DEN, IMN, MIL, MNT and SAG), as well as an artificial population formed by 10 individuals from the Alaska genetic cluster (admixture membership >0.99).

2.9 | Flow cytometry

Flow cytometry was run to estimate *C. tetragona*'s genome size and check the ploidy of a few populations on Disko Island, Greenland. We cut up silica dried *C. tetragona* and *Solanum lycopersicum* (genome size standard variety 'Stupické polní rané' grown from seed from the Doležel lab (Temsch et al., 2022)) leaf tissue in our nuclei extraction buffer, washed and fixed the nuclei before dying with propidium iodide. Flow cytometry was run on a Cytoflex flow cytometer in the UBC Biomedical Research Center. Our detailed flow cytometry protocol can be found in Appendix S1 – Flow cytometry.

3 | RESULTS

3.1 | SNP calling

Variant calling was performed on GBS data for 371 *Cassiope* samples from 36 populations (Figure 1b, Table S3), resulting in a total of 559,815 variant sites. The 10 *C. mertensiana* individuals were only used as an outgroup for the ABBA-BABA tests and in demographic modelling because read mapping to the de novo *C. tetragona* reference genome was poor for many of these individuals (missing data amounts are listed in Table S4).

3.2 | Population structure

Population structure analyses were based on 9285 SNPs and 330 individuals after filtering. In PCAs, approximately 27% of genetic variation was associated with the first principal component (PC1) separating ssp. *saximontana* (purple) from ssp. *tetragona* (Figure 4a, Table S6). Kluane Lake, Yukon (blue) individuals are apparent hybrids between the two subspecies. Approximately 6% of the variation occurs along PC2 separating Russia (red) from the rest of the Arctic. There appear to be a few individuals from Greenland (orange) found between Russia (red) and the rest of the Arctic. Despite the low amount of variation explained by PC3 (2.63%) and PC4 (1.63%), PC3 separates Alaska (blue) from Russia, Europe and Greenland (red/yellow/orange) and PC4 splits Greenland (orange) and Europe (yellow) in opposite directions from Alaska and Russia. We also ran the PCA with a MAF = 5 and MAC = 2 (Figure S5) and found the same clustering patterns with very different SNP filtering.

ADMIXTURE clustering results for $K = 1$–14 are consistent across multiple tests (with $K = 5$ consistently showing the last large drop in CV scores; Table S7). At $K = 5$, ADMIXTURE results suggest unique clusters currently in BC, Russia, Alaska, Europe and Greenland. Populations in Northern Canada (NWT, Baffin and Ellesmere) appear to have mixed ancestry from Alaska and Europe/Greenland (Figure 1a). SplitsTree and TreeMix plots supported the PCA and ADMIXTURE population structure (Appendix S2 – SplitsTree, TreeMix).

3.3 | Population statistics (Nucleotide diversity (π), F_{ST} and genetic distance)

The windows and site-based approaches for calculating nucleotide diversity gave comparable results. The statistics were based on 26,350 SNPs and 330 individuals after filtering. A significant positive relationship was found between ice retreat times and the amount of diversity both when including individuals with subspecies *saximontana* ancestry and when not (without *saximontana*, $n = 31$, $r^2 = 0.31$, $p = 0.00032$, Figure S8). This relationship was

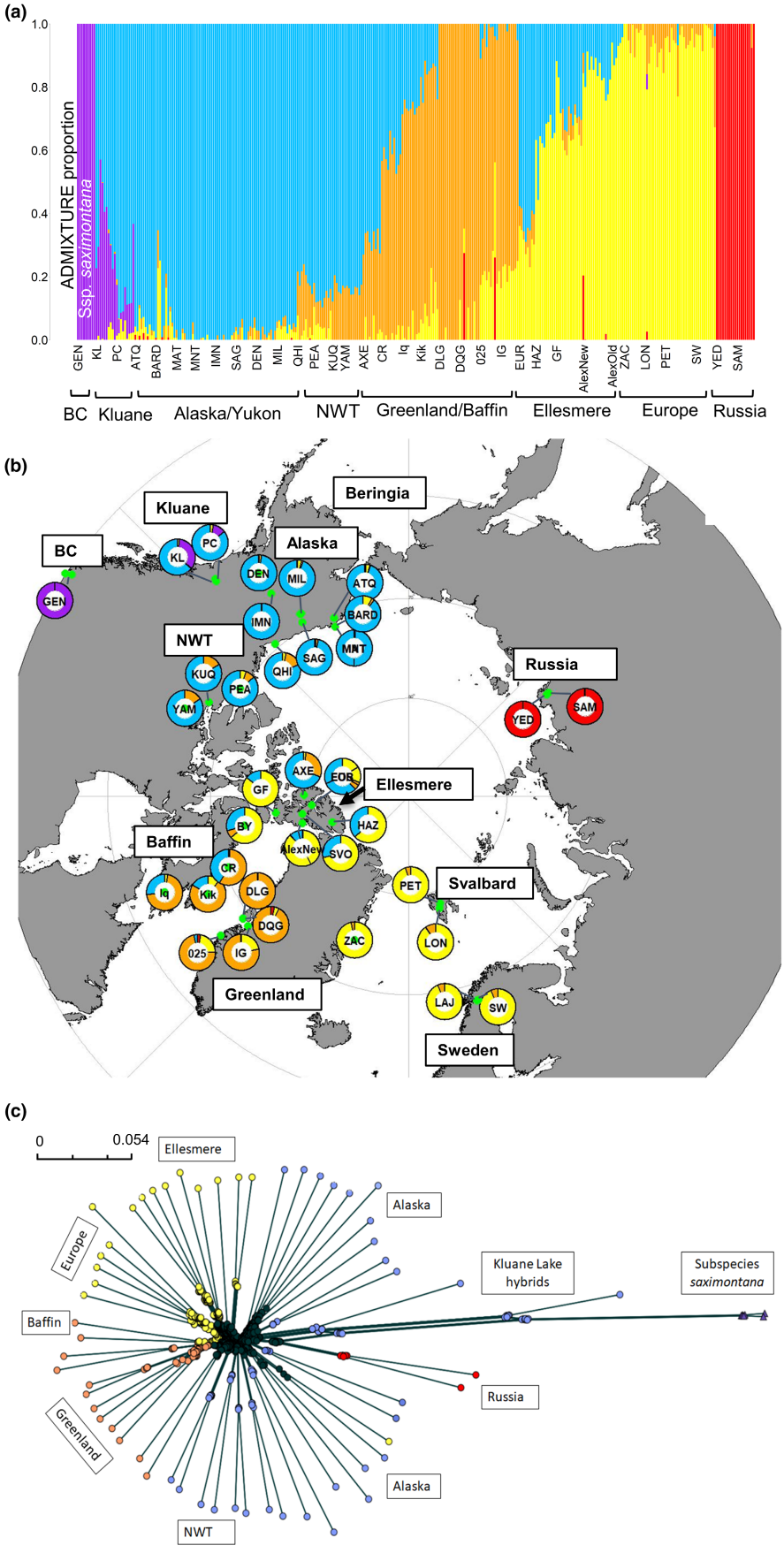

FIGURE 1 Population structure. Relationship between populations using ADMIXTURE results based on 9285 SNPs and 330 individuals. (a) Barplot shows individual genotypes as vertical bars ordered by populations and grouped by genetic cluster. Colours represent the proportion of that individual's genotype associated with each cluster (Top: $K = 5$, CV error = 0.262). (b) Geographic map of collection sites. Each site is represented by a pie chart that is colour-coded by its ADMIXTURE groups ($K = 5$). See Table S3 for information about all sample locations (PlotSvalbard package (Vihtakari, 2020)). (c) SplitsTree neighbor-net tree network built with the uncorrected P distance using two individuals randomly selected from every population (see Table S3 for collection location of each 3-letter code). Colours based on results from $K = 5$ ADMIXTURE groups. Network based on 25 572 SNPs and 60 individuals. Scale bar shows the uncorrected P distance.

