Population genomic evidence for plant glacial survival in Scandinavia

Kristine Bakke Westergaard[1], Niklaus Zemp[2], Leo P. Bruederle[3], Hans K. Stenøien[4], Alex Widmer[5], Simone Fior[5]

[1] Norwegian institute for nature research, P.O. Box 5585 Torgarden, 7485 Trondheim, Norway, [2] ETH Zurich, Genetic Diversity Centre (GDC), Universitätstrasse 16, 8092 Zurich, Switzerland, [3] Department of Integrative Biology, CB171, University of Colorado Denver, POB 173363, Denver, CO 80217-3364, USA, [4] NTNU University Museum, Norwegian University of Science and Technology, 7491 Trondheim, Norway, [5] ETH Zurich, Institute of Integrative Biology, Universitätstrasse 16, 8092 Zurich, Switzerland

Correspondence: Kristine Bakke Westergaard, Norwegian institute for nature research, P.O. Box 5585 Torgarden, NO-7485 Trondheim, Norway. E-mail: Kristine.westergaard@nina.no

Keywords: Arctic-alpine phylogeography, *Carex scirpoidea*, ddRAD-seq, demographic inference, glacial survival, Pleistocene refugia

Running title: Phylogeography of *Carex scirpoidea*

ABSTRACT

Quaternary glaciations have played a major role in shaping the genetic diversity and distribution of plant species. Strong paleoecological and genetic evidence supports a postglacial recolonization of most plant species to northern Europe from southern, eastern, and even western glacial refugia. Although highly controversial, the existence of small *in situ* glacial refugia in northern Europe has recently gained molecular support. We used genomic analyses to examine the phylogeography of a species that is critical in this debate. *Carex scirpoidea* Michx ssp. *scirpoidea* is a dioecious, amphi-Atlantic arctic-alpine sedge that is widely distributed in North America, but absent from most of Eurasia, apart from three extremely disjunct populations in Norway, all well within the limits of the Weichselian ice sheet. Range-wide population sampling and variation at 5307 SNPs show that the three Norwegian populations comprise unique evolutionary lineages diverged from Greenland with high between-population divergence. The Norwegian populations have low within-population genetic diversity consistent with having experienced genetic bottlenecks in glacial refugia, and host private alleles likely accumulated in long-term isolated populations. Demographic analyses support one single, pre-Weichselian colonization into Norway from East-Greenland, and subsequent divergence of the three populations in separate refugia. Other refugial areas are identified in Northeast-Greenland, Minnesota/Michigan, Colorado and Alaska. Admixed populations in British Columbia and West-Greenland indicate postglacial contact. Taken together, evidence from this study strongly indicate *in situ* glacial survival in Scandinavia.

INTRODUCTION

A long-standing debate in arctic-alpine plant biogeography concerns the relative contributions of different Pleistocene refugia, geographical barriers, and dispersal in generating and maintaining patterns of distribution and genetic diversity in species (e.g. Blytt 1876; Brochmann *et al.* 2003; Eidesen *et al.* 2013; Pellissier *et al.* 2016; Provan & Bennett 2008; Shafer *et al.* 2010). During Quaternary glaciations, plant populations survived in suitable habitats in refugial areas differing geographically and demographically, with palaeoecological and genetic evidence supporting observations that many arctic-alpine plant species survived in macrorefugia outside the extents of European and American ice sheets (e.g. Birks 2008; Brochmann *et al.* 2003; Eidesen *et al.* 2007a; Eidesen *et al.* 2007b; Skrede *et al.* 2006). In Nordic biogeography, *in situ* glacial survival of arctic-alpine plant species within the maximum limits of the Weichselian ice sheet has traditionally been considered necessary (Blytt 1876, 1888; Sernander 1896; Warming 1888), particularly in order to explain biogeographic disjunctions across the North Atlantic Ocean (Dahl 1963; Gjærevoll 1990; Hultén 1937, 1958). Furthermore, the majority of arctic-alpine plant species in the North-Atlantic region lack obvious traits to promote long-distance dispersal; thus, post-glacial dispersal across the North Atlantic Ocean was considered virtually impossible (summarized in Brochmann *et al.* 2003). Recently, it has been suggested that even boreal cold-tolerant trees survived the last glacial maximum (LGM) in micro-environmentally favorable ice-free pockets in western Norway (Parducci *et al.* 2012; Quinzin *et al.* 2017). On the other hand, the alternative *tabula rasa* hypothesis of post-glacial immigration (Nathorst 1892; Ostenfeld 1926) has previously received overwhelming support from molecular studies, as the ability of arctic-alpine plant species to track their ecological niches over vast distances and barriers

has been extensively documented. For example, Westergaard *et al*. (2010) reported that very long-distance dispersal events best explain the extreme Beringian-Atlantic disjunctions in *Saxifraga rivularis*. Furthermore, other arctic-alpine plant species lacking obvious morphological adaptations for long-distance dispersal have crossed the Atlantic postglacially (e.g. Schönswetter *et al.* 2008; Schönswetter *et al.* 2007; Westergaard *et al.* 2011a), and colonized North Atlantic islands multiple times from different source areas (Alsos *et al.* 2015; Alsos *et al.* 2007).

At the heart of this discussion on Nordic discontinuous distributions are the so-called West-Arctic plant species, a subset of about 30 amphi-Atlantic vascular plant species occurring disjunctly across the North Atlantic Ocean (e.g. Dahl 1963; Hultén 1958; Nordal 1987). These species are widely distributed in North America, with few and isolated European occurrences, while being absent from areas eastwards in Eurasia and from the Central European mountains. Their contemporary European distribution lies within the area that was glaciated during the Weichselian glaciation, and their highly disjunct distribution was traditionally considered to provide evidence for *in situ* glacial survival. Compared to other species now present in Scandinavia which are also found elsewhere in Europe, the European populations of the West-Arctic species are expected to be less prone to genetic swamping as a consequence of admixture with postglacial immigrants from populations that survived in areas outside the ice sheets, and may thus still retain genetic footprints of *in situ* glacial survival. Furthermore, long-term isolation is expected to have resulted in strong genetic divergence among populations in different regions. If they are post-glacial immigrants from

North America, on the other hand, a higher level of genetic similarity is expected across the North Atlantic Ocean. A combination of glacial survival and recent immigration is also possible, resulting in co-occurrence of highly diverged genetic lineages and potential admixture. Phylogeographic histories of three rare, West-Arctic species (*Arenaria humifusa*, *Carex rufina*, and *Sagina caespitosa*) revealed distinct genetic groups on each side of the North Atlantic Ocean, consistent with the expectations of *in situ* glacial survival in North Europe (Westergaard *et al.* 2011a; Westergaard *et al.* 2011b). Interestingly, despite the lack of obvious morphological adaptations facilitating dispersal, genetic evidence indicates that all three species have considerable long-distance dispersal abilities.

In this paper we contribute to this debate by focusing on *Carex scirpoidea* Michx ssp. *scirpoidea* (Cyperaceae; hereafter *C.s*. *scirpoidea*), a species critical to the discussion of discontinuous distributional ranges. This arctic-alpine sedge has a wide, but island-like distribution in North America, including Beringia, but is absent from most of Eurasia apart from three extremely disjunct, small population groups in northern Norway **(Figure 1)**. It is the only dioecious West-Arctic species with sexual seed production – most individuals are either male or female, with a small percentage of plants (<5%) having a few pistillate or staminate flowers, respectively, with little if any regional variation. As such, the establishment of new populations of *C.s*. *scirpoidea* is challenged by the requirement of successful dispersal and establishment of two or more diaspores, making it the least likely post-glacially long-distance dispersed West-Arctic species. Two other West-Arctic, but monoecious *Carex* species (*Carex arctogena* and *C*. *macloviana*) have recently been studied

within the context of their bipolar disjunctions, which has been explained by long-distance dispersal (Márquez-Corro *et al.* 2017; Villaverde *et al.* 2015). Compared to other studied West-Arctic species, however, the dioecy of *C.s*. *scirpoidea* makes it a highly unlikely post-glacial, long-distance disperser across the Atlantic from North America to Norway.

In Norway, *C.s. scirpoidea* has long rhizomes, grows mostly in mats, and prefers sloping, eutrophic, herbaceous mountain vegetation influenced by seeping water, and moist heath on solifluction soils (Høiland 1986; Skifte 1985). It is only known from three populations, all in Nordland: (1) Solvågtind, Saltdal municipality, where four small subpopulations are known from 550-850 m a.s.l., (2) Kjelvatn, Ballangen municipality, where three small subpopulations are known from 830-1000 m a.s.l., and (3) Kjerringa, Gildeskål municipality, where two small subpopulations are known from 600-700 m a.s.l. Following Svendsen *et al*. (2004), none of these locations are found in areas known to be ice-free during the Last Glacial Maximum (LGM; 25,000-10,000 years ago). However, the geometry and vertical extent of the Scandinavian ice-sheet during the Weichselian have long been debated, and is thought to have been highly dynamic in space and time, intermittently exposing ice-free areas (Kolstrup & Olsen 2012). In North America, *C.s. scirpoidea* is predominantly caespitose (i.e., turf making), and occupies a variety of habitats including riparian zones, tundra, meadows, gravelly beaches, solifluction slopes, moist to dry rock slopes, and calcareous peatlands (Dunlop 2003; Dunlop & Crow 1999). Despite the ecological differences between Norway and North America, Norwegian plants have not been recognized as taxonomically distinct (Dunlop & Crow 1999). Because there are so few populations, and they occur in habitats

negatively affected by increasing temperature due to climate change, *C.s. scirpoidea* is listed as near threatened (NT) on the Norwegian Red-list (Solstad *et al.* 2015). From a conservation point of view, it is of interest to determine whether Norwegian populations represent ancient, cryptic refugia, or young and newly established founder populations.

Here we explore the phylogeography and population history of *C.s. scirpoidea* using double-digest restriction-site associated DNA (ddRAD) variation in population samples collected throughout much of the species' distribution. We specifically ask whether the three highly disjunct Norwegian populations are *in situ* glacial survivors, or originate from post-glacial, long-distance dispersal events from North America. It is also possible that glacial survival in Norway is confounded by recent immigration from North America, which would be expected to result in co-occurrence and potential mixing of divergent lineages in Norway.

