Journal of Biogeography (J. Biogeogr.) (2016)

ORIGINAL ARTICLE

A new Eurasian phylogeographical paradigm? Limited contribution of southern populations to the recolonization of high latitude populations in *Juniperus communis* L. (Cupressaceae)

E. V. Hantemirova[1]*, B. Heinze[2], S. G. Knyazeva[3], A. M. Musaev[4], M. Lascoux[5] and V. L. Semerikov[1]

[1]*Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, 8 Marta Str., 202, 620144 Ekaterinburg, Russia,* [2]*Department of Forest Genetics, Austrian Federal Research Centre for Forests, Seckendorff-Gudent-Weg, 8, 1130 Vienna, Austria,* [3]*Forest Institute, Siberian Branch of the Russian Academy of Sciences, 660036 Krasnojarsk Akademgorodok, Russia,* [4]*Mountain Botanical Garden of the Dagestan Scientific Centre of the Russian Academy of Sciences, 45 M. Gadgiev Street, 367000 Makhachkala, Republic of Dagestan, Russia,* [5]*Department of Ecology & Genetics, EBC, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden*

*Correspondence: Elena V. Hantemirova, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, 8 Marta Str., 202, 620144 Ekaterinburg, Russia. E-mail: hantemirova@ipae.uran.ru

ABSTRACT

Aim The aims of this population genetics study of the common juniper across Eurasia were to (1) assess the contribution of southern mountain ranges to the post-glacial recolonization of high latitudes and (2) test whether recent expansion or high gene flow could explain the low genetic differentiation in Northern Eurasia.

Location Northern Eurasia and mountain regions of Central Europe and Asia.

Methods Six hundred and twenty-two individuals were sampled in 42 populations. Two chloroplast DNA (cpDNA) fragments were investigated (*trn*T-*trn*L and *16*S-*trn*A). Analyses of the distribution of haplotypes across the continent included a suite of phylogeographical and phylogenetic tests. Putative geographical distribution in the past was reconstructed using environmental niche modelling.

Results Eighty-four haplotypes clustered into four main clades (GL1-GL4). The largest clade, GL3, corresponds to populations from the Alps, northern Europe, Western Caucasus and Siberia. These populations were moderately differentiated (28%) compared to the total range (76%) and Fu's F_s statistic was negative, indicating a population expansion. Some haplotypes within GL3 form subclades with a restricted geographical distribution, suggesting a local origin of the mutation and limited dispersal. In line with these findings, modelling of ecological niches found no significant reduction in the expected range during the LGM. Remarkably, populations from the eastern part of North Caucasus, the Himalayas, Tien Shan and south Siberia were distinctly different from populations in the rest of the range.

Main conclusions As in Siberian larch species, the pattern of genetic diversity at cpDNA across the natural range of *J. communis* suggests that colonization of northern Europe and Siberia started from a limited area and predated the last glaciation. It is likely that juniper survived the subsequent glacial epoch at high latitudes in cryptic refugia serving as secondary centres of recolonization. Southern mountain refugia contribution, to the recolonization of high latitudes was, at best, limited.

Keywords

chloroplast DNA, common juniper, Cupressaceae, Eurasia, glacial cycles, phylogeography

Journal of Biogeography

INTRODUCTION

Pleistocene glacial cycles are one of the most important factors determining the demographic history and, consequently, the population genetic structure of species of the Northern Hemisphere (Hewitt, 2000). However, their effect largely depends on the environmental characteristics of the species and the geographical region (Petit *et al.*, 2003; Lascoux *et al.*, 2004). Cold-resistant species, such as birches, larches or spruces, responded to Pleistocene climate dynamics very differently than thermophilic ones, like oaks and beeches: while the latter could only survive glacial episodes in southern refugia, the former could persist close to the ice sheets (Willis & Van Andel, 2004; Binney *et al.*, 2009). Indeed, while post-glacial migration from southern refugia left a strong genetic footprint in thermophilic species, with a clear structuring of chloroplastic haplotypes, in boreal species such a signal was conspicuously absent [e.g. *Betula pubescens*/*B. pendula* (Maliouchenko *et al.*, 2007), *Salix caprea* (Palme *et al.*, 2003)] or mixed and hard to interpret, and giving a better match with the earlier, pre-LGM epoch (e.g. *Larix* sp., Semerikov *et al.*, 2013; Polezhaeva *et al.*, 2010). Finally, in contrast to temperate or boreal species, arctic species that are adapted to periglacial habitats may expand during glacial phases and contract during interglacials (e.g. Prost *et al.*, 2013). The effect of glacial cycles on genetic variability and demographic history of species with a wide amplitude of environmental requirements, as opposed to species that are confined to the Arctic or to cold-sensitive ones, remains poorly understood.