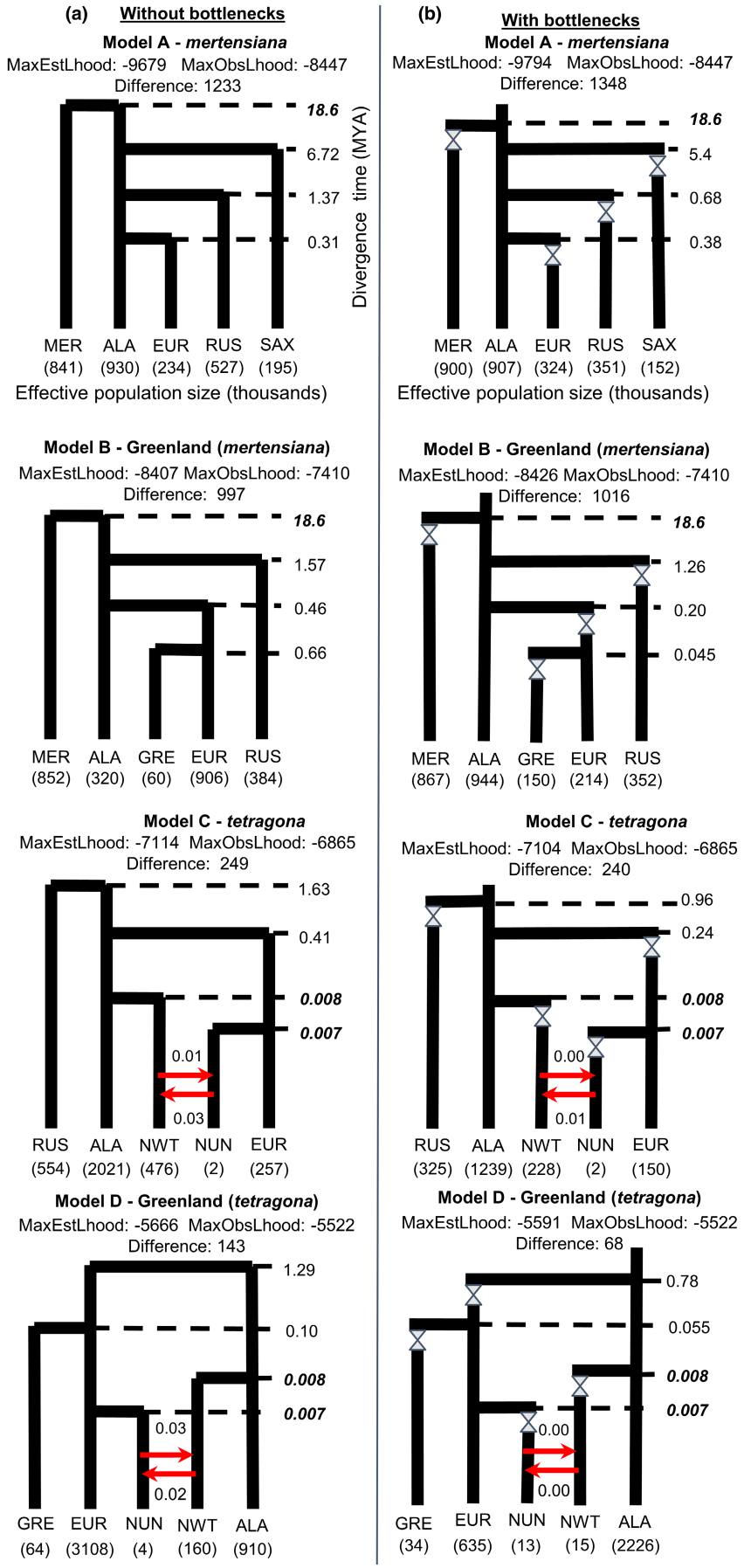

FIGURE 2 Coalescent demographic analysis results from fastsimcoal2. The four models that were run in fastsimcoal2 are shown with the best fit values not including (a) and including (b) a bottleneck during the population founding event. The four models (A-D) are independent and not compared to each other. Models A and B use the fixed time of 18.6 MYA for the split between *Cassiope tetragona* and *C. mertensiana* (MER). Models C and D use the fixed time of the ice retreat from northern Canada (Nunavut (NUN) and the Northwest Territories (NWT)) to be 7-8 thousand years ago. Models A and C do not include Greenland (GRE) (due to computational limitations) while models B and D remove one of the populations in models A and C to incorporate Greenland.

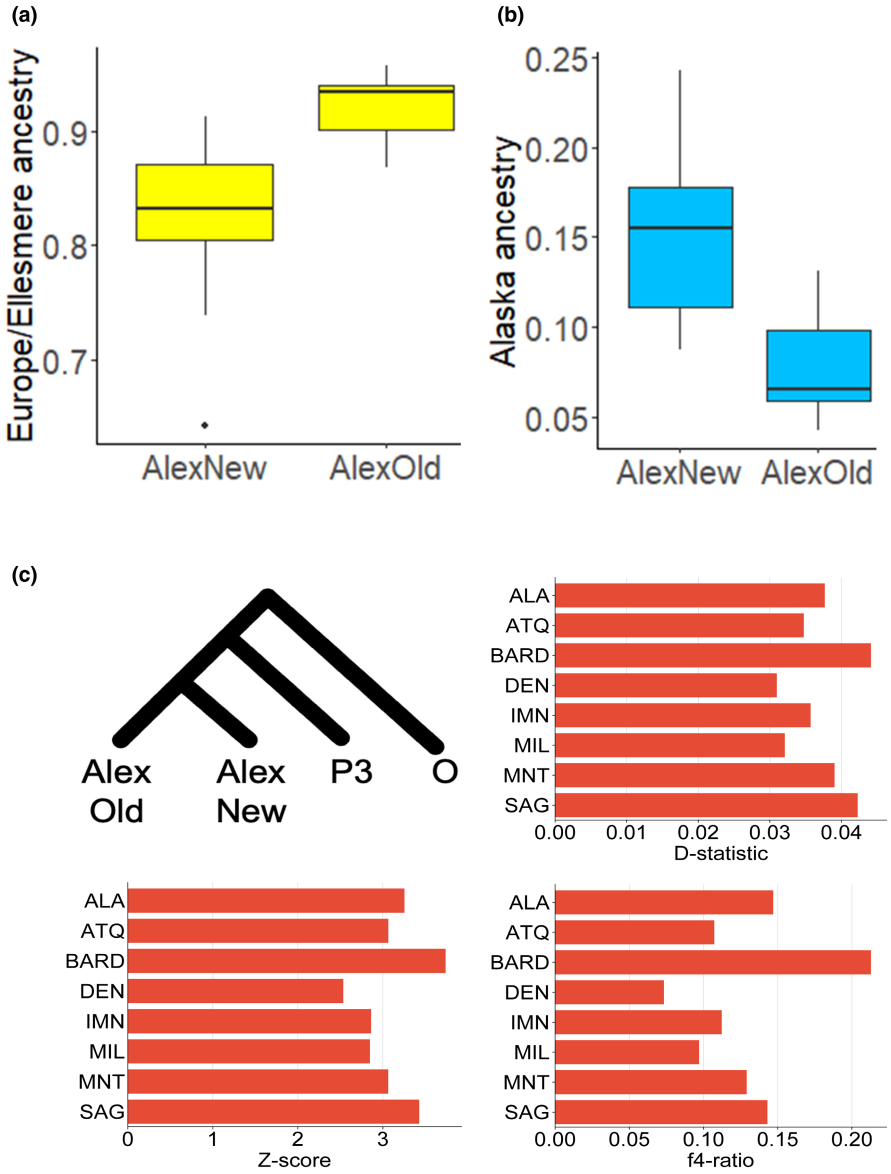

FIGURE 3 Comparison between historic and present-day samples. Samples of *C. tetragona* from Alexandra Fiord, Ellesmere Island, Canada including living individuals (AlexNew) and historic samples (AlexOld) from the ice edge of the retreating Twin Glacier. Barplots of ADMIXTURE ancestry proportions for the 23 individuals from Alexandra Fiord (based on 9285 SNPs from 330 individuals, $K=5$; see Figure 1a). (a) Relative proportions of European ancestry observed in the present-day and historic samples (Mann-Whitney-Wilcoxon Test $p=0.0023$). (b) Relative proportions of Alaskan ancestry observed in the present-day and historic samples (Mann-Whitney-Wilcoxon Test $p=0.0046$). (c) D-statistic, *Z*-score, and f4-ratio for admixture scenarios tested with the ABBA-BABA test. The tree shows the phylogenetic relations assumed. In each scenario, P1 and P2 were represented by the historic and present-day samples, respectively, and as P3 we used multiple populations from Alaska (ATQ, BARD, DEN, IMN, MIL, MNT, and SAG) as well as an artificial population formed by 10 individuals from the Alaska genetic cluster (ALA; admixture membership >0.99).

driven by the higher diversity and older populations in Beringia because when samples from Beringia were removed the relationship was no longer significant ($n=31$, $r^2=0.072$, $p=0.095$) (Figure 5a, Table S8).