MATERIALS AND METHODS

Sampling and DNA extraction

With collection permits from all local authorities (see Acknowledgements), we collected 306 individuals of *C.s. scirpoidea* from 24 populations sampled from across its distribution in North America and Norway **(Figure 1a, Table 1)**. Because *C.s. scirpoidea* is known to reproduce clonally (Bernard 1990), the individuals were sampled several meters apart, as far from each other as possible given the spatial extent of the population. Silica-dried leaf material was cut into small pieces, frozen, and ground at 30 sec/30 Hz in a mixer mill

(MM301; Retsch GmbH & Co.) using three ceramic beads (2.8 mm Zirconium Oxide beads, Omni International). Genomic DNA was extracted using the NucleoSpin® 8 Plant II kit (Mackerey-Nagel), following the manufacturer's instructions, incubating the samples using buffer PL1 for 30 min at 65°C. The amount of extracted DNA was quantified on a Qubit 2.0 using the HS Assay kit (Thermo Fisher Scientific).

ddRAD-seq library production

ddRAD-seq libraries were prepared using a customized version of the Peterson *et al*. (2012) protocol, including seven replicates for five individuals (i.e. 1-2 replicates per individual). Digestion of 130 ng high quality genomic DNA was done in a 50 µl reaction volume, first with 1 µl EcoRI-HF (20U) and 5 µl Buffer CutSmart (New England Biolabs, Inc.) for 30 min at 37°C, followed by 0.5 µl Taqα1 (New England Biolabs, Inc.) for 30 min at 65°C. The double digest was cleaned using 1X volume of Agencourt AMPure XP beads (Beckman Coulter, Inc.), before ligation in a 30 µl reaction volume using 1 µl P1 Adapter, 1 µl P2 Adapter, 3 µl T4 DNA ligase buffer 10X, and 1 µl T4 DNA ligase (400U/µl). Forty individually barcoded samples were multiplexed in a pooled library that was processed using 1X AMPure XP beads to remove unligated adapters. DNA concentrations were measured on a Qubit 2.0 using the HS Assay kit, and 570 bp libraries were selected using first 1.6X AMPure beads diluted 1:1.7, and subsequently 0.12X undiluted AMPure beads. Libraries were then washed with Dynabeads M-270 Streptavidin beads (Invitrogen) to select for P2-biotin labeled adapters. Unique Illumina indexes were ligated to each library during PCR amplification performed with a Phusion Polymerase Kit (New England Biolabs, Inc.) for seven cycles, and sets of two libraries

were multiplexed in each sequencing lane. Libraries were further cleaned using 1X AMPure XP beads and checked for DNA quantity on a Qubit 2.0 using the HS Assay kit, and for optimal fragment sizes on a Bioanalyzer using the HS DNA Assay kit. Libraries were sequenced in four lanes of 100 bp paired-end reads on an Illumina HiSeq 2500, adding 5% PhiX, at the Genomic Technologies Facility of the University of Lausanne, Switzerland.

Reference construction, read mapping, variant calling and filtering

Raw sequences were demultiplexed using the *process_radtags* component of *STACKS* v.1.26 (Catchen *et al.* 2013) before *de novo* assembly of a reference catalogue and variant calling was performed following the *dDocent* pipeline (Puritz *et al.* 2014). To build the reference catalogue, parameters were chosen to bypass most of sequencing errors and provide effective clustering of divergent alleles within loci. The *dDocent* pipeline concatenates forward and reverse reads to generate sets of unique sequences that are then clustered into reference contigs by the software *Rainbow* (Chong *et al.* 2012) and *CD-HIT* (Fu *et al.* 2012). Parameters that can substantially affect the resulting contigs include the number of reads set to retain unique sequences, i.e. K, the threshold similarity used by *Cd-hit* to cluster sequences, i.e. -c, as well as the individuals included as representative of the allelic diversity across the sampling. A larger set of unique sequences is retained for lower K values and by including a larger number of individuals, which on one hand maximizes the allelic diversity used to generate reference contigs, but on the other hand may lead to overall splitting of alleles belonging to the same locus if these are maintained separate from restrictive values of the -c parameter in *Cd-hit*. To assess the potential effect of these variables on our

197 population genetics data set, we produced four different reference catalogues using the
198 parameter combinations K=2 -c=0.8 and K=5 -c=0.9 on sets of reads from all individuals or
199 including only one randomly chosen individual for each of the 24 populations. We then
200 generated alternative population genetic data sets as described below and compared
201 estimates of observed and expected heterozygosity for each population as inferred using
202 Vcftools v. 0.1.11 (Danecek *et al.* 2011), which resulted in qualitatively consistent results
203 across populations and reference catalogues. The reference obtained with parameters K=5 -
204 c=0.9 on reads from all individuals, including 237'682 contigs, was retained for the
205 downstream analyses. To generate the population genetic data set, sequencing reads were
206 quality filtered with *Trimmomatic* v.0.33 (Bolger *et al.* 2014) to remove Illumina adapters,
207 bases below quality 20 at both ends of the reads, and low-quality bases at the end of the
208 reads assessed using a sliding window with average quality and window size set to 10 and 5,
209 respectively. Paired reads longer than 50 bp were mapped on the reference catalogue with
210 *BWA-MEM* (Li 2013) with default settings, while variant calling was performed using
211 *Freebayes* v. 1.1 (Garrison & Marth 2012) setting minimum quality and base quality to 5,
212 minimum repeat entropy to 1, and disabling prior expectations on observations. The
213 resulting variant call file (vcf) was conservatively filtered following recommendations of the
214 *dDocent* pipeline
215 (https://github.com/jpuritz/dDocent/blob/master/tutorials/Filtering%20Tutorial.md).
216 Sixteen individuals from eight different populations were removed from the data set
217 because of high proportion of missing data (i.e. >60%; see **Table 1**), and the vcf was filtered
218 to retain only variants present in at least 90% of individuals with minor allele frequency
219 (maf) of 0.05, and in 90% of individuals in each population. Additional filters to remove

variants resulting from sequencing errors, paralogs, multicopy loci, or artefacts of library preparation were applied as recommended in the pipeline. To increase the accuracy of the calls, and to reduce linkage disequilibrium challenges in the Structure analyses without having to thin data to a single SNP per locus, SNPs were haplotyped using the *rad_haplotyper* v 1.1.5 (Willis *et al.* 2017) leading to 5,307 SNPs and 2,796 haplotypes across 290 individuals. The final vcf was converted to other formats as needed using PGD Spider (Lischer & Excoffier 2012).

Genetic structure and admixture

A principal components analysis (PCA) for the 5,307 SNPs was conducted using SNPRelate (Zheng *et al.* 2012).

To explore the most likely number of genetically homogeneous groups (*K*) and overall structuring in the dataset, we ran genetic cluster algorithms in STRUCTURE 2.3.4 (Falush *et al.* 2003; Pritchard *et al.* 2000) using all 2,796 haplotypes and the admixture model without specifying any a priori population membership information. We first ran an explorative analysis using default settings for each value of K from 1-25 with a burn-in of 10 000 generations followed by 30 000 generations. According to Wang (2017), unbalanced population sizes and the assumption that populations are descendants of recent ancestral populations may yield inaccurate estimates of both *K* and assignment probabilities when using the default ancestry prior, the default value of alpha, and the correlated frequency model. Using the alternative population specific ancestry prior, a smaller initial alpha value

(alpha = 1/assumed optimal *K*), and the uncorrelated frequency model, STRUCTURE should yield more accurate inferences (Wang 2017). Thus, after identifying likely values of *K*, we ran STRUCTURE with ten replicate runs for *K*=1-10 using default settings and a non-random seed. We then ran STRUCTURE for ten replicate runs for *K*=1-10 using the alternative ancestry prior and an alpha value of 0.15 based on an optimal *K* around 7 (as inferred during the explorative analysis using the default parameters). The most likely number of evolutionary clusters *K*(s) was inferred in Structure Harvester (Earl & vonHoldt 2012), using both the ΔK statistic of Evanno *et al*. (2005) and calculations of Pr[X|*K*] (the probability of obtaining the genotype data X given K; Pritchard *et al.* 2000). To compare clustering results from STRUCTURE at multiple values of *K*, we aligned and visualized bar plots using the CLUMPAK (Cluster Markov Packager Across K) web server identifying distinct 'modes', i.e. groups of runs giving highly similar results in the space of possible solutions (Kopelman *et al.* 2015).

Genetic diversity

We calculated summary statistics for nucleotide diversity (π) and Fst among population pairs using vcftools (Danecek *et al.* 2011). As a measure of absolute differentiation, dxy (Nei & Li 1979) was calculated based on the allele frequencies as suggested in Smith & Kronforst (2013). Summary statistics were averaged per fragment assuming RAD fragment length of 200 bp, and mean dxy and Fst pairwise distance heatmaps and nucleotide diversity plots were generated and displayed graphically using *ggplot* in R. Population 10 (see **Table 1**) was excluded from the data set as only two individuals were sequenced.

To estimate the number of private alleles for single populations, we produced a separate vcf following the pipeline described above, except for the maf filter to retain alleles present in low frequency in the overall dataset. Using the seven replicates, the genotyping error rate per coverage class was estimated in Tiger (https://bitbucket.org/phaentu/tiger/wiki/Home) to be maximum 0.994%. Numbers of private alleles were inferred for each target population against all other individuals (metapopulation) by computing allele frequencies from genotype likelihoods in the popStat function of vcflib (https://github.com/vcflib/vcflib). Given the uneven number of individuals representing each population, ten individuals were randomly selected 100 times for the target population, and number of private alleles was averaged across replicates. Private alleles were inferred as SNPs with an allele frequency of <1% in the metapopulation, which accounts for the estimated genotyping error, and ≥5%, ≥10%, ≥20%, ≥90% or 100% in the target population. The three first frequencies correspond to observing one, two, or four alleles, while the two last correspond to near fixed or fixed alleles in the subsampled target population. To account for differences in population diversity, we corrected the number of private alleles by the ratio of π of the metapopulation and the target population. Populations 10, 17, and 22 (**Table 1**) were excluded as target populations as they included less than eight individuals.

Estimating demographic history of the Atlantic populations

Based on results from the phylogeographic analyses, two competing evolutionary scenarios may explain the history of the highly disjunct Norwegian populations **(Figure S1)**. In scenario 1, Norway was colonized twice from East Greenland, first by lineages that today constitute

the Kjelvatn (1) and Solvågtind (3) populations, and later by lineages that today constitute the Kjerringa (2) population. In scenario 2, Norway was colonized only once from East Greenland, that is, all known Norwegian populations resulted from one single colonization event. To evaluate the most likely colonization history of Norway by *C.s. scirpoidea*, we compared the two scenarios using an approximate Bayesian computing (ABC) approach as implemented in the DIYABC version 2.1.0 software (Cornuet *et al.* 2014). Summary statistics were derived from a merged East Greenland population with samples from Holm Bugt (4) and Mestersvig (5) treated as a single population, and from each of the Kjelvatn (1), Kjerringa (2) and Solvågtind (3) populations. We used RAD-locus diversities within populations, and Fst and Nei's distances between populations (all based on mean of complete distributions) as summary statistics to compare to simulated values for the two scenarios.