Juniperus communis L. is a good example of a species with such a wide ecological tolerance. This evergreen, cold-tolerant and xeromorphic gymnosperm is able to grow above the polar and alpine tree lines, reaching the 72 latitude on the Taimyr peninsula and 4300 m a. s. l. in the Himalayas, producing variable growth forms ranging from a shrub to an erect tree. *Juniperus communis* has an extensive circumboreal range and several morphologically barely distinguishable varieties (Farjon, 2001; Thomas *et al.*, 2007; Adams, 2014). It is generally found under the canopy of coniferous forests, often together with *Pinus sylvestris*. Its pollen is wind-dispersed supporting elevated gene flow, as in other boreal trees.

A previous rangewide allozyme study of the common juniper (Hantemirova *et al.*, 2012) revealed one large and a few smaller geographical groups that are markedly distinct from each other. The smaller groups correspond to Alaska, the north-east of Siberia and the south of the Russian Far East, part of south Siberia, and the Caucasus. The large 'northern group' corresponds to the main part of the range, including most of Europe and Siberia. Within-population diversity is high with expected heterozygosity, H_e, ranging from 0.134 to 0.260, and F_{ST} among populations is low with values around 5%. This general structure, with one large homogeneous group, surrounded by smaller, highly differentiated ones, is not unique to *J. communis* and is observed among other boreal or temperate plants and animals, for example, the wood lemming, the brown bear or black alder (Fedorov *et al.*, 2008; Korsten *et al.*, 2009; Havrdova *et al.*, 2015). Generally, the genetic structure of these species fits well with the expectations of the rear edge hypothesis (Hampe & Petit, 2005) that assumes a stable, genetically well differentiated set of rear edge populations (ancient populations persisting over long periods of time) contributing to highly dynamic, expanding leading edge ones that are genetically more similar among them. An alternative explanation of the observed allozyme structure in common juniper is a long-lasting survival in the northern part of the range of numerous 'microrefugia' which would provide sufficient gene flow to homogenize populations across this territory, with a limited contribution from the rear edge populations. This, for instance, seems to be the case in *Larix sibirica* (Semerikov *et al.*, 2013). These two hypotheses differ in their expectations. Under the rear edge hypothesis, total genetic diversity and population differentiation of common juniper at the leading edge are expected to be lower than among the rear edge populations. Within-population diversity should also be higher and isolation by distance should be weak. The microrefugia hypothesis does not predict a decrease in total genetic diversity in the north but predicts some isolation by distance. Under the rear edge hypothesis the haplotype diversity of the leading edge populations is also expected to be a subsample of the diversity in the rear edge populations, whereas under the microrefugia hypothesis the presence of endemic locally evolved groups of haplotypes is expected in the north.

A chloroplast DNA (cpDNA) study of *J. communis* restricted to Europe discovered low inter-population diversity, probably because of significant gene flow (Vargas, 2003; Filipowicz *et al.*, 2006). In contrast, high inter-population differentiation within European juniper populations was found with AFLP markers, although without correlation with geographical distances (Michalczyk *et al.*, 2010). These studies did not explicitly test the two hypotheses introduced above. In other species from the genus *Juniperus* [*J. przewalskii* Komarov (Zhang *et al.*, 2005; Li *et al.*, 2011), *J. tibetica* complex (Opgenoorth *et al.*, 2010), *J. sabina* L. (Guo *et al.*, 2010) from Qinghai–Tibet Plateau and *J. phoenica* L. from the Mediterranean (Boratynski *et al.*, 2009)], a strong phylogeographical signal and significant inter-population differentiation were observed.

The aims of this study were twofold: (1) to test whether southern montane populations (Caucasus, the Himalayas, Tien Shan and south Siberia) contributed to the recolonization of northern latitudes, (2) to test whether the genetic pattern observed in junipers corresponds better to the expectations of the rear edge hypothesis or the micro-refugia hypothesis. To that end we developed chloroplast DNA markers, which are paternally inherited in Cupressaceae (Neale *et al.*, 1989; Neale *et al.* 1991; Mogensen, 1996; Kondo *et al.*, 1998), and genotyped individuals sampled across the Eurasian part of the range of *J. communis*.