Generally, low F_{ST} values (near zero) indicate largely random mixing among populations, while values closer to one are indicative of isolated populations connected by little gene flow. F_{ST} was calculated for all the geographically separated sampling locations (Figure 5b, for all *C. tetragona* populations listed in Table S3) and for the regional ADMIXTURE clusters for $K=5$ (Figure S6a). As expected, the largest F_{ST} values were found when comparing the ssp. *tetragona* in the Arctic and the ssp. *saximontana* cluster in BC (GEN) ($F_{ST}\sim0.6$) (Figure 5b). Within spp. *tetragona*, the Russian populations (SAM and YED) showed the highest F_{ST} values (~0.3) in pairwise comparisons followed by the Disko Island in Greenland ($F_{ST}\sim0.2$; DLG, Table S3, dark green). The rest of the Arctic clusters (Alaska, Yukon, NWT, Nunavut, Greenland, Svalbard and Sweden) all had lower F_{ST} values ($F_{ST}<0.15$) when compared with geographically close populations than with populations further away. Inbreeding coefficients were not significantly different from zero in the majority of populations, implying that populations were in HWE (Figure S7). Genetic distance significantly increased with geographic distance in both the 'Alaska' and 'Europe clusters' but not in the 'Greenland' cluster (Mantel test *p*-values, Alaska $=1.0\times10^{-4}$, Europe $=0.0012$, Greenland $=0.8697$) (Figure S6b,c). There is a positive relationship between F_{ST} and geographic distance in kilometres in populations from Alaska east to western Ellesmere and from Europe west into western Ellesmere.

3.4 | Demographic analysis (fastsimcoal2)

Demographic models were compared to observed site frequency spectra built from 10,338 SNPs after filtering. In all the models tested, the major genetic groups diverged well before the current

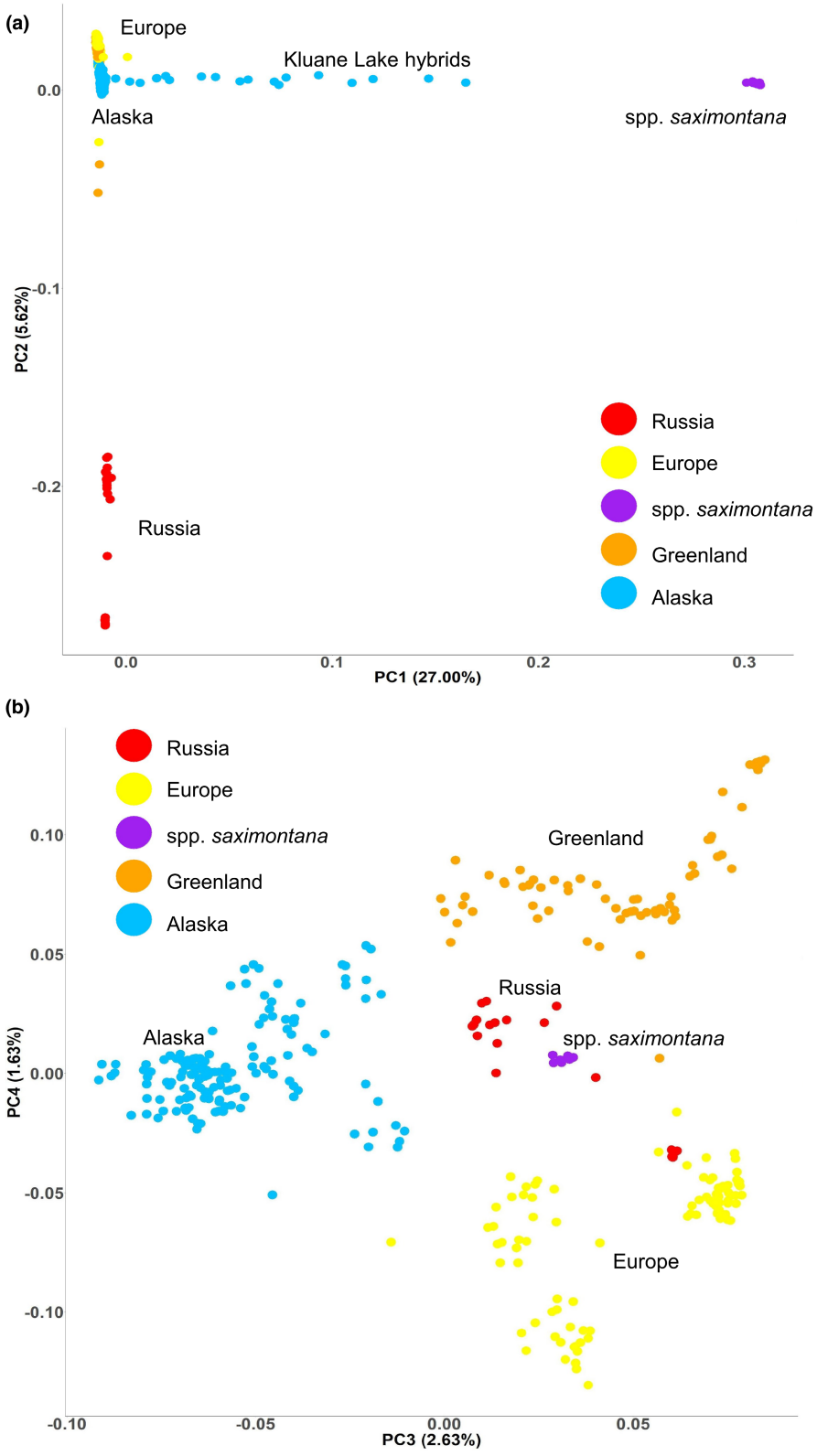

FIGURE 4 Principal component analysis (PCA) run on 9 285 variants and 349 *Cassiope tetragona* individuals. Individuals are colour-coded according to the $K = 5$ ADMIXTURE groupings (outgroup *C. mertensiana* is excluded). Percentage of variance accounted for by each component shown with the axis label. (a) The second principal component (PC2) versus PC1. (b) PC4 versus PC3.

interglacial (Figure 2 and Tables S9 and S10), supporting the presence of independent refugia in BC (known subspecies), Alaska, Russia, Europe and Greenland. Divergence times were mostly consistent with genetic differentiation values between groups (F_{ST}). In our models ssp. *saximontana* diverged first from Alaska spp. *tetragona* about 5.4–6.7 mya ($F_{ST} = 0.75$). Russia diverged next 1.0–1.6 mya ($F_{ST} = 0.22$), followed by Europe 200–1200 kya ($F_{ST} = 0.14$), and then, Greenland split from Alaska or Europe ~60–600 kya ($F_{ST} = 0.20$) (Figure 2, Tables S9 and S10). The lower F_{ST} value between Europe and Alaska may be explained by the ongoing gene flow through the Canadian hybrids. Estimates of divergence times between the four models were largely consistent