Several short trial runs (200K simulations) with increasingly wider prior ranges were performed, culminating in one long run (2 million simulations) using the following upper prior ranges (all lower prior ranges being equal to one): time since first colonization T1: 3 million generations, time since divergence of the Kjerringa population T2: 2 million generations, time since divergence of the Solvågtind and Kjelvatn population T2: 2 million generations. The simulated values of T1-T3 were independent of one another for the two scenarios. We assumed that colonization of Norway involved a period of bottlenecks (lasting x generations), where effective population size of Norwegian immigrant population was reduced to Nx individuals. The prior effective population sizes of the four populations ranged

from one to 100K (Kjelvatn and Solvågtind), and one to 3 million (Kjerringa and Grønland), respectively. During bottleneck(s) we assumed that the colonizing population(s) had no more than at most 50K individuals, and that this lasted for a maximum of 1K generations. We compared the posterior probabilities of the two scenarios by counting the number of times a given scenario was found among the 500 simulated data sets being most similar to the observed summary statistics (direct measure), as well as using a logistic regression approach described in Fagundes *et al.* (2007) and Beaumont (2008) using the 1000 simulations most similar to the observed data set.

RESULTS

Population assignment and admixture

In a principal component analysis (PCA) based on SNPs, the first two axes explained 31.1% of the variation in the data **(Figure 2)**. The resulting plot notably resembled a geographic map of the *C.s. scirpoidea* distribution, where the first axis explained a substantial amount of the genetic variation (21.5%) and clearly showed the populations arranged along an East-West axis. The second axis (9.6%) mainly separated the highly disjunct Colorado population (13) from the others, while the third axis (4.9%) separated the Norwegian populations Kjelvatn (1) and Solvågtind (3), but not Kjerringa (2) and the East Greenland populations (4, 5).

The Δ*K* analyses of the two STRUCTURE runs with default and custom parameter settings both identified *K*=2 as the most likely number of genetically homogeneous groups among

329 our 24 populations of *C.s. scirpoidea*. CLUMPAK confirmed that individual assignment to the
330 two groups was highly correlated across the STRUCTURE runs (r = 0.99; **Figure 3a**). One
331 group contained all of the Norwegian and Greenlandic populations and prevailed in the
332 populations from Minnesota and Michigan (the Eastern group), while the other group
333 contained the populations from Alaska, Yukon and prevailed in the populations from British
334 Columbia and Colorado (the Western group). The population from Northeast Canada was
335 divided between the two groups.

336

337 We explored the STRUCTURE results as inferred for higher K's to achieve resolution of the
338 Norwegian populations. This occurred at *K*=9, which corresponds to the *K*s with highest
339 likelihoods inferred from the highest mean value of Pr[X|*K*]. Using this estimator, the
340 STRUCTURE run with default parameter settings identified an optimal *K*=9, while the
341 STRUCTURE run with custom parameter settings identified alternative resolutions at an
342 optimal *K*=10 (major mode 6/10, minor modes 3/10 and 1/10; **Figure 3b-d**). Methods based
343 on mean likelihoods are known to be biased against lower *K* values and yield models which
344 may be over-parametrized with minor gene pools resulting in alternative results. Indeed,
345 these minor gene pools may not be biologically meaningful and should be interpreted with
346 caution. A combined interpretation of the results obtained for increasing *K* values showed a
347 hierarchical resolution of genetically homogeneous groups in the western range of the
348 distribution including Yukon and Alaska, in the distant relict population in Colorado, the
349 populations from Minnesota and Michigan, and in the eastern range including East
350 Greenland. Populations occurring at intermediate locations showed admixture with adjacent

groups, in particular, three populations from British Columbia showed considerable admixture with Yukon and Alaska, while a fourth population showed admixture between the Yukon-Alaska group and the Minnesota-Michigan group. Similarly, populations from East-Canada and West-Greenland shared large proportions of ancestry with populations from both Minnesota-Michigan and East-Greenland. For *K*=9, the Norwegian population Kjerringa (2) was part of the East Greenland group, while the two other Norwegian populations Kjelvatn (1) and Solvågtind (3) formed a distinct, genetically homogeneous group with little evidence of admixture from other groups. For *K*=10, the distribution of individual assignments for the major mode (6 out of 10 runs) resembled the *K*=9 results with two exceptions: the Norwegian population Kjelvatn (1) formed a distinct cluster, and a large proportion of the ancestry in West Greenland was attributed to a separate cluster. The minor mode (3/10) of *K*=10 identified the Norwegian populations as three distinct clusters; notably, Kjerringa (3) was not part of the same cluster as the East Greenland populations.

Genetic diversity and differentiation

Nucleotide diversity (π) varied considerably among geographic regions **(Figure 4)**. Mean π was lowest in the highly disjunct populations in Norway (Kjelvatn, 0.0005; Kjerringa, 0.0006; Solvågtind, 0.0004), and Colorado (0.0005). A heatmap of the dxy-values **(Figure S2)** displayed highest values of absolute divergence between populations in the eastern and western parts of the distribution area of *C.s. scirpoidea*, coinciding with the two STRUCTURE groups identified by the Δ*K* statistics, and further supports a split into two deep evolutionary groups in our dataset.

A similar geographic pattern of genetic differentiation is shown in the heatmap of Fst-values **(Figure S3)**, where highest Fst-values were found between populations from the Eastern and Western STRUCTURE groups identified by the Δ*K* statistics. Differentiation within the Eastern group was relatively high, and notably Fst increased from East Greenland to Norway. The Colorado population (13) also had high Fst values, while there was little differentiation between populations in Yukon and Alaska.

The corrected numbers of private alleles occurring at lower frequencies (thresholds ≥5% and ≥10%) were generally higher in populations from the Western group compared to the Eastern group **(Table S2, Figure 5, Figure S4a)**, and conspicuously high in the highly disjunct Colorado population (13). Within the Eastern group, the Norwegian populations had more private alleles than the East Greenland populations. Fixed or near fixed private alleles occurred only in populations from Norway (1 and 3) and Colorado (13) (**Table S2, Figure S4c-d**).

Demography of the Norwegian populations

Both the direct and logistic regression approaches yielded support for scenario 2 (i.e., one single colonization of Norway from East Greenland and subsequent divergence of the populations). In the direct approach, 88% of the 500 simulations most similar to observed data were made within the scenario 2 framework, while in the regression approach, 100% of

394 the 1000 most similar simulations were from scenario 2. The posterior estimates of the
395 parameters are given in **Table S1**. According to these ABC analyses, the first colonization of
396 Norway happened 170 000 generations ago (2.5% lower credible interval 91 400 generations
397 ago). North American *C.s. scirpoidea* is thought to have a lifespan of 10-20 years (Shackleford
398 2003, and references therein), so by applying a highly conservative generation time of one
399 year, the demographic analyses supports a pre-Weichselian (>115 000 years ago)
400 colonization of Norway.

401

402 **DISCUSSION**

403 *Glacial survival of* Carex scirpoidea ssp. scirpoidea *in Norway*

404 The three Norwegian populations of *C.s. scirpoidea* (populations 1-3) make up unique and
405 highly divergent evolutionary groups with low within-population genetic diversities and a
406 relatively high number of private alleles, as expected from the classic pattern of small
407 populations that have experienced genetic bottlenecks and drift in isolated *in situ* glacial
408 refugia (see e.g. Hewitt 2004). Importantly, absolute divergence (dxy) between the
409 Norwegian and East Greenland populations (populations 4 and 5) is comparable to levels of
410 divergence between other populations in the Eastern group, and the numbers of private
411 alleles in the Norwegian populations are higher compared to the East Greenland
412 populations. Notably, two Norwegian populations (1 and 3) host fixed private alleles.
413 Overall, this evidence contrasts with a scenario of Norwegian populations originating from a
414 postglacial recolonization from the Eastern group, in which case a decrease in genetic

415 divergence and number of private alleles is expected in comparison with the putative source
416 of colonization (i.e. East Greenland). Demographic analyses support a pre-Weichselian
417 colonization by *C.s. scirpoidea* into Norway from East Greenland, and subsequent divergence
418 of the three populations in separate refugia (**Figure S1 and Table S1**).

419

420 Our study provides the first genomic data consistent with *in situ* glacial survival of a vascular
421 plant species in mainland Scandinavia. Molecular evidence suggesting *in situ* glacial survival
422 in the East Atlantic region has previously been presented for three other West-Arctic
423 vascular plant species: *Arenaria humifusa*, *Sagina caespitosa*, and *Carex rufina*. However,
424 their refugial areas were most likely in known ice-free areas in the Arctic archipelago
425 Svalbard or East Greenland (*A*. *humifusa*), or possibly in southern Norway or even further
426 south (*S*. *caespitosa*; Westergaard *et al.* 2011b), or could not be elaborated further (*C*.
427 *rufina*; Westergaard *et al.* 2011a). Many other molecular studies have focused on more
428 common species and have demonstrated high dispersal capability and postglacial
429 immigration into northern Europe, leaving the glacial survival hypothesis superfluous (e.g.
430 Brochmann *et al.* 2003). Notably, *C.s. scirpoidea* has its only current European populations
431 well within the maximum limits of the Weichselian ice sheet, strongly limiting the possibility
432 of genetic swamping from conspecific, post-glacial immigrants that recolonized from refugial
433 areas outside the former ice sheets. Our results on *C*.*s*. *scirpoidea* increase our knowledge on
434 glacial refugia in the North Atlantic region by adding long-sought evidence of plant survival
435 within the maximum extent of the Weichselian ice sheet.

436

Although dioecy coupled with long distances and a narrow niche in Norway make *C.s. scirpoidea* a highly unlikely long-distance, post-glacial disperser, Bayesian clustering and principal coordinate analyses (**Figures 2, 3**) support a shared ancestry between the Norwegian and the East Greenland populations. The Kjerringa (2) population from Norway grouped with the East Greenland populations in all but one Structure run (*K*=10, minor mode 3/10), which suggested two possible evolutionary scenarios describing the colonization history of the Norwegian populations. Norway was either colonized twice from East Greenland, by lineages that today constitute the Kjelvatn (1) + Solvågtind (3) and the Kjerringa (2) population, or all three Norwegian populations resulted from a single colonization event followed by population divergence. Our demographic analyses yielded overwhelming support for the latter scenario; thus, we suggest that the STRUCTURE results reflect the stochastic variation of retained ancestry during the divergence of the Norwegian populations.