MATERIALS AND METHODS

Population sampling

In total, 42 populations of *J. communis* were sampled (Fig. 1, see Appendix S1 in Supporting Information): 25 Russian populations (*J. communis* var. *communis* and *J. communis* var. *saxatilis*), one Swedish, three Alpine, two Himalayan and two from the Tien Shan mountains (Kazakhstan and Kyrgyzstan, Central Asia). Eight populations of *J. communis* var. *oblonga* from the North Caucasus and one population of *J. communis* var. *depressa* Pursh from Alaska were also sampled. Each sample contained from 5 to 32 randomly selected plants, thus 622 plants were analysed in total. Part of this material was previously analysed with isozyme markers (Hantemirova *et al.*, 2012).

DNA extraction, PCR amplification, RFLP and sequencing

Genomic DNA was extracted from fresh or silica gel dried needles using the CTAB method according to Devey *et al.* (1996), and used as a template for PCR amplification of two fragments of chloroplast DNA. Region *trn*T-*trn*L was amplified with primers a, b, c, d of Taberlet *et al.* (1991), and the *16*S-*trn*A intergenic spacer region using the primer pair, 5′-gaaggctgggctagtgactg-3′ and 5′-gagataagcggactcgaacc-3′ (V.L. Semerikov, unpublished), and an additional internal primer 5′-gaaaccaatgccataagcaa- 3′. The individual PCR product was subjected to digestion with *Taq*I enzyme for the *16*S-*trn*A region and *Hinf*I for the *trn*T-*trn*L region using the manufacturer's instructions (Sibenzyme, Novosibirsk, Russia). Products of restriction reactions were separated individually in 6% denaturing polyacrylamide gels with subsequent silver staining. In each population all haplotypes (banding patterns) identified with this restriction fragment length polymorphism (RFLP) procedure in each of the fragments were sequenced in at least one individual. Indel polymorphism in the *16*S-*trn*A spacer allowed to distinguishing the majority of all haplotypes by the RFLP method. However, in some populations (the Alps, Kola peninsula, Estonia, Polar Ural, Sweden, Tver, Yakutia) many different RFLP patterns with similar band sizes were revealed in the *16*S-*trn*A region, and therefore all samples were sequenced in each of those populations. To that end the PCR bands were cut out of 1.0% agarose gels, isolated with SiO_2 particles (Vogelstein & Gillespie, 1979) and then used as templates for direct sequencing using BigDye v. 3.1 kit (Gene Analyzers 3130; Applied Biosystems, Foster City, CA, USA), and using PCR primers as the sequencing primers. Fragments *trn*T-*trn*L and