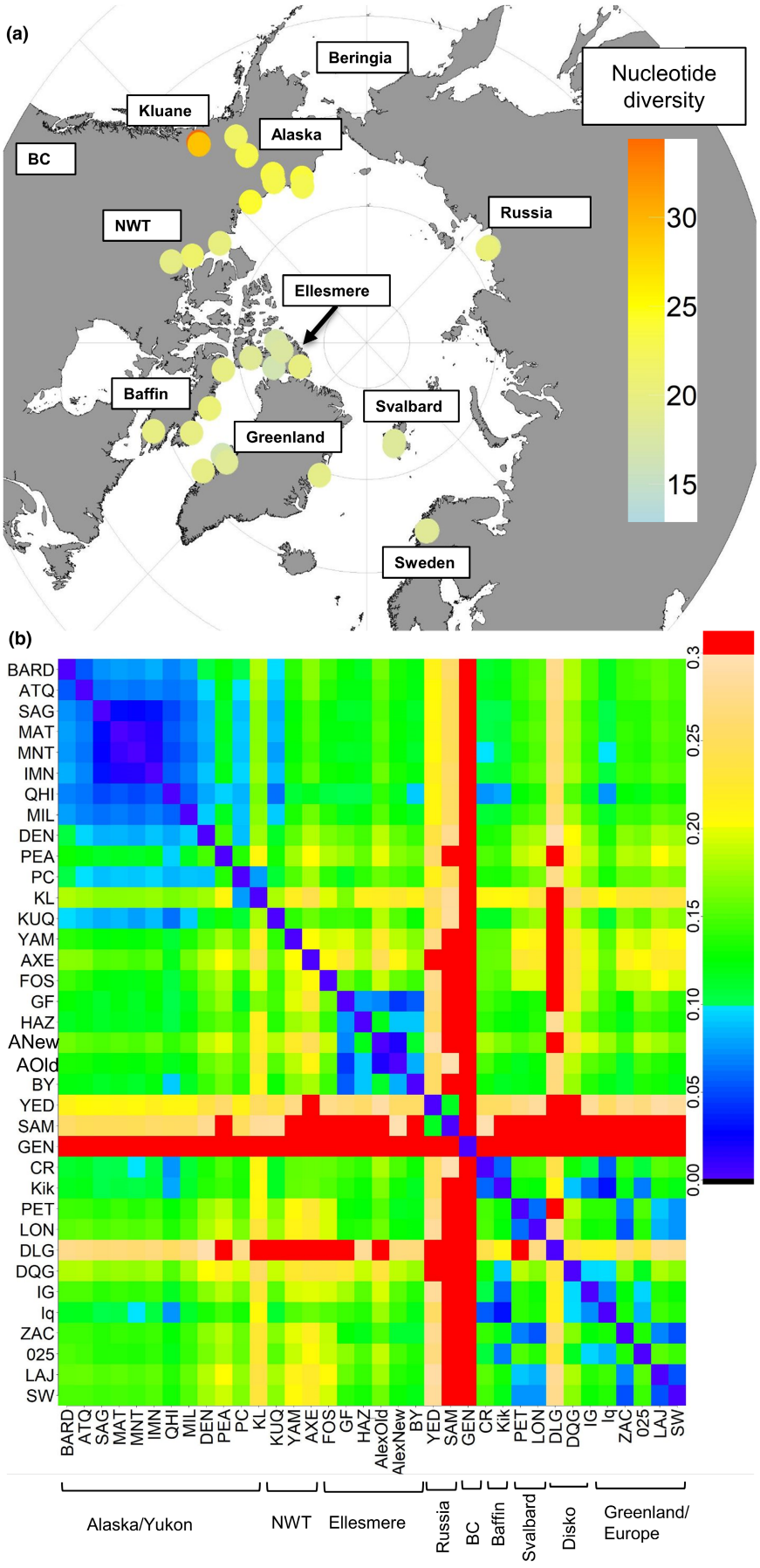

FIGURE 5 Population statistics. (a) Sum of window-based nucleotide diversity (π) mapped for each Arctic population (See Table S3 for names and location information of sample sites) (PlotSvalbard package (Vihtakari, 2020)). (b) F_{ST} values for all pairs of sampled locations (ordered by distance from Barrow, Alaska). F_{ST} values >0.3 are shown in red (British Columbia *saximontana* population = GEN).

within an order of magnitude even when different generation times were tested and bottlenecks were introduced for each population founding event (Tables S9 and S10).

3.5 | Flow cytometry (genome size)

We determined the genome size of both populations found on Disko Island, Greenland (DLG and DQG) to test whether a difference in ploidy was causing the high F_{ST} values between these very close populations. Flow cytometry results (Table S11) showed that *C. tetragona* has a haploid genome size of roughly 1.6 Gbp and there were no clear genome size differences between the two populations on Disko Island.

3.6 | Historic samples

Carbon dating estimated samples from under the Twin Glacier, Alexandra Fiord, Ellesmere Island (AlexOld), to be between 270 and 430 years old, and samples from Teardrop Glacier, Sverdrup Pass (SVO) to be 290–520 years old (exact values Table S12). Glaciers in North America began expanding in the mid-1500s (Forbes et al., 2020) and this expansion stopped in the late 19th to early 20th century. The historic samples therefore were likely living plants existing just prior to the glacial expansion. ADMIXTURE results comparing only the historic and present-day Alexandra Fiord populations (AlexNew and AlexOld) indicated that present-day samples have a higher 'Alaskan' ancestry compared to the historic populations (Figure 3a,b). A Mann–Whitney *U* test confirmed this significant difference in ancestry proportions. Historic and present-day ancestry comparison for the Alaskan/Ellesmere cluster had a $p = 0.0046$ and the European/Ellesmere cluster had a $p = 0.0023$. After filtering, we retained 29,824 variants for the ABBA-BABA tests. Results consistently show that modern samples from Alexandra Fiord share more alleles with populations from Alaska than historical samples (Figure 3c; Table S13).

4 | DISCUSSION

Using GBS data from 38 populations of *Cassiope tetragona*, including historic samples dated from the Little-Ice-Age, we were able to deduce the existence of multiple major Arctic glacial refugia. We found *C. tetragona* lineages that diverged from the Alaska/Beringia population 1000–1500 kya (Russia), 200–1200 kya (Europe) and potentially 60–600 kya (Greenland). These locations likely represent refugia for other Arctic species (Allen et al., 2015; Brochmann & Brysting, 2008; Marr et al., 2008; Marr et al., 2013). A previous study (Eidesen et al., 2007), using AFLP markers with limited resolution, proposed *C. tetragona* recolonized most of the Arctic after the LGM from a Alaskan/Beringian glacial refugium. Here, we found evidence supporting the presence of multiple Pleistocene refugia, from which *C. tetragona* recolonized the Arctic. We were also able to date the split between ssp. *saximontana* in southern British Columbia and ssp. *tetragona* to approximately 6 mya. In northern Canada, we observed mixing of Alaska, Europe and Greenland genetic clusters likely starting during the current interglacial (<10 kya). Moreover, using historical samples from the Little-Ice-Age (250–500 years old), we found that the Alaskan genotype is still spreading and expanding eastwards, as evidenced by the discovery of a higher proportion of Alaskan/foreign SNPs found in present-day Alexandra Fiord populations that are missing in the equivalent historic populations from a few hundred years ago.

4.1 | One or many refugia?

A review of Arctic plant genetic clusters using AFLPs predicted that both Alaska and Siberia (Beringia) were likely major refugia, but that much of North America/Europe was recolonized during the current interglacial from Alaska/Beringia (Eidesen et al., 2013). Consistent with an eastwards migration out of Beringia in the last 11 kya, we found nucleotide diversity (π) was highest in Alaska and lowest in the Canadian High Arctic. We also observed a significant relationship between the date of ice retreat and the amount of diversity, similar to that found by Pellissier et al. (2016). This suggests that recently deglaciated locations were recolonized not long after becoming ice free and/or have potentially smaller effective population sizes. *Cassiope tetragona* was previously assumed to have rapidly spread eastwards across North America, into Greenland and Europe recently (i.e., during the last interglacial; Eidesen et al., 2007). However, our population genomics and demographic analyses support at least two more ssp. *tetragona* refugia. In total, we found refugia in Alaska, Russia, Europe and Greenland, and a westward expansion of the last two. Populations in some of these refugia date back to the early Pleistocene (1.0–1.6 mya), while others formed more recently in the mid to late Pleistocene.