Where were the *in situ* glacial refugia for *C.s*. *scirpoidea* situated in Norway? Our results do not fit with the classic glaciation model of a thick, single-domed ice sheet covering most of Scandinavia at LGM, leaving no *in situ* refugia available for plants (Svendsen *et al*. 2004). However, the vertical extent of the ice at LGM has been reconstructed in a variety of models as dynamic, thin, multi-domed, and asymmetric ice sheets with available refugial areas (Arnold *et al.* 2002; Kolstrup & Olsen 2012; Linge *et al.* 2006; Olsen 1997). The hypothesis of such highly dynamic ice cover in space and time is coupled with findings of a unique and rare mitochondrial haplotype of spruce with a high frequency in western Norway, and chloroplast

459 DNA of pine and spruce in late-glacial lake sediments from the known ice-free Andøya
460 refugium in northwestern Norway, indicating LGM survival of boreal conifers in northern
461 Scandinavia (Parducci *et al.* 2012). In contrast to both pine and spruce and most other
462 molecular studies of rare and common vascular plants in the North Atlantic region (e.g. Alsos
463 *et al.* 2015; Alsos *et al.* 2007; Eidesen *et al.* 2013), our data provide no support for postglacial
464 dispersal of *C.s. scirpoidea* in Norway. In fact, there are no signs of recent admixture among
465 the three Norwegian populations as they form separate evolutionary groups with relatively
466 high Fst-values and a relatively high number of population-specific private alleles. It is
467 generally assumed that long-distance seed dispersal and establishment is important for the
468 survival of plant species, as it enhances species range expansion and migration during
469 climate change (e.g. Alsos *et al.* 2007; Nathan 2006), especially in dynamic landscapes with
470 high turnover of habitat patches (e.g. Hanski 1998). The lack of metapopulation dynamics
471 between the three relatively close-lying Norwegian populations (Kjelvatn – Solvågtind 170
472 km, Kjerringa – Solvågtind 50 km, Kjelvatn – Kjerringa 200 km) is striking, especially when
473 compared to other rare species that lack morphological adaptations to long-distance
474 dispersal, but with a demonstrated post-glacial contact across the North Atlantic Ocean
475 (Birkeland *et al.* 2017; Westergaard *et al.* 2011b). We do not demonstrate nor claim that *C.s.*
476 *scirpoidea* survived the entire Weichselian glaciation at its current locations in Norway, as its
477 *in situ* glacial refugia could have been located somewhere in the vicinity. It is plausible, for
478 example, that the species expanded to a larger distribution in Norway under more favorable
479 conditions during the peak warming of the Holocene thermal maximum (ca 8000-4000 y BP),
480 and subsequently experienced genetic bottlenecks when retreating into the current, small
481 stations. Alternative explanations for the genetic patterns of *in situ* glacial refugia of the

Norwegian populations would include a highly complex hypothesis of refugia outside the Weichselian ice sheet, followed by post-glacial dispersal into Norway with subsequent extinction in the glacial refugia and potentially also along the dispersal route. We find this to be a less parsimonious explanation for the patterns observed.

Other refugia and postglacial contact zones

The overall geographic structure of the postglacial genetic groups found in *C.s. scirpoidea* resembles the structure previously reported in other widespread arctic-alpine plant species. For *Cassiope tetragona* ssp. *tetragona*, genetic groups were identified in Siberia, Beringia, North Canada, East Canada/West Greenland, and East Greenland/Scandinavia (Eidesen *et al.* 2007b). For *Vaccinium vitis-idaea*, a similar geographical pattern was reported (Alsos *et al.* 2012), while for *Betula nana* s. lat. and *Vaccinium uliginosum*, the Beringian group extended more across Canada, and populations from West and East Greenland formed a common group (Alsos *et al.* 2007; Eidesen 2007; Eidesen *et al.* 2007a). For the circumpolar *Saxifraga oppositifolia*, an important model for arctic-alpine plant phylogeography, several studies have evaluated its large-scale range dynamics using different molecular markers and sampling. These have identified ancestral clades in southern Europe and central and/or eastern Eurasia including Beringia, with recent contact zones in the Tatra mountains (western Carpathians), Northern Greenland, and Taymyr (Abbott *et al.* 2000; Winkler *et al.* 2012).

In *C.s. scirpoidea*, two distinctly divergent evolutionary groups were identified using the Δ*K* estimator: one Eastern group that contained populations from Norway and Greenland, and prevailed in the populations from Minnesota and Michigan, and one Western group that contained all populations from Alaska and Yukon and prevailed in populations from British Columbia and Colorado. Using the Pr[X|*K*] estimator, nine or ten evolutionary groups were identified **(Figure 3)**. Although the Pr[X|*K*] estimator has been reported to be more accurate in recapitulating ancestral populations than the Δ*K* estimator (Wang 2017), it may yield over-parametrized models and indeed several optimal values of *K* may exist that correspond to a number of evolutionary groups at different hierarchical levels (Evanno *et al.* 2005). When interpreting the most likely time level each dataset represents, the present-day spatial patterns of genetic variation are often interpreted in relation to the most recent glaciation. Thus, we argue that *K*=2 represents the deepest division of individuals into two historic lineages, while *K*=9-10 represents divergent evolutionary lineages formed in several different glacial refugia during and after the Weichselian-Wisconsin glaciations. None of them has expanded extensively after the last glaciation, although two postglacial meeting zones are evident from the highly admixed populations in West Greenland/East Canada and British Columbia.

In the Eastern group, populations from East Greenland (4, 5) form one well-defined group, while populations from the Northern Lakes and Forests Ecoregion of Minnesota (11) and Michigan (12) form another, both with medium levels of genetic diversity. For the Minnesota/Michigan group, a periglacial refugium south of the Laurentide ice sheet has

been proposed for *C.s. scirpoidea* and other arctic-alpine plants (Dunlop 1990 and references therein). Today, Minnesota populations of *C.s. scirpoidea* are found growing in sedge meadows and shallow prairie swales associated with the ancient beach ridges of the large glacial Lake Agassiz. On the other hand, East Greenland is strongly isolated between two major barriers against gene flow (Greenlandic ice cap and North Atlantic Ocean), and the existence of glacial refugia in this region has been proposed for several arctic-alpine plants (Eidesen *et al.* 2013; Funder 1979; Westergaard *et al.* 2011b). The most likely refugial area for the East Greenland populations would have been the extensive ice-free uplands and dry shelves that were present at the time of the LGM 25 000 – 10 000 years ago (Brochmann *et al.* 2003 and references therein). In our results, separate genetic clustering of the Minnesota and East Greenland populations of *C.s. scirpoidea* could potentially arise from a strong depletion of genetic diversity at the colonization front from North America. Though the small number of alleles in the East Greenland populations would support this hypothesis, there is no evidence of substantially decreased genetic diversity between the two groups. Instead, the increased genetic diversity of admixed populations in West Greenland and East Canada is consistent with postglacial expansion and admixture of lineages that diverged during long-term *in situ* survival in the Upper Midwest (e.g., Driftless Area) and East Greenland. Indeed, increased genetic diversity is expected in contact areas of evolutionary groups expanding from isolated refugia (Petit *et al.* 2003).

The highly divergent population from Western Cordilleran Colorado (13) has a higher number of private alleles than any other population included in this study, including several

547 fixed private alleles **(Table S2, Figure 5, Figure S4)**. It likely survived the last glacial period in
548 the well-known southern high-elevation refugium in the Rocky Mountains, together with
549 other boreal and arctic plant species like *Kobresia myosuroides* and *Dryas octopetala*, all
550 remaining disjunct from their main ranges (Cooper 2004). Similar patterns of divergent
551 southern populations are found in many other arctic-alpine plant species, e.g. *Ranunculus*
552 *glacialis* (Schönswetter *et al.* 2003), *Arabis alpina* (Koch *et al.* 2006), *Oxyria digyna* (Allen *et*
553 *al.* 2012), *Saxifraga oppositifolia* (Winkler *et al.* 2012), and *Sibbaldia procumbens* (Allen *et al.*
554 2015).

555

556 In the Western group, the seven populations from the Tundra, Taiga and Boreal Cordilleran
557 Ecoregions of Yukon and Alaska (18-24) form a well-defined evolutionary group **(Figure 3)**.
558 The populations have average genetic diversity **(Figure 4)**, are poorly differentiated
559 genetically **(Figure S3)** and have many private alleles **(Table S2, Figure 5, Figure S4)**, which is
560 consistent with glacial survival in a large Beringian refugium followed by continuous gene
561 flow among populations. This coincides well with the proposed Beringian refugium for *C.s.*
562 *scirpoidea* based on present-day distribution patterns (Dunlop 1990). The Yukon/Alaska
563 group shows expansion after the last glaciation and introgression with an ancestral element
564 in the populations from the Boreal Cordilleran and Marine West Coast Forest Ecoregions of
565 British Columbia (14-17), as well as with populations from Minnesota/Michigan **(Figure 3)**.
566 Dunlop (1990) hypothesized a cryptic LGM refugium in British Columbia or south of the ice in
567 the Cordilleran, and several other studies of tundra plant species have found unique

haplotypes in British Columbia (e.g. Allen *et al.* 2012; Allen *et al.* 2015; Guest & Allen 2014; Marr *et al.* 2013).

Implications for conservation in Norway

We provide molecular evidence that the three Norwegian populations have survived the last glaciation in separate *in situ* refugia with likely reduced postglacial genetic interchange. Furthermore, we presume that their long isolation has led to local adaptation to their niches (Höglund 2009). Species have often survived past climate changes by range shifts in elevation or altitude; however, this requires sufficient dispersal abilities and availability of suitable habitats. *C.s*. *scirpoidea* has no apparent specialized dispersal adaptations except relatively small seeds, and the Norwegian populations have niches that differ compared to North American populations. Our data thus suggest that distinct management units (MUs, sensu Moritz 1995; Waples & Gaggiotti 2006) should be recognized for each of the Norwegian populations. In fragmented populations of rare species, low levels of genetic diversity are expected to decrease further, while genetic differentiation could increase as a consequence of genetic drift, bottlenecks, and strong natural selection in narrow niches (e.g. Allendorf & Luikart 2007; Honnay & Jacquemyn 2007). Demographic and environmental stochasticity may further exacerbate the accumulation of deleterious mutations, which can be a significant source of extinction vulnerability in small sexual populations, known as genetic meltdown (Lynch *et al.* 1995). Since populations of *C.s. scirpoidea* in Norway are confined to very small habitat patches negatively affected by increasing temperature due to climate change, they are more prone to extinction caused by such stochastic events,

environmental stress, and subsequent competition. Although two of the Norwegian populations are found within protected areas (Làhko National Park and Junkerdalsura Nature Reserve/Junkerdal National Park), our data call for further management efforts to preserve them. This could include compensating efforts like population monitoring and *ex situ* preservation of seeds or plants, or even mitigating efforts to protect their habitat.