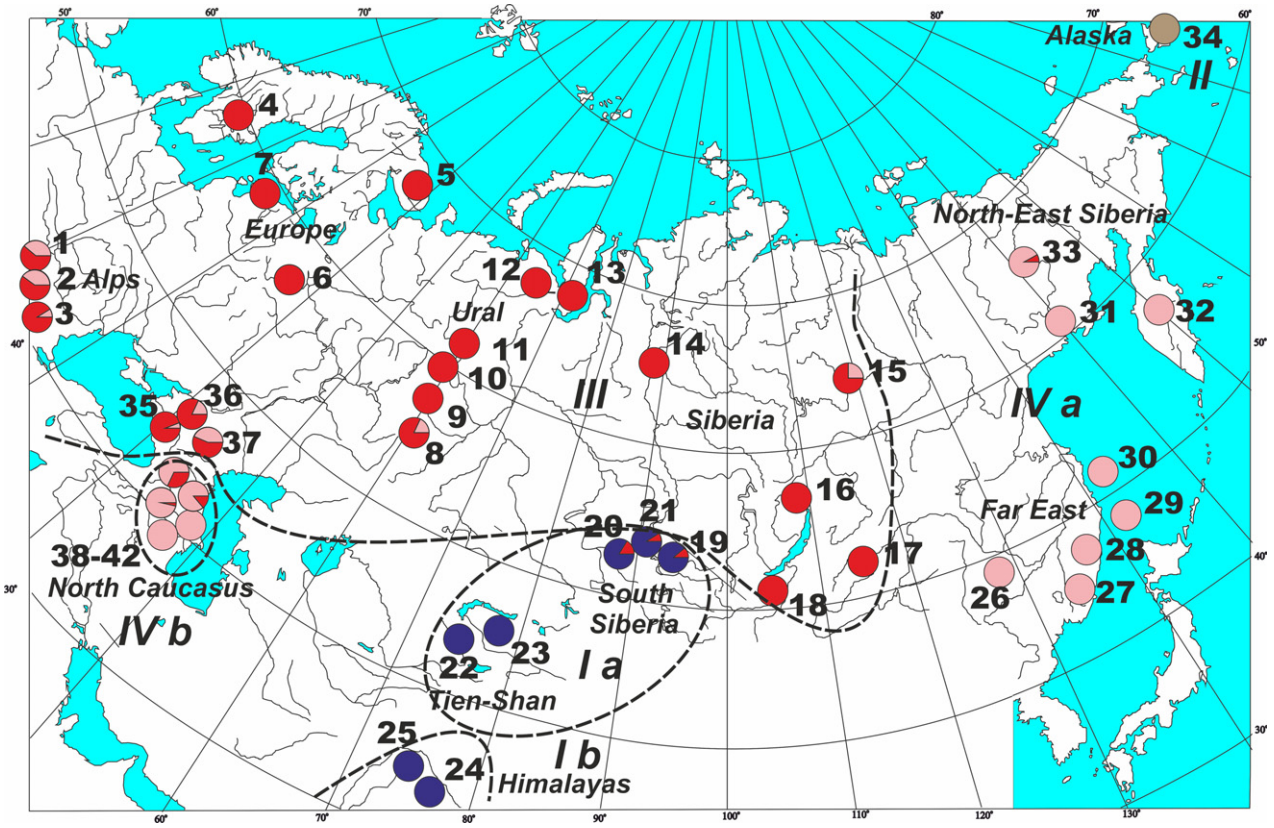

Figure 1 Geographical location of sampled populations of *Juniperus communis* in Eurasia and Alaska (for numbers see Appendix S1) and distribution of the main genetic lineages (GLs). GL1, blue: GL2, brown; GL3, red; GL4, pink. SAMOVA population groups are delineated by dashed line.

*16*S-*trn*A were analysed with RFLP in 622 individuals of *J. communis*. Among them fragment *16*S-*trn*A was additionally sequenced in 404 individuals and fragment *trn*T-*trn*L, which is less variable, was sequenced in 180 individuals.

Data analyses

The sequences of the two regions examined were edited and manually aligned in BioEdit 7.2.5 (Hall, 1999). All sequences have been deposited in GenBank under accession numbers KX066269–KX066350, and KX066351–KX066362. In all subsequent analyses each indel was treated as a single 1/0 character independently of its length.

Spatial analysis of molecular variance (samova 1.0; Dupanloup *et al.*, 2002) was used to identify population clusters. This method aims to identify clusters of populations that are maximally differentiated. A simulated annealing process is run until obtaining the configuration of K groups that maximizes F_{CT} (the proportion of total genetic variance due to differences among groups of populations) and that minimizes the number of single-population groups. The analyses were repeated with increasing values of K (2-15), until the F_{CT} values reached a plateau.

The existence of a phylogeographical structure was tested following Pons & Petit (1996). Two measures of population differentiation (G_{ST}, N_{ST}) were obtained using permut 1.0 (http://www.pierroton.inra.fr/genetics/labo/Software/) with 1000 permutations, and compared to each other.

Arlequin 3.5.1.2 (Excoffier & Lischer, 2010) was used to calculate molecular diversity indices, including the number of haplotypes, haplotype diversity (H_e) and nucleotide diversity (π), as well as for analyses of molecular variance (AMOVA, Excoffier *et al.*, 1992). Measures of DNA divergence between populations and groups, F_{ST}, were calculated and their significance level was determined using 10,000 permutations. The groups used here were the ones identified with samova. Relationships between cpDNA haplotypes were investigated using network 4.6.1.2 (Bandelt *et al.*, 1999). The phylogenetic relationships among cpDNA haplotypes were evaluated with MrBayes 3.2.3. (Huelsenbeck & Ronquist, 2001). *Juniperus rigida* was used as an outgroup, and the GTR+G+I nucleotide substitution model was chosen with the Modeltest 3.7 software program and Akaike's information criterion (Posada & Crandall, 1998). Two independent Markov chain Monte Carlo (MCMC) chains with a length of 10 million steps were used, saving the parameters every 1000 steps. To monitor the adequacy of the length of the Markov chain, the average standard deviation of split frequencies was used. After 2.5 million generations, the value of this measure was < 0.01, and due to this, the first 2.5 million generations were discarded as burn-in. The remaining 7500 trees were used to construct a consensus tree.