Many previous studies have proposed potential refugia locations based on species distribution models and past climate predictions (Pellissier et al., 2016), chloroplast DNA or AFLP data (Abbott et al., 2000; Alsos et al., 2009; Eidesen et al., 2007; Eidesen et al., 2013; Marr et al., 2008; Winkler et al., 2012), and pollen and fossil evidence (Brochmann & Brysting, 2008). Using chloroplast DNA data or AFLPs, some studies exploring Arctic-alpine refugia have found no evidence of a single Beringia refugium, instead supporting origins in Asia, Europe and/or southern North America for species such as *Bistorta vivipara*, *Sibbaldia procumbens*, *Oxyria digyna*, *Arabis alpina*, *Saxifraga oppositifolia* and *Ranunculus glacialis* (Allen et al., 2012; Allen et al., 2015; Koch et al., 2006; Marr et al., 2013; Ronikier et al., 2012; Winkler et al., 2012). The strong divide observed between Europe and North America in *Salix herbacea* using AFLPs could also point to separate refugia (Alsos et al., 2009). Our results support previous evidence that there were regions in Europe, Alaska and Russia that all had suitable, isolated conditions for Arctic/alpine species during the glaciations.

4.2 | Russian refugia

The Russia/Siberia cluster appears much older than the Canadian Arctic genetic clusters having diverged from other populations in the early Pleistocene (~1.0–1.6 mya) possibly from Beringia/Alaska or vice versa. The divergence times that we estimated were fairly consistent across models which used completely distinct calibration events (the outgroup divergence at 18.6 mya and the more recent range expansion into North America 7–9 kya). The relatively high F_{ST} values for the Russian populations imply little gene flow with other locations as previously noted (Eidesen et al., 2007, 2013). Sampling in Russia and Asia was significantly less thorough than in the rest of the world; so it is quite possible that the Russian cluster extends eastward, and hybrid genotypes do exist in western Beringia that were not sampled in this study.

4.3 | European refugia

European populations appear to have had their own refugia after the early to mid-Pleistocene (divergence from Alaskan populations happened 200–1200 kyrs ago). We modelled the founding of Europe from Alaska since F_{ST} values between Alaska and Europe were lower than between Russia and Europe. The decrease in European ancestry moving west into Northern Canada also supports a westward expansion from Europe after the LGM. During the current interglacial, as a result of this expansion, the European genotypes appear to have mixed with the Alaskan genotypes across Northern Canada creating some of the Canadian hybrids.

4.4 | Greenland refugia

West Greenland/Disko Island may also have been a younger, secondary refugium (~60–600 kyrs) that was originally populated from either the Alaskan or European refugia. We modelled a split from Europe since it had a higher European Admixture percentage. This genetic cluster is found currently in high proportions in western Greenland and Baffin Island and is more differentiated from populations in Europe and Alaska than the rest of Northern Canada. A clear division was reported between eastern and western Greenland in Eidesen et al. (2007), which was interpreted to be a result of the large Greenland Ice Sheet. However, there is fossil evidence of *C. tetragona* in Greenland during a previous interglacial so this division may be a result of a secondary, young refugia in West Greenland (Bennike & Böcher, 1994). This split pattern across Greenland has been seen in several other Arctic species including *Salix herbacea* (Alsos et al., 2009, Eidesen et al., 2013). While there is also older fossil evidence of *C. tetragona* in Greenland as far back as 2 mya (Eidesen et al., 2013), our demographic results suggest that the current Greenland genetic group is younger (~60–600 kyrs). Other genetic groups in Greenland could be unrepresented in our study or became extinct.

Two populations were sampled on Disko Island in West Greenland (DLG and DQG) and surprisingly show entirely unique genotypes. The DQG population, only about 2.5 km away from DLG, is similar to the rest of Greenland and has relatively high F_{ST} (weighted 0.152, mean 0.110) when compared to DLG. This suggests either a lack of effective gene flow between populations on the island, which could result from some strong reproductive barriers, or historically separated populations that are now near each other (possibly introduced by humans from a location not sampled in this study (Ware et al., 2012)). At $K=10$, DLG forms its own unique ADMIXTURE cluster. We wondered if this unique population could be explained by a whole genome duplication and paralogy; however, based on flow cytometry, these populations have the same genome size.

While the main refugia were known to have had ice free regions during the glaciations (e.g. parts of Europe, Alaska, Russia and southern BC), other small refugia may have existed on nunataks in Greenland (Beatty & Provan, 2010; Westergaard et al., 2011). Although we estimate the approximate ages of various genetic clusters and their locations now, we cannot be sure of the corresponding refugium locations that would have had suitable conditions for plant survival during the glaciations. Based on the unique genotype on Disko Island, there is a chance that an isolated population of *C. tetragona* survived some of the late Pleistocene there or somewhere further south.

4.5 | Ellesmere/Baffin divide

We observed higher F_{ST} values between Baffin Island (south) and Ellesmere Island (north) than would have been expected from their proximity. These clusters appeared to share some genetic ancestry with Greenland and Europe clusters, respectively. However, they do not appear to share recent ancestry with each other, supporting the existence of a refugium in Europe and a young, secondary refugium in Greenland, as well as potential long-term barriers to gene flow between north and south Nunavut. The cause of the apparent lack of mixing between these geographically close populations is unclear, but it might be due to differences in timing of deglaciation between the Innuitian and Laurentide Ice Sheets on Ellesmere and Baffin Island, respectively (Batchelor et al., 2019). The barrier of the Greenland Ice Sheet may also have resulted in the European refugia founding Ellesmere in the north and an isolated West Greenland population founding Baffin in the south. A similar divide on Ellesmere Island was observed in the chloroplast sequences from *Saxifraga oppositifolia* (Abbott & Comes, 2004).

4.6 | Ongoing postglacial Canadian hybridization

A comparison of the Little-Ice-Age individuals (270–520 years old) with present-day populations from Ellesmere Island implies ongoing admixture between Alaskan and European genotypes in Northern Canada. Foreign SNPs appear to have been introduced to Alexandra Fiord in the last few hundred years. Based on the ADMIXTURE results

at $K=5$ and $K=10$, most of these recently introduced foreign SNPs have originated in Alaska and are still mixing with the European lineage present in northern Nunavut. ABBA-BABA tests confirmed that extant populations carry more Alaskan ancestry than historical populations. Possible explanations for this ongoing mixing include (1) dominant wind currents are still present and continue to facilitate gene flow from Alaska (Hole & Macias-Fauria, 2017; Kling & Ackerly, 2021), (2) glacial retreat continues to provide space for foreign genotypes to establish (Losapio et al., 2021), (3) sexual reproduction/recombination may have been less frequent during the Little-Ice-Age due to shorter colder growing seasons, and (4) changing abiotic barriers continue to allow foreign genotypes to establish (Billings, 1987; Lewontin & Birch, 1966; Malcolm et al., 2002; McGraw, 1985). In addition, the increased Alaskan ancestry in Alexandra Fiord in the last few centuries suggests that the genotype from Alaska is spreading faster than the other genotypes, gradually replacing them.

To determine whether missing data caused this difference between historic and present-day samples (Ewart et al., 2019), we filtered out any individuals with missing data at more than 10% of SNPs (Results Table S4). We observed slightly higher homozygosity in the two historic populations, which could be a result of allele dropout due to DNA degradation and more variation at restriction enzyme sites preventing DNA digestion (Andrews et al., 2016). However, this seems unlikely to completely explain the clear drop in frequency of Alaskan SNPs in the historic samples.

5 | CONCLUSIONS

We provide strong evidence for independent Pleistocene refugia in central/western North America, Beringia, Russia, Europe and possibly a secondary, younger refugium in West Greenland. While Beringia was a refugium for North America, it appears Europe had its own unique lineage that also recolonized North America from the east during the Holocene. Multiple early to mid-Pleistocene refugia for *C. tetragona* may indicate that other Arctic plant species likely also survived the glaciations in similar locations; something future studies should investigate with high-resolution genomic data. The Arctic biome is logistically challenging to study and currently lacks many of the genomic resources available for many lower latitude species. With the Arctic climate warming at four times the global rate (Rantanen et al., 2022), Arctic species are the first to deal with climatic shifts making it important to record a baseline that may be rapidly shifting as well as to help predict potential future responses in lower latitude species (Colella et al., 2020). Collaborative scientific networks, such as the International Tundra Experiment (Henry et al., 2022), allow widespread sampling and data collection for genomic studies, reducing some of the logistical constraints.