ACKNOWLEDGEMENTS

We thank our colleagues for help with providing plant samples; authorities for permits to collect samples and help with field logistics in Norway (Midtre Nordland National Park Board for research permit to Làhko National Park, County Governor of Nordland for research permit to Junkerdalsura Nature Reserve, Ballangen municipality for helicopter landing permits), Greenland (Expedition permits C-14-25 and C-16-26, Survey licences G14-037 and G16-051, permission to visit Mestersvig military station from the Joint Arctic Command), and USA (Minnesota Department of Natural Resources for Special Use Permit to Pembina Wildlife management area, The Nature Conservatory of Colorado for research permit to High Creek Fen); Tina Dahl, Heidi Myklebost, and Theis Westergaard for field assistance; NHMO DNA Bank at the Natural History Museum, University of Oslo for permission to loan and use plant material; Claudia Michel, Beatrice Arnold, and the Genetic Diversity Center at ETH Zürich for laboratory and bioinformatic assistance; the Bioinformatics for Adaptation Genomics Winter School for constructive training; Dr. Jonathan Puritz for knowledgeable advice. This work was supported by a personal postdoctoral grant 231120/F20 to KBW from the Research Council of Norway.

612

613 Data on genetic material contained in this paper are published for non-commercial use only.
614 Utilization by third parties for purposes other than non-commercial scientific research may
615 infringe the conditions under which the genetic resources were originally accessed and
616 should not be undertaken without contacting the corresponding author of the paper and/or
617 seeking permission from the original provider of the genetic material.

618

619 **REFERENCES**

620 Abbott RJ, Smith LC, Milne RI*, et al.* (2000) Molecular analysis of plant migration and refugia in the
621 Arctic. *Science* **289**, 1343-1346.
622 Allen GA, Marr KL, McCormick LJ, Hebda RJ (2012) The impact of Pleistocene climate change on an
623 ancient arctic–alpine plant: multiple lineages of disparate history in *Oxyria digyna*. *Ecology*
624 *and Evolution* **2**, 649-665.
625 Allen GA, Marr KL, McCormick LJ, Hebda RJ (2015) Geographical origins, migration patterns and
626 refugia of *Sibbaldia procumbens*, an arctic–alpine plant with a fragmented range. *Journal of*
627 *Biogeography* **42**, 1665-1676.
628 Allendorf FW, Luikart G (2007) *Conservation and the genetics of populations.* Blackwell Publishing,
629 Oxford, UK.
630 Alsos IG, Ehrich D, Eidesen PB*, et al.* (2015) Long-distance plant dispersal to North Atlantic islands:
631 colonization routes and founder effect. *AoB PLANTS* **7**, plv036-plv036.
632 Alsos IG, Ehrich D, Thuiller W*, et al.* (2012) Genetic consequences of climate change for northern
633 plants. *Proceedings of the Royal Society B: Biological Sciences* **279**, 2042-2051.
634 Alsos IG, Eidesen PB, Ehrich D*, et al.* (2007) Frequent long-distance plant colonization in the changing
635 Arctic. *Science* **316**, 1606-1609.
636 Andersen BG, Borns HWJ (1997) *The Ice Age World. An introduction to Quaternary history and*
637 *research with emphasis on North America and northern Europe during the last 2.5 million*
638 *years.* Scandinavian University Press, Oslo.
639 Arnold NS, van Andel TH, Valen V (2002) Extent and dynamics of the Scandinavian Ice Sheet during
640 Oxygen Isotope Stage 3 (65,000-25,000 yr bp). *Quaternary Research* **57**, 38-48.
641 Beaumont MA (2008) Joint determination of topology, divergence time, and immigration in
642 population trees. In: *Simulation, Genetics, and Human Prehistory* (eds. Matsumura S, Forster
643 P, Renfrew C), pp. 135-154. McDonald Institute for Archaeological Research, Cambridge.
644 Bernard JM (1990) Life history and vegetative reproduction in *Carex*. *Canadian Journal of Botany* **68**,
645 1441-1448.
646 Birkeland S, Skjetne IEB, Brysting AK, Elven R, Alsos IG (2017) Living on the edge: conservation
647 genetics of seven thermophilous plant species in a high Arctic archipelago. *AoB PLANTS* **9**,
648 plx001-plx001.
649 Birks HH (2008) The Late-Quaternary history of arctic and alpine plants. *Plant Ecology & Diversity* **1**,
650 135-146.
651 Blytt A (1876) *Essay on the immigration of the Norwegian flora during alternating rainy and dry*
652 *periods.* Cammermeyer, Christiania (Oslo), Norway.
653 Blytt A (1888) Die Theorie der wechselnden kontinentalen und insularen Klimate. *Botanische*
654 *Jahrbücher* **2**, 1-50.
655 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data.
656 *Bioinformatics* **30**, 2114-2120.
657 Brochmann C, Gabrielsen TM, Nordal I, Landvik J, Elven R (2003) Glacial survival or *tabula rasa*? The
658 history of North Atlantic biota revisited. *Taxon* **52**, 417-450.
659 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for
660 population genomics. *Molecular Ecology* **22**, 3124-3140.
661 Chong Z, Ruan J, Wu C-I (2012) Rainbow: an integrated tool for efficient clustering and assembling
662 RAD-seq reads. *Bioinformatics* **28**, 2732-2737.
663 Cooper DJ (2004) The modern environment, flora and vegetation of South Park, Colorado. In:
664 *Biodiversity response to climate change in the middle Pleistocene. The Porcupine Cave fauna*
665 *from Colorado.* (ed. Barnosky AD), pp. 27-38. University of California Press.

Cornuet JM, Pudlo P, Veyssier J*, et al.* (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187-1189.

Dahl E (1963) Plant migrations across the North Atlantic Ocean and their importance for the palaeogeography of the region. In: *North Atlantic Biota and Their History* (eds. Löve Á, Löve D), pp. 173-188. Pergamon Press, Oxford.

Danecek P, Auton A, Abecasis G*, et al.* (2011) The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158.

Dunlop DA (1990) *The biosystematics of Carex section Scirpinae (Cyperaceae)*, University of New Hampshire.

Dunlop DA (2003) *Carex* Linnaeus sect. *Scirpinae* (Tuckerman) Kükenthal. In: *Flora of North America Vol. 23*.

Dunlop DA, Crow GE (1999) The taxonomy of *Carex* section *Scirpinae* (Cyperaceae). *Rhodora* **101**, 163-199.

Dyke AS, Andrews JT, Clark PU*, et al.* (2002) The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. *Quaternary Science Reviews* **21**, 9-31.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Eidesen PB (2007) *Arctic-alpine plants on the move. Individual and comparative phylogeographies reveal responses to climate change.* PhD thesis, University of Oslo.

Eidesen PB, Alsos IG, Popp M*, et al.* (2007a) Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species. *Molecular Ecology* **16**, 3902-3925.

Eidesen PB, Carlsen T, Molau U, Brochmann C (2007b) Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic. *Journal of Biogeography* **34**, 1559-1574.

Eidesen PB, Ehrich D, Bakkestuen V*, et al.* (2013) Genetic roadmap of the Arctic: plant dispersal highways, traffic barriers and capitals of diversity. *New Phytologist* **200**, 898-910.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Fagundes NJR, Ray N, Beaumont M*, et al.* (2007) Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17614-17619.

Falush D, Stephens M, Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152.

Funder S (1979) Ice-age plant refugia in East Greenland. *Palaeogeography, Palaeoclimatology, Palaeoecology* **28**, 279-295.

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

Gjærevoll O (1990) Alpine plants. In: *Maps of distribution of Norwegian vascular plants. Volume II* (eds. Berg RY, Fægri K, Gjærevoll O). Tapir Publishers, Trondheim.

Guest HJ, Allen GA (2014) Geographical origins of North American *Rhodiola* (Crassulaceae) and phylogeography of the western roseroot, *Rhodiola integrifolia*. *Journal of Biogeography* **41**, 1070-1080.

Hanski I (1998) Metapopulation dynamics. *Nature* **396**, 41-49.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences* **359**, 183-195.

Honnay O, Jacquemyn H (2007) Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology* **21**, 823-831.

716 Hultén E (1937) *Outline of the history of Arctic and boreal biota during the Quaternary period.* J.
717 Cramer, Lehre, New York.
718 Hultén E (1958) *The amphi-Atlantic plants and their phytogeographical connections.* Almqvist &
719 Wiksell, Stockholm.
720 Höglund J (2009) *Evolutionary conservation genetics.* Oxford University Press, Oxford.
721 Høiland K (1986) Utsatte planter i Nord-Norge. Spesiell del. In: *Økoforsk rapport 1986: 2*, pp. 1-163.
722 Koch MA, Kiefer C, Ehrich D*, et al.* (2006) Three times out of Asia Minor: the phylogeography of
723 *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* **15**, 825-839.
724 Kolstrup E, Olsen L (2012) Palaeoenvironmental developments in the central Scandinavian mountains
725 during deglaciation – a discussion. *Norsk Geografisk Tidsskrift - Norwegian Journal of*
726 *Geography* **66**, 30-51.
727 Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for
728 identifying clustering modes and packaging population structure inferences across K.
729 *Molecular Ecology Resources* **15**, 1179-1191.
730 Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
731 *arXiv:1303.3997v2*.
732 Linge H, Brook EJ, Nesje A*, et al.* (2006) In situ ^{10}Be exposure ages from southeastern Norway:
733 implications for the geometry of the Weichselian Scandinavian ice sheet. *Quaternary Science*
734 *Reviews* **25**, 1097-1109.
735 Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting
736 population genetics and genomics programs. *Bioinformatics* **28**, 298-299.
737 Lynch M, Conery J, Burger R (1995) Mutation accumulation and the extinction of small populations.
738 *American Naturalist* **146**, 489-518.
739 Márquez-Corro JI, Escudero M, Martín-Bravo S, Villaverde T, Luceño M (2017) Long-distance dispersal
740 explains the bipolar disjunction in *Carex macloviana*. *American Journal of Botany* **104**, 663-
741 673.
742 Marr KL, Allen GA, Hebda RJ, McCormick LJ (2013) Phylogeographical patterns in the widespread
743 arctic-alpine plant *Bistorta vivipara* (Polygonaceae) with emphasis on western North
744 America. *Journal of Biogeography* **40**, 847-856.
745 Moritz C (1995) Uses of molecular phylogenies for conservation. *Philosophical Transactions of the*
746 *Royal Society of London Series B-Biological Sciences* **349**, 113-118.
747 Nathan R (2006) Long-distance dispersal of plants. *Science* **313**, 786-788.
748 Nathorst AG (1892) Kritische Bemerkungen über die Geschichte der Vegetation Grönlands. *Englers*
749 *Botanische Jahrbücher* **14**, 183-221.
750 Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction
751 endonucleases. *Proceedings of the National Academy of Sciences of the United States of*
752 *America* **76**, 5269-5273.
753 Nordal I (1987) Tabula rasa after all? Botanical evidence for ice-free refugia in Scandinavia reviewed.
754 *Journal of Biogeography* **14**, 377-388.
755 Olsen L (1997) Rapid shifts in glacial extension characterise a new conceptual model for glacial
756 variations during the Mid and Late Weichselian in Norway. *Norges Geologiske Undersøkelser*
757 *Bulletin* **433**, 54-55.
758 Ostenfeld CEH (1926) The flora of Greenland and its origin. *Biologiske Meddelelser* **6**, 1-71.
759 Parducci L, Jørgensen T, Tollefsrud MM*, et al.* (2012) Glacial survival of boreal trees in northern
760 Scandinavia. *Science* **335**, 1083-1086.
761 Pellissier L, Eidesen PB, Ehrich D*, et al.* (2016) Past climate-driven range shifts and population genetic
762 diversity in arctic plants. *Journal of Biogeography* **43**, 461-470.
763 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive
764 method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS*
765 *ONE* **7**, e37135.