To detect historical population expansion events in *J. communis*, mismatch distributions were calculated using Arlequin 3.5.1.2. A total of 1000 parametric bootstrap replicates were used to generate an expected distribution under a model of sudden demographic expansion (Rogest & Harpending, 1992). Fu's F_s and raggedness indices were used to test for deviation from neutrality.

We used environmental niche modelling based on the current distribution of *J. communis* in Eurasia and present and past climatic parameter distributions to reconstruct putative species ranges during the Last Glacial Maximum (LGM: 21,000 years before present) and the Last Interglacial (LIG: *c.* 120,000 years before present). To do this we used the machine learning method based on maximum entropy implemented in the program maxent 3.3.3 (Phillips *et al.*, 2006). *Juniperus communis* distribution data (from Europe) were retrieved from the GBIF database (http://www.gbif.org, last accessed 20 February 2012). The data set was expanded by the addition of 66 of our field records and 396 records from the territory of the former USSR (Sokolov *et al.*, 1977). To reduce number of species occurrence to 10,000 we used thin.max.r R script, a part of the ENMTools 1.4.4 (Warren *et al.*, 2010).

The environmental data describing the baseline climate (19 BioClim layers for the period 1950–2000 at a spatial resolution of 2.5 arc min), the LGM climate (BioClim layers derived from the Coupled Model Intercomparison Project Phase 5) and for the Last Interglacial climate (Otto-Bliesner *et al.*, 2006) were retrieved from the WorldClim database (Hijmans *et al.*, 2005). To reduce an effect of association between climate parameters, we computed correlation between all pairs of 19 parameters for the geographical points of the species occurrence. For the groups of parameters with correlation 0.8 or more, we included only one parameter, leaving nine: bio1 – Annual Mean Temperature, bio2 – Mean Diurnal Range, bio3 – Isothermality, bio4 – Temperature Seasonality, bio5 – Max Temperature of Warmest Month, bio8 – Mean Temperature of Wettest Quarter, bio12 – Annual Precipitation, bio15 – Precipitation Seasonality, bio18 – Precipitation of Warmest Quarter. To account for bias associated with different accessibility of sites, a layer with accessibility was used. It was prepared by assigning a value of 20 to all Western European sites. The value of 2 was assigned to the rest of Eurasia. Default settings of Maxent were used.

RESULTS

DNA sequence variation

Juniperus communis is highly variable: 84 cpDNA haplotypes were identified based on the combined polymorphisms at the *trn*T-*trn*L and *16*S-*trn*A loci. The total alignment length of the sequences was 2857 bp across all 404 individuals from the 42 populations. Sequences of the *trn*T-*trn*L fragment included polymorphic sites apparently arising from point mutations. A high level of polymorphism was detected in the *16*S-*trn*A spacer and represented insertions, deletions and substitutions. Haplotype frequencies in each population are presented in Appendix S1. More than half of the haplotypes

were rare, including 36 singletons and 11 haplotypes observed only twice or three times in a single population (see Appendix S2). The number of haplotypes and haplotype diversity (H_e) per population ranged from 1 to 13 and from 0 to 0.96, respectively, with the greatest values for the number of haplotypes in North Tyrol. In addition to the Alps, some northern populations (Yakutia, Polar Ural, Estonia, Sweden, Kola Peninsula) and the Tver population also showed high haplotype diversity levels (Fig. 2). The Swedish population also has high nucleotide diversity π (0.18). The distribution of the number of private haplotypes (haplotypes specific to a particular population) did not show any clear geographical pattern: a high number was observed in the Polar Urals (4) but also in the Alps (North Tyrol; 7; see Fig. 3).

Phylogenetic relationships among haplotypes and their geographical distribution

The distribution of the 84 haplotypes was not random and showed strong geographical patterns. Populations were highly differentiated (G_{ST} = 0.640). The value for N_{ST} (0.796) was higher than that for G_{ST} ($P < 0.001$), indicating significant phylogeographical structure across the entire distribution of the species and low gene-flow relative to the mutation rate.

The Bayesian phylogenetic