AFFILIATIONS

[1]Department of Botany, University of British Columbia, Unceded xʷməθkʷəy̓əm (Musqueam) Territory, Vancouver, British Columbia, Canada
[2]Department of Geography, University of British Columbia, Unceded xʷməθkʷəy̓əm (Musqueam) Territory, Vancouver, British Columbia, Canada
[3]Centre for Arctic Knowledge and Exploration, Canadian Museum of Nature, Ottawa, Ontario, Canada
[4]Norwegian Institute for Nature Research, Trondheim, Norway
[5]Terrestrial Ecology Section, Department of Biology, University of Copenhagen, Copenhagen, Denmark
[6]WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland
[7]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland
[8]Département des Sciences de l'Environnement, Université du Québec-à-Trois-Rivières et Centre d'Études Nordiques, Trois-Rivières, Québec, Canada
[9]School of GeoSciences, University of Edinburgh, Edinburgh, UK
[10]Department of Biological Science, Florida International University, Miami, Florida, USA
[11]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[12]Geography Department, Humboldt-Universität zu Berlin, Berlin, Germany
[13]Government of Northwest Territories, Yellowknife, Canada
[14]Biology Department, Grand Valley State University, Allendale, Michigan, USA
[15]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
[16]Gothenburg Global Biodiversity Centre, Gothenburg, Sweden
[17]University of Texas at El Paso, El Paso, Texas, USA
[18]Department of Bioscience and Arctic Research Centre, Aarhus University, Roskilde, Denmark
[19]Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark

ACKNOWLEDGEMENTS

We thank Hana Christoffersen, Ed Rastetter, Erin Gleeson, Gaius R. Shaver (Arctic LTER: NSF grant DEB-1637459) for sampling *Cassiope* tissue from their sites, Armando Moreno Geraldes for help and guidance regarding the GBS data analysis and Rob Elphinstone for comments regarding the study. This research was enabled in part by WestGrid (https://www.westgrid.ca/) and Compute Canada (www.computecanada.ca). The widespread sample collection was enabled by the International Tundra Experiment network (https://www.gvsu.edu/itex/), US National Science Foundation 1504224 (to RDH), the peoples of Utqiagvik and Atqasuk, BC Parks, Canadian Museum of Nature, 2017 Students on Ice Arctic Expedition and Canada C3 expedition. Sampling permit information for all samples can be found in Table S3. Funding for this project came from GenomeBC (sequencing), ArcticNet, Natural Sciences and Engineering Research Council of Canada (to GHRH) with logistic support from Polar Continental Shelf Program and Royal Canadian Mounted Police (to GHRH) (Ellesmere field collections), University of British Columbia, Weston Family Foundation, Natural Sciences and Engineering Research Council of Canada (CGSM) and Vanier (funding for CE). The Strategic Research Area: BECC – Biodiversity and Ecosystem services in a Changing Climate (to MPB and RGB) for financing the monitoring programme at Latnjajaure Field Station.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

Sequence data are available in the GenBank/SRA database under accession number SUB11222726 and BioProject ID PRJNA824830. Detailed data processing notes and scripts can be found here: https://github.com/celphin/Population_genomics_Cassiope.

ORCID

Cassandra Elphinstone https://orcid.org/0000-0002-2968-1431

REFERENCES

Abbott, R. J., & Brochmann, C. (2003). History and evolution of the arctic flora: In the footsteps of Eric Hultén. *Molecular Ecology*, *12*(2), 299–313. https://doi.org/10.1046/j.1365-294X.2003.01731.x

Abbott, R. J., & Comes, H. P. (2004). Evolution in the Arctic: A phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (Purple saxifrage). *New Phytologist*, *161*(1), 211–224. https://doi.org/10.1046/j.1469-8137.2003.00953.x

Abbott, R. J., Smith, L. C., Milne, R. I., Crawford, R. M. M., Wolff, K., & Balfour, J. (2000). Molecular analysis of plant migration and refugia in the Arctic. *Science*, *289*(5483), 1343–1346. https://doi.org/10.1126/science.289.5483.1343

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*(1), 246. https://doi.org/10.1186/1471-2105-12-246

Allen, G. A., Marr, K. L., McCormick, L. J., & Hebda, R. J. (2012). The impact of Pleistocene climate change on an ancient arctic-alpine plant: Multiple lineages of disparate history in *Oxyria digyna*. *Ecology and Evolution*, *2*(3), 649–665. https://doi.org/10.1002/ece3.213

Allen, G. A., Marr, K. L., McCormick, L. J., & Hebda, R. J. (2015). Geographical origins, migration patterns and refugia of *Sibbaldia procumbens*, an arctic–alpine plant with a fragmented range. *Journal of Biogeography*, *42*(9), 1665–1676. https://doi.org/10.1111/jbi.12543

Alsos, I. G., Alm, T., Normand, S., & Brochmann, C. (2009). Past and future range shifts and loss of diversity in dwarf willow (*Salix herbacea* L.) inferred from genetics, fossils and modelling. *Global Ecology and Biogeography*, *18*(2), 223–239. https://doi.org/10.1111/j.1466-8238.2008.00439.x

Alsos, I. G., Müller, E., & Eidesen, P. B. (2013). Germinating seeds or bulbils in 87 of 113 tested Arctic species indicate potential for ex situ seed bank storage. *Polar Biology*, *36*(6), 819–830. https://doi.org/10.1007/s00300-013-1307-7

Alsos, I. G., Rijal, D. P., Ehrich, D., Karger, D. N., Yoccoz, N. G., Heintzman, P. D., Brown, A. G., Lammers, Y., Pellissier, L., Alm, T., Bråthen, K. A., Coissac, E., Merkel, M. K. F., Alberti, A., Denoeud, F., Bakke, J., & PhyloNorway Consortium. (2022). Postglacial species arrival and diversity buildup of northern ecosystems took millennia. *Science Advances*, *8*(39), eabo7434. https://doi.org/10.1126/sciadv.abo7434

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Batchelor, C. L., Margold, M., Krapp, M., Murton, D. K., Dalton, A. S., Gibbard, P. L., Stokes, C. R., Murton, J. B., & Manica, A. (2019). The configuration of northern hemisphere ice sheets through the quaternary. *Nature Communications*, *10*, 3713. https://doi.org/10.1038/s41467-019-11601-2

Beatty, G. E., & Provan, J. (2010). Refugial persistence and postglacial recolonization of North America by the cold-tolerant herbaceous plant *Orthilia secunda*. *Molecular Ecology*, *19*(22), 5009–5021. https://doi.org/10.1111/j.1365-294X.2010.04859.x

Bennike, O., & Böcher, J. (1994). Land biotas of the last interglacial/glacial cycle on Jameson land, East Greenland. *Boreas*, *23*(4), 479–487. https://doi.org/10.1111/j.1502-3885.1994.tb00615.x

Bergsma, B. M., Svoboda, J., & Freedman, B. (1984). Entombed plant communities released by a retreating glacier at Central Ellesmere Island, Canada. *Arctic*, *37*(1), 49–52.

Billings, W. D. (1987). Constraints to plant growth, reproduction, and establishment in Arctic environments. *Arctic and Alpine Research*, *19*(4), 357–365. https://doi.org/10.1080/00040851.1987.12002616

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brochmann, C., & Brysting, A. K. (2008). The Arctic – An evolutionary freezer? *Plant Ecology and Diversity*, *1*(2), 181–195. https://doi.org/10.1080/17550870802331904

Chong, Z., Ruan, J., & Wu, C.-I. (2012). Rainbow: An integrated tool for efficient clustering and assembling RAD-seq reads. *Bioinformatics*, *28*(21), 2732–2737. https://doi.org/10.1093/bioinformatics/bts482

Colella, J. P., Talbot, S. L., Brochmann, C., Taylor, E. B., Hoberg, E. P., & Cook, J. A. (2020). Conservation genomics in a changing Arctic. *Trends in Ecology & Evolution*, *35*(2), 149–162. https://doi.org/10.1016/j.tree.2019.09.008

Dalton, A. S., Margold, M., Stokes, C. R., Tarasov, L., Dyke, A. S., Adams, R. S., Allard, S., Arends, H. E., Atkinson, N., Attig, J. W., Barnett, P. J., Barnett, R. L., Batterson, M., Bernatchez, P., Borns, H. W., Breckenridge, A., Briner, J. P., Brouard, E., Campbell, J. E., … Wright, H. E. (2020). An updated radiocarbon-based ice margin chronology for the last deglaciation of the north American ice sheet complex. *Quaternary Science Reviews*, *234*, 106223. https://doi.org/10.1016/j.quascirev.2020.106223