766 Petit R, Aguinagalde I, de Beaulieu J-L*, et al.* (2003) Glacial refugia: Hotspots but not melting pots of
767 genetic diversity. *Science* **300**, 1563-1565.
768 Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus
769 genotype data. *Genetics* **155**, 945-959.
770 Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology &*
771 *Evolution* **23**, 564-571.
772 Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for
773 population genomics of non-model organisms. *PeerJ* **2**, e431.
774 Quinzin MC, Normand S, Dellicour S, Svenning J-C, Mardulyn P (2017) Glacial survival of trophically
775 linked boreal species in northern Europe. *Proceedings of the Royal Society B: Biological*
776 *Sciences* **284**.
777 Schönswetter P, Elven R, Brochmann C (2008) Trans-Atlantic dispersal and large-scale lack of genetic
778 structure in the circumpolar, arctic-alpine sedge *Carex bigelowii* s. l. (Cyperaceae). *American*
779 *Journal of Botany* **95**, 1006-1014.
780 Schönswetter P, Paun O, Tribsch A, Niklfeld H (2003) Out of the Alps: colonization of northern Europe
781 by East Alpine populations of the glacier buttercup *Ranunculus glacialis* L. (Ranunculaceae).
782 *Molecular Ecology* **12**, 3373-3381.
783 Schönswetter P, Suda J, Popp M, Weiss-Schneeweiss H, Brochmann C (2007) Circumpolar
784 phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA
785 sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and*
786 *Evolution* **42**, 92-103.
787 Sernander R (1896) Några ord med anledning af Gunnar Andersson, Svenska Växtvärldens historia.
788 *Botaniska Notiser* **1896**, 114-128.
789 Shackleford R (2003) Conservation assessment for bulrush sedge (*Carex scirpoidea* Michx.):
790 subspecies *scirpoidea* and *convoluta* (Kükenth.) Dunlop, p. 54. USDA Forest Service, Eastern
791 Region.
792 Shafer ABA, Cullingham CI, Côté SD, Coltman DW (2010) Of glaciers and refugia: a decade of study
793 sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*
794 **19**, 4589-4621.
795 Skifte O (1985) Nye funn av grønlandsstarr - *Carex scirpoidea* Michx. - i Nordland fylke. *Blyttia* **43**, 16-
796 21.
797 Skrede I, Eidesen PB, Portela RP, Brochmann C (2006) Refugia, differentiation and postglacial
798 migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.).
799 *Molecular Ecology* **15**, 1827-1840.
800 Smith J, Kronforst MR (2013) Do *Heliconius* butterfly species exchange mimicry alleles? *Biology*
801 *Letters* **9**, 20130503.
802 Solstad H, Elven R, Mjelde M*, et al.* (2015) Karplanter. Pteridophyta, Pinophyta, Magnoliophyta. In:
803 *Norsk rødliste for arter 2015* eds. Henriksen S, Hilmo O). Artsdatabanken, Norge.
804 Svendsen JI, Alexanderson H, Astakhov VI*, et al.* (2004) Late Quaternary ice sheet history of northern
805 Eurasia. *Quaternary Science Reviews* **23**, 1229-1271.
806 Villaverde T, Escudero M, Martín-Bravo S*, et al.* (2015) Direct long-distance dispersal best explains
807 the bipolar distribution of *Carex arctogena* (*Carex* sect. *Capituligerae*, Cyperaceae). *Journal of*
808 *Biogeography* **42**, 1514–1525.
809 Wang J (2017) The computer program STRUCTURE for assigning individuals to populations: easy to
810 use but easier to misuse. *Molecular Ecology Resources* **17**, 981-990.
811 Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic
812 methods for identifying the number of gene pools and their degree of connectivity.
813 *Molecular Ecology* **15**, 1419-1439.
814 Warming E (1888) Om Grønlands vegetation. *Meddelelser om Grønland* **12**, 1-245.

Westergaard KB, Alsos IG, Engelskjøn T, Flatberg KI, Brochmann C (2011a) Trans-Atlantic genetic uniformity in the rare snowbed sedge *Carex rufina*. *Conservation Genetics* **12**, 1367-1371.
Westergaard KB, Alsos IG, Popp M*, et al.* (2011b) Glacial survival may matter after all: nunatak signatures in the rare European populations of two west-arctic species. *Molecular Ecology* **20**, 376-393.
Westergaard KB, Jørgensen MH, Gabrielsen TM, Alsos IG, Brochmann C (2010) The extreme Beringian/Atlantic disjunction in *Saxifraga rivularis* (Saxifragaceae) has formed at least twice. *Journal of Biogeography* **37**, 1262-1276.
Willis SC, Hollenbeck CM, Puritz JB, Gold JR, Portnoy DS (2017) Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. *Molecular Ecology Resources* **17**, 955-965.
Winkler M, Tribsch A, Schneeweiss GM*, et al.* (2012) Tales of the unexpected: Phylogeography of the arctic-alpine model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Molecular Ecology* **21**, 4618-4630.
Zheng X, Levine D, Shen J*, et al.* (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326-3328.

DATA ACCESSIBILITY

The Illumina paired-end sequences for each individual are available via the EMBL Nucleotide Archive under accession ID PRJEB28490.

AUTHOR CONTRIBUTIONS

KBW, LB, HS and AW designed the research; KBW and LB did the fieldwork; KBW did the lab work with input from SF; KBW, NZ, HS and SF performed the data analyses. KBW wrote the manuscript with input from all co-authors. All authors read and approved the final manuscript.

TABLES AND FIGURES

Table 1. Collection information for the 24 investigated populations of *Carex scirpoidea* Michx. ssp. *scirpoidea* (Cyperaceae); *n* is the number of individuals collected/retained in the ddRADseq analyses.

Pop no	Pop ID	Country	Sampling locality	*n*	Lat	Long	Year	Collector(s)*	DNA bank acc no (O-DP)
1	KBW14-7	Norway	Nordland, Kjelvatn	20/20	68.18	17.21	2014	KBW, HM	
2	KBW14-8	Norway	Nordland, Kjerringa	20/20	66.78	14.21	2014	KBW, HM	
3	KBW14-9	Norway	Nordland, Solvågtind	20/20	66.83	15.42	2014	KBW, HM	
4	KBW14-1	Greenland (DK)	Traill Ø, Holm Bugt	18/12	72.52	-23.98	2014	KBW, TD	
5	KBW14-2	Greenland (DK)	Mestersvig	19/17	72.23	-23.98	2014	KBW, TD	
6	KW06-24/25	Greenland (DK)	Qeqertarsuaq, Østerdalen	10/10	69.25	-53.51	2006	KBW, KIF, BBF	47294-47314
7	KW06-33	Greenland (DK)	Kangerlussuaq	10/10	67.00	-50.67	2006	KBW, KIF, BBF	47393-47402
8	KW06-35	Greenland (DK)	Nuuk, Qinngorput	10/10	64.17	-51.66	2006	KBW, KIF	47415-47424
9	KW06-37	Greenland (DK)	Narsqaq	10/10	60.92	-46.05	2006	KBW, KIF	47437-47446
10	262-07	Canada	Quebec, Salluit	2/2	62.26	-75.73	2007	KIF, BBF	46619-46628
11	KBW14-3	USA	Minnesota, Pembina WMA	19/17	48.08	-96.45	2014	KBW, LPB	
12	0711201301	USA	Michigan, Escanaba river	9/8	45.90	-87.21	2013	LPB	
13	KBW14-6	USA	Colorado, High Creek Fen	20/20	39.10	-105.97	2014	KBW, LPB	
14	KM-2	Canada	British Columbia, Summit Lake	10/10	58.67	-124.64	2008	KM	47150-47159
15	KM08-2	Canada	British Columbia, Nimbus Peak	10/10	49.76	-122.67	2008	KM, RH, WM	47129-47138
16	KM-1	Canada	British Columbia, Little Blue Sheep Lake	10/10	58.74	-128.25	2008	KM	47140-47149
17	0809200801	Canada	British Columbia, Haines Hwy Summit	6/4	59.67	-136.54	2008	LPB	
18	KW86	Canada	Yukon, Keele Range	8/8	66.97	-140.80	2007	BAB, JL, CAK, LM	48157-48166
19	0719201001	Canada	Yukon, Quill Creek	10/10	61.46	-139.52	2010	LPB	
20	KBW14-4	USA	Alaska, Eagle Summit	20/20	65.48	-145.41	2014	KBW, LPB	
21	KBW14-5	USA	Alaska, Twelve Mile Summit	20/19	65.40	-145.96	2014	KBW, LPB	
22	KW06-62	USA	Alaska, Rainbow Ridge	5/4	63.31	-145.65	2007	MC	47614-47623
23	KW06-94	USA	Alaska, Brooks Range, Galbraith Lake	10/10	68.46	-149.42	2007	MC	47636-47645
24	0821200801	USA	Alaska, Nome, Anvil Mountain	10/9	64.57	-165.36	2008	LPB	

* Collectors: BAB = B.A. Bennett, BBF = B.B. Flatberg, CAK = C. A. Kennedy, HM = H. Myklebost, JL = J. Line, KBW = K.B. Westergaard, KIF = K.I. Flatberg, KM = K. Marr, LPB = L.P. Bruederle, LM = L. Mennell, MC = M. Carlson, RH = R. Hebda, TD = T. Dahl, WM = W. MacKenzie.

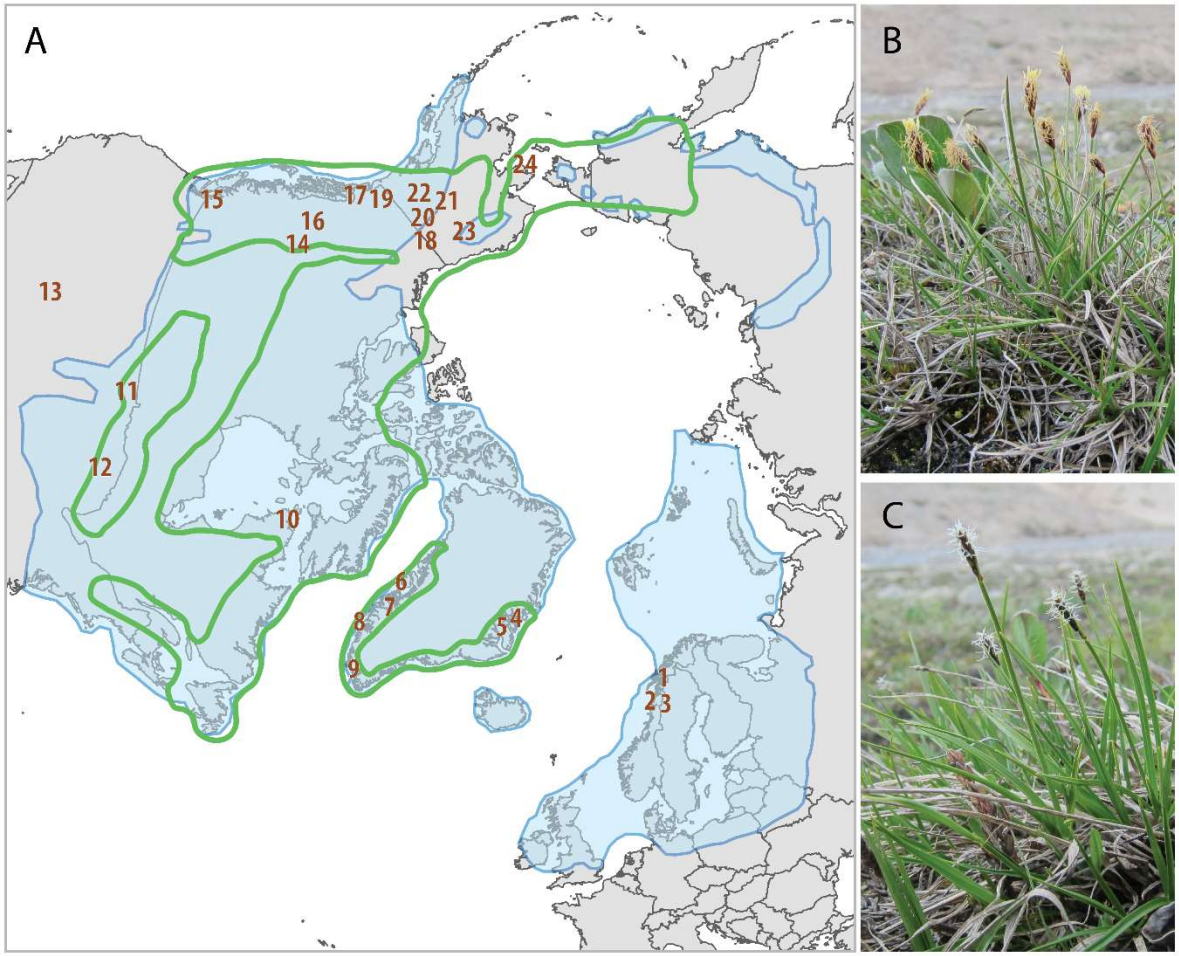
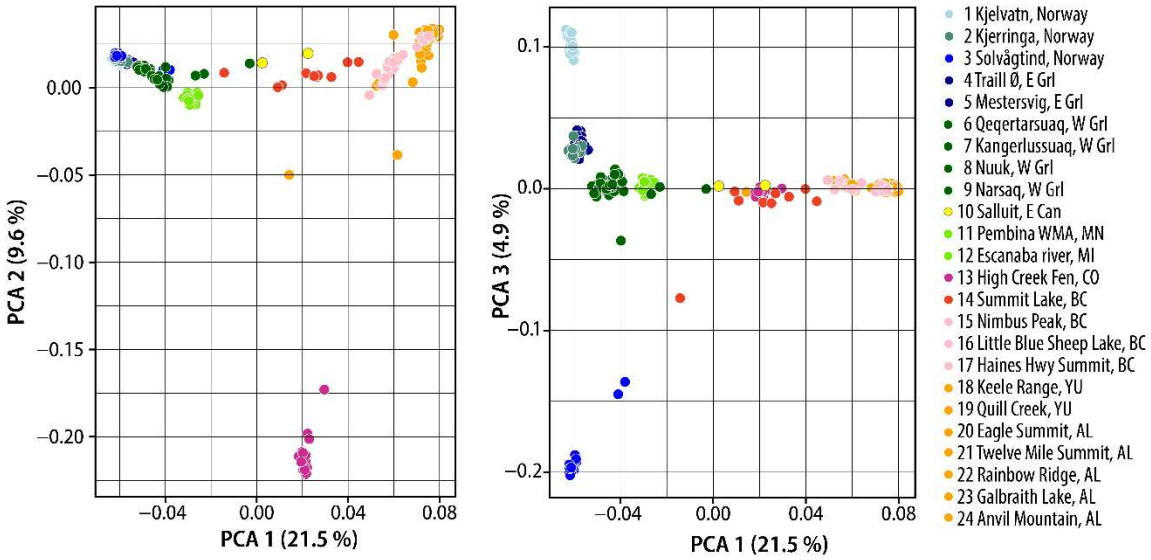


Figure 1. A) Outline of geographic distribution (green line) and sampling locations (population numbers, see **Table 1**) of *Carex scirpoidea* Michx. ssp. *scirpoidea* (Cyperaceae). Blue shade indicates the Last Glacial Maximum ice extent (Andersen & Borns 1997; Dyke *et al.* 2002; Svendsen *et al.* 2004). B) Male and C) female individuals, Northeast Greenland National Park. Photos: K. B. Westergaard.

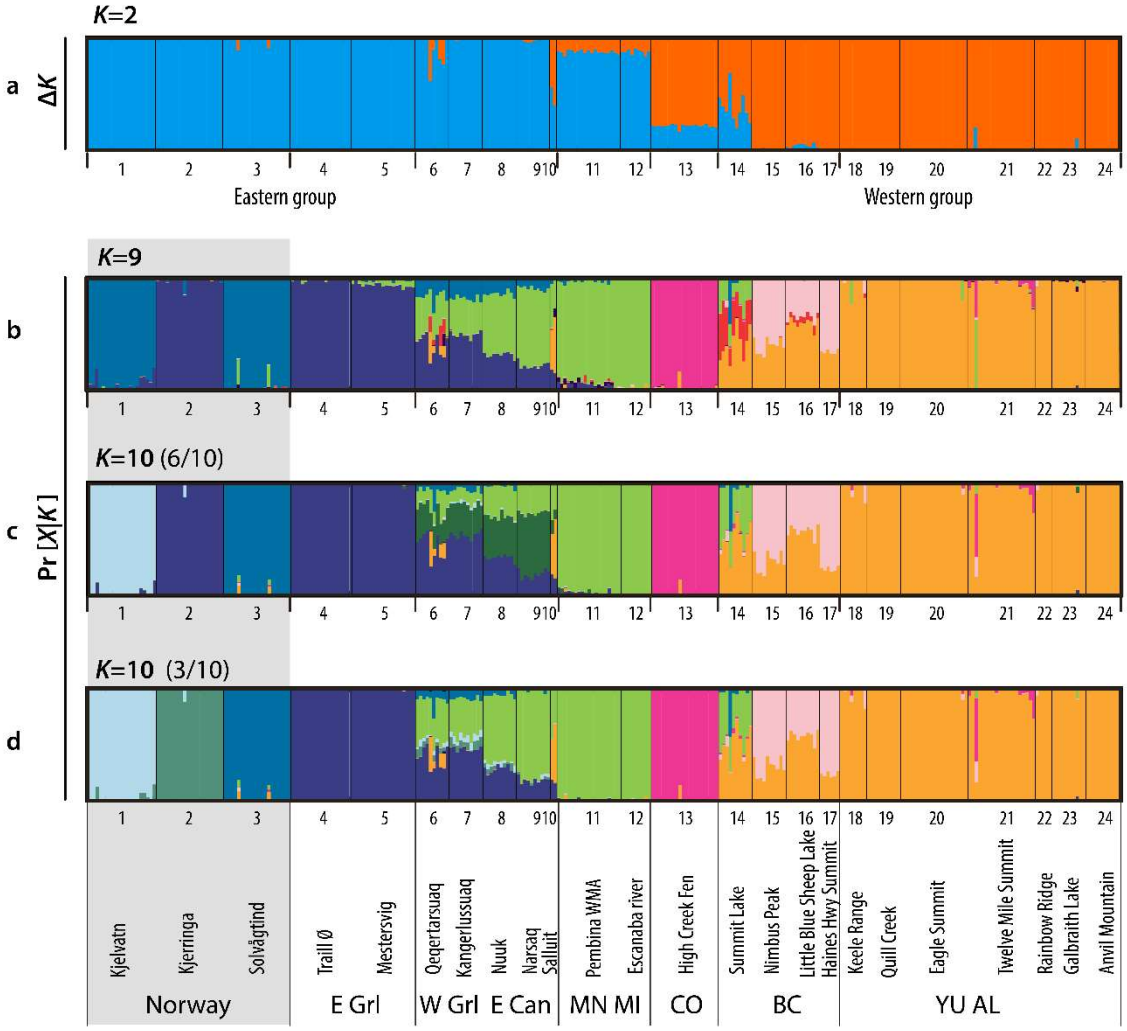
857



858

859 **Figure 2.** 2-dimensional PCA plots based on 5,307 SNPs in 24 populations of *Carex scirpoidea*
860 Michx. ssp. *scirpoidea* (Cyperaceae), see Table 1 for population information. The 24
861 populations are color coded according to their STRUCTURE group (*K*=10; see **Figure 3**), and
862 ordered geographically from East to West; Norway, East Greenland (E Grl), West Greenland
863 (W Grl), East Canada (E Can), Minnesota, USA (MN), Michigan, USA (MI), Colorado, USA (CO),
864 British Columbia, Canada (BC), and Yukon, Canada (YU) and Alaska, USA (AL).

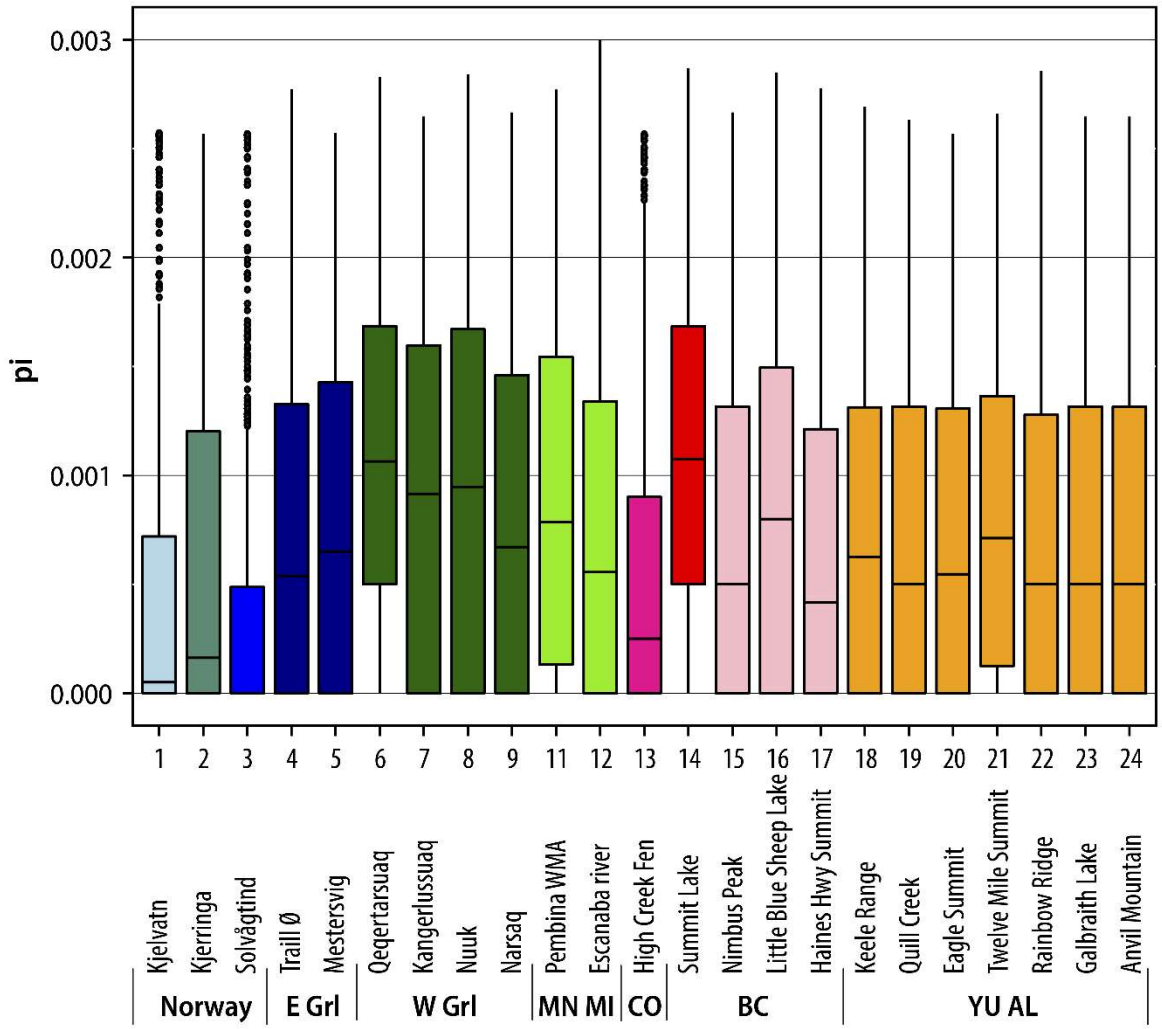
865



866

867 **Figure 3.** Bar plots showing the results from STRUCTURE analyses of 290 individuals of *Carex*
868 *scirpoidea* Michx. ssp. *scirpoidea* (Cyperaceae) from 24 populations (see **Table 1**), using
869 2,796 haplotyped SNPs; a) assignment of individuals to an Eastern and Western STRUCTURE
870 group (*K*=2), as identified by Δ*K* analyses; b) assignment of individuals to nine STRUCTURE
871 groups using default parameter settings (*K*=9), as identified by the highest mean value of
872 Pr[X|*K*]; c) assignment of individuals to ten STRUCTURE groups using custom parameter
873 settings (*K*=10, major mode, 6 out of 10 runs), as identified by the highest mean value of
874 Pr[X|*K*]; d) assignment of individuals to ten STRUCTURE groups using custom parameter
875 settings (*K*=10, minor mode 3/10), as identified by the highest mean value of Pr[X|*K*]. Minor
876 mode 1/10 not shown. The 24 populations are ordered geographically from East to West;
877 Norway (highlighted with gray shading), East Greenland (E Grl), West Greenland (W Grl), East
878 Canada (E Can), Minnesota, USA (MN), Michigan, USA (MI), Colorado, USA (CO), British
879 Columbia, Canada (BC), and Yukon, Canada and Alaska, USA (YU AL).

880



881

882 **Figure 4.** Boxplots showing the nucleotide diversity (π) of 290 individuals of *Carex scirpoidea*
883 Michx. ssp. *scirpoidea* (Cyperaceae) from 23 populations (excluding population 10 with only
884 two individuals; see **Table 1**). Dots indicate outlier values. The populations are color coded
885 according to their STRUCTURE group (*K*=10; see **Figure 3**), and ordered geographically from
886 East to West; Norway, East Greenland (E Grl), West Greenland (W Grl), East Canada (E Can),
887 Minnesota, USA (MN), Michigan, USA (MI), Colorado, USA (CO), British Columbia, Canada
888 (BC), and Yukon, Canada and Alaska, USA (YU AL).

889

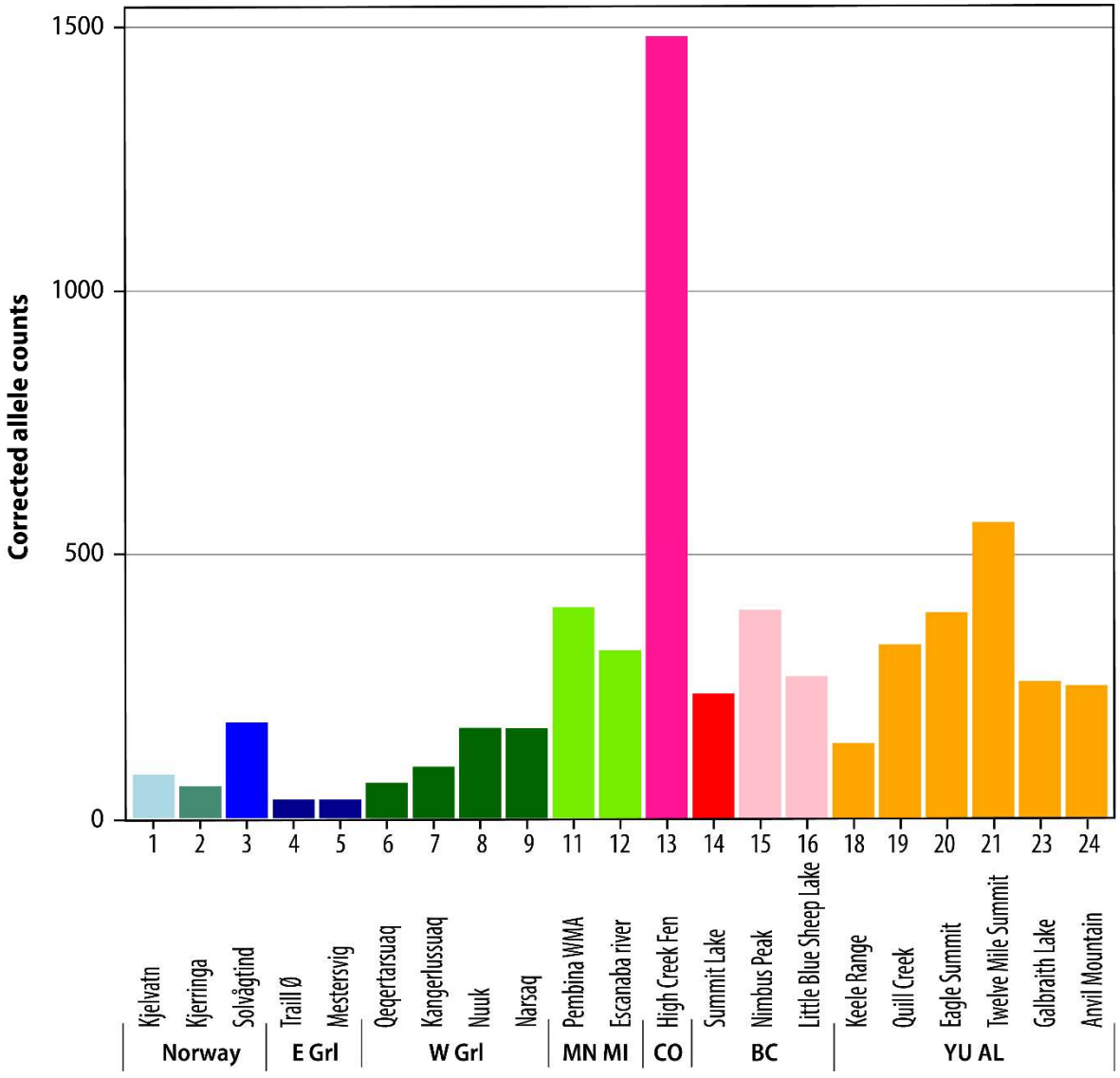


Figure 5. Number of corrected private alleles for each of 21 populations of *Carex scirpoidea* Michx. ssp. *scirpoidea* (Cyperaceae) inferred with an allele frequency of <1% in the metapopulation and ≥10% in the target population (corresponding to two observed alleles; populations 10, 17 and 22 with less than eight individuals were excluded); see **Table 1** for population information and **Table S2** for corrected numbers of private alleles). The populations are color coded according to their STRUCTURE group (*K*=10; see **Figure 3**), and ordered geographically from East to West; Norway, East Greenland (E Grl), West Greenland (W Grl), East Canada (E Can), Minnesota, USA (MN), Michigan, USA (MI), Colorado, USA (CO), British Columbia, Canada (BC), and Yukon, Canada and Alaska, USA (YU AL).