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252. https://doi.org/10.1093/molbev/msr048

Dyke, A., & Prest, V. (1987). Late Wisconsinan and Holocene history of the Laurentide ice sheet. *Géographie Physique et Quaternaire*, *41*(2), 237–263. https://doi.org/10.7202/032681ar

Eidesen, P. B., Carlsen, T., Molau, U., & Brochmann, C. (2007). Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic. *Journal of Biogeography*, *34*(9), 1559–1574. https://doi.org/10.1111/j.1365-2699.2007.01719.x

Eidesen, P. B., Ehrich, D., Bakkestuen, V., Alsos, I. G., Gilg, O., Taberlet, P., & Brochmann, C. (2013). Genetic roadmap of the Arctic: Plant dispersal highways, traffic barriers and capitals of diversity. *New Phytologist*, *200*(3), 898–910. https://doi.org/10.1111/nph.12412

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Ewart, K. M., Johnson, R. N., Ogden, R., Joseph, L., Frankham, G. J., & Lo, N. (2019). Museum specimens provide reliable SNP data for population genomic analysis of a widely distributed but threatened cockatoo species. *Molecular Ecology Resources*, *19*(6), 1578–1592. https://doi.org/10.1111/1755-0998.13082

Excoffier, L., Marchi, N., Marques, D. A., Matthey-Doret, R., Gouy, A., & Sousa, V. C. (2021). fastsimcoal2: Demographic inference under complex evolutionary scenarios. *Bioinformatics*, *37*(24), 4882–4885. https://doi.org/10.1093/bioinformatics/btab468

Forbes, V., Ledger, P. M., Cretu, D., & Elias, S. (2020). A sub-centennial, Little Ice Age climate reconstruction using beetle subfossil data from Nunalleq, southwestern Alaska. *Quaternary International*, *549*, 118–129. https://doi.org/10.1016/j.quaint.2019.07.011

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. ArXiv:1207.3907 [q-Bio] http://arxiv.org/abs/1207.3907

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., … Pääbo, S. (2010). A draft

sequence of the Neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*(10), e1000695. https://doi.org/10.1371/journal.pgen.1000695

Havstrom, M., Callaghan, T. V., Jonasson, S., & Svoboda, J. (1995). Little ice age temperature estimated by growth and flowering differences between subfossil and extant shoots of *Cassiope tetragona*, an arctic heather. *Functional Ecology*, *9*(4), 650–654. https://doi.org/10.2307/2390157

Henry, G., Hollister, R., Klanderud, K., Björk, R., Bjorkman, A., Elphinstone, C., Jónsdóttir, I., Molau, U., Petraglia, A., Oberbauer, S., Rixen, C., & Wookey, P. (2022). The International Tundra Experiment (ITEX): 30 years of research on tundra ecosystems. *Arctic Science*, *8*, 550–571. https://doi.org/10.1139/as-2022-0041

Hole, G. M., & Macias-Fauria, M. (2017). Out of the woods: Driftwood insights into Holocene pan-Arctic sea ice dynamics. *Journal of Geophysical Research: Oceans*, *122*(9), 7612–7629. https://doi.org/10.1002/2017JC013126

Hou, Y., Nowak, M. D., Mirré, V., Bjorå, C. S., Brochmann, C., & Popp, M. (2016). RAD-seq data point to a northern origin of the arctic–alpine genus Cassiope (Ericaceae). *Molecular Phylogenetics and Evolution*, *95*(Supplement C), 152–160. https://doi.org/10.1016/j.ympev.2015.11.009

Hughes, A. L. C., Gyllencreutz, R., Lohne, Ø. S., Mangerud, J., & Svendsen, J. I. (2016). The last Eurasian ice sheets – A chronological database and time-slice reconstruction, DATED-1. *Boreas*, *45*(1), 1–45. https://doi.org/10.1111/bor.12142

Hultén, E. (1937). Outline of the history of arctic and boreal biota during the quaternary period: Their evolution during and after the glacial period as indicated by the equiformal progressive areas of present plant species. J. Cramer.

Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics (Oxford, England)*, *14*(1), 68–73.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Ikeda, H., Eidesen, P. B., Yakubov, V., Barkalov, V., Brochmann, C., & Setoguchi, H. (2017). Late Pleistocene origin of the entire circumarctic range of the arctic-alpine plant *Kalmia procumbens*. *Molecular Ecology*, *26*(20), 5773–5783. https://doi.org/10.1111/mec.14325

Jones, G. A., & Henry, G. H. R. (2003). Primary plant succession on recently deglaciated terrain in the Canadian high Arctic. *Journal of Biogeography*, *30*(2), 277–296. https://doi.org/10.1046/j.1365-2699.2003.00818.x

Kling, M. M., & Ackerly, D. D. (2021). Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(17), e2017317118. https://doi.org/10.1073/pnas.2017317118

Koch, M. A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C., & Mummenhoff, K. (2006). Three times out of Asia minor: The phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology*, *15*(3), 825–839. https://doi.org/10.1111/j.1365-294X.2005.02848.x

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*(7), 1812–1819. https://doi.org/10.1093/molbev/msx116

LaCava, M. E. F., Aikens, E. O., Megna, L. C., Randolph, G., Hubbard, C., & Buerkle, C. A. (2020). Accuracy of de novo assembly of DNA sequences from double-digest libraries varies substantially among software. *Molecular Ecology Resources*, *20*(2), 360–370. https://doi.org/10.1111/1755-0998.13108

Lewontin, R. C., & Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution*, *20*(3), 315–336. https://doi.org/10.1111/j.1558-5646.1966.tb03369.x

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Losapio, G., Cerabolini, B. E. L., Maffioletti, C., Tampucci, D., Gobbi, M., & Caccianiga, M. (2021). The consequences of glacier retreat are uneven between plant species. *Frontiers in Ecology and Evolution*, *8*, 520. https://doi.org/10.3389/fevo.2020.616562

Malcolm, J. R., Markham, A., Neilson, R. P., & Garaci, M. (2002). Estimated migration rates under scenarios of global climate change. *Journal of Biogeography*, *29*(7), 835–849. https://doi.org/10.1046/j.1365-2699.2002.00702.x

Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, *21*(2), 584–595. https://doi.org/10.1111/1755-0998.13265

Mallik, A. U., Wdowiak, J. V., & Cooper, E. J. (2011). Growth and reproductive responses of *Cassiope tetragona*, a circumpolar Evergreen shrub, to experimentally delayed snowmelt. *Arctic, Antarctic, and Alpine Research*, *43*(3), 404–409.

Marr, K., Allen, G., & Hebda, R. (2008). Refugia in the cordilleran ice sheet of western North America: Chloroplast DNA diversity in the Arctic-alpine plant *Oxyria digyna*. *Journal of Biogeography*, *35*, 1323–1334. https://doi.org/10.1111/j.1365-2699.2007.01879.x

Marr, K. L., Allen, G. A., Hebda, R. J., & McCormick, L. J. (2013). Phylogeographical patterns in the widespread arctic–alpine plant *Bistorta vivipara* (Polygonaceae) with emphasis on western North America. *Journal of Biogeography*, *40*(5), 847–856. https://doi.org/10.1111/jbi.12042

McGraw, J. B. (1985). Experimental ecology of Dryas Octopetala ecotypes: Relative response to competitors. *New Phytologist*, *100*(2), 233–241. https://doi.org/10.1111/j.1469-8137.1985.tb02775.x

Molau, U. (1997). Phenology and reproductive success in Arctic plants: Susceptibility to climate change. In W. C. Oechel, T. V. Callaghan, T. G. Gilmanov, J. I. Holten, B. Maxwell, U. Molau, & B. Sveinbjörnsson (Eds.), *Global Change and Arctic Terrestrial Ecosystems* (pp. 153–170). Springer. https://doi.org/10.1007/978-1-4612-2240-8_8

Moyers, B. T., Owens, G. L., Baute, G. J., & Rieseberg, L. H. (2017). The genetic architecture of UV floral patterning in sunflower. *Annals of Botany*, *120*(1), 39–50. https://doi.org/10.1093/aob/mcx038

O'Kane, K. (2018). *Plant succession in the High Arctic: Patterns and mechanisms*. University of British Columbia. https://doi.org/10.14288/1.0365968

Owens, G. L., Baute, G. J., & Rieseberg, L. H. (2016). Revisiting a classic case of introgression: Hybridization and gene flow in Californian sunflowers. *Molecular Ecology*, *25*(11), 2630–2643. https://doi.org/10.1111/mec.13569

Pellissier, L., Eidesen, P. B., Ehrich, D., Descombes, P., Schönswetter, P., Tribsch, A., Westergaard, K. B., Alvarez, N., Guisan, A., Zimmermann, N. E., Normand, S., Vittoz, P., Luoto, M., Damgaard, C., Brochmann, C., Wisz, M. S., & Alsos, I. G. (2016). Past climate-driven range shifts and population genetic diversity in arctic plants. *Journal of Biogeography*, *43*(3), 461–470. https://doi.org/10.1111/jbi.12657

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, *2*, e431. https://doi.org/10.7717/peerj.431

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, *3*(1), 168. https://doi.org/10.1038/s43247-022-00498-3

Rayback, S. A., & Henry, G. H. (2006). Reconstruction of summer temperature for a Canadian high Arctic site from retrospective analysis

of the dwarf shrub, *Cassiope tetragona*. *Arctic, Antarctic, and Alpine Research*, *38*(2), 228–238.

Ronikier, M., Schneeweiss, G. M., & Schönswetter, P. (2012). The extreme disjunction between Beringia and Europe in *Ranunculus glacialis* s. L. (Ranunculaceae) does not coincide with the deepest genetic split – A story of the importance of temperate mountain ranges in arctic–alpine phylogeography. *Molecular Ecology*, *21*(22), 5561–5578. https://doi.org/10.1111/mec.12030

Stroeven, A. P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., Goodfellow, B. W., Harbor, J. M., Jansen, J. D., Olsen, L., Caffee, M. W., Fink, D., Lundqvist, J., Rosqvist, G. C., Strömberg, B., & Jansson, K. N. (2016). Deglaciation of Fennoscandia. *Quaternary Science Reviews*, *147*, 91–121. https://doi.org/10.1016/j.quascirev.2015.09.016

Temsch, E. M., Koutecký, P., Urfus, T., Šmarda, P., & Doležel, J. (2022). Reference standards for flow cytometric estimation of absolute nuclear DNA content in plants. *Cytometry, Part A*, *101*(9), 710–724. https://doi.org/10.1002/cyto.a.24495

Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, T., & Pavoine, S. (2018). *Multivariate analysis of ecological data with ade4* (1st ed.). Springer.

Vihtakari, M. (2020). PlotSvalbard: PlotSvalbard—Plot research data from Svalbard on maps [R package version 0.9.2]. https://github.com/MikkoVihtakari/PlotSvalbard

Wang, Y., Pedersen, M. W., Alsos, I. G., De Sanctis, B., Racimo, F., Prohaska, A., Coissac, E., Owens, H. L., Merkel, M. K. F., Fernandez-Guerra, A., Rouillard, A., Lammers, Y., Alberti, A., Denoeud, F., Money, D., Ruter, A. H., McColl, H., Larsen, N. K., Cherezova, A. A., … Willerslev, E. (2021). Late quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature*, *600*(7887), 86–92. https://doi.org/10.1038/s41586-021-04016-x

Ware, C., Bergstrom, D. M., Müller, E., & Alsos, I. G. (2012). Humans introduce viable seeds to the Arctic on footwear. *Biological Invasions*, *14*(3), 567–577. https://doi.org/10.1007/s10530-011-0098-4

Weijers, S., Buchwal, A., Blok, D., Löffler, J., & Elberling, B. (2017). High Arctic summer warming tracked by increased *Cassiope tetragona* growth in the world's northernmost polar desert. *Global Change Biology*, *23*(11), 5006–5020. https://doi.org/10.1111/gcb.13747

Westergaard, K. B., Alsos, I. G., Popp, M., Engelskjøn, T., Flatberg, K. I., & Brochmann, C. (2011). Glacial survival may matter after all: Nunatak signatures in the rare European populations of two west-arctic species. *Molecular Ecology*, *20*(2), 376–393. https://doi.org/10.1111/j.1365-294X.2010.04928.x

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Winkler, M., Tribsch, A., Schneeweiss, G. M., Brodbeck, S., Gugerli, F., Holderegger, R., Abbott, R. J., & Schönswetter, P. (2012). Tales of the unexpected: Phylogeography of the arctic-alpine model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Molecular Ecology*, *21*(18), 4618–4630. https://doi.org/10.1111/j.1365-294X.2012.05705.x

Wullschleger, S. D., Breen, A. L., Iversen, C. M., Olson, M. S., Näsholm, T., Ganeteg, U., Wallenstein, M. D., & Weston, D. J. (2015). Genomics in a changing arctic: Critical questions await the molecular ecologist. *Molecular Ecology*, *24*(10), 2301–2309. https://doi.org/10.1111/mec.13166

Zeng, J., Zou, Y., Bai, J., & Zheng, H. (2002). Preparation of total DNA from "recalcitrant plant taxa.". *Acta Botanica Sinica*, *44*(6), 694–697.

Zheng, X., Gogarten, S., Laurie, C., & Weir, B. (2020). SNPRelate: Parallel Computing Toolset for Relatedness and Principal Component Analysis of SNP Data (1.22.0). Bioconductor version: Release (3.11) https://doi.org/10.18129/B9.bioc.SNPRelate

BIOSKETCH

Cassandra Elphinstone is a PhD Candidate in the Department of Botany at the University of British Columbia (UBC). Cassandra's research focuses on tundra molecular ecology. She studies Arctic and alpine plant responses to global change using genomics (DNA methylation, population structure, gene expression) and ecology. She is interested in the role of chromosome structure (tandem repeats, inversions, centromeres) in adaptation. Awards and networks include student representative for the International Tundra Experiment Steering Committee, Schulich Leader, Students on Ice Leacross Bern Alumni, Killam Doctoral Scholar, NSERC Vanier and Weston Family Doctoral Award in Northern Research.

Author Contributions: Cassandra Elphinstone, Marco Todesco, Greg H.R. Henry and Loren H. Rieseberg conceived the ideas. Cassandra Elphinstone, Greg H.R. Henry, Paul C. Sokoloff, Annika Hofgaard, Casper T. Christiansen, Esther R. Frei, EG, Esther R. Frei, Esther Lévesque, Gergana N. Daskalova, HLC, HT, Isla H. Myers-Smith, Jacob A. Harris, Julia Boike, Jeffery M. Saarela, Joachim Obst, Julia Boike, Karin Clark, Katie MacIntosh, Katlyn R. Betway-May, Mats P. Björkman, Michael L. Moody, Niels Martin Schmidt, Per Molgaard, Robert D. Hollister, Robert G. Björk, Roger D. Bull, Sofie Agger, Marco Todesco, Vincent Maire and WCase collected the field samples (see Table S3 for specifics) and edited the initial draft. Cassandra Elphinstone and Marco Todesco conducted the DNA extractions and GBS library preparation. Winnie Cheung provided support in the laboratory. Jean-Sébastien Légaré helped set up some of the pipelines on Compute Canada. Cassandra Elphinstone and Fernando Hernández analysed the data. Cassandra Elphinstone, Marco Todesco, Loren H. Rieseberg and Greg H.R. Henry led the writing.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Elphinstone, C., Hernández, F., Todesco, M., Légaré, J.-S., Cheung, W., Sokoloff, P. C., Hofgaard, A., Christiansen, C. T., Frei, E. R., Lévesque, E., Daskalova, G. N., Thomas, H. J. D., Myers-Smith, I. H., Harris, J. A., Saarela, J. M., May, J. L., Obst, J., Boike, J., Clark, K., … Rieseberg, L. H. (2024). Multiple Pleistocene refugia for Arctic Bell-Heather revealed with genomic analyses of modern and historic plants. *Journal of Biogeography*, *51*, 1958–1973. https://doi.org/10.1111/jbi.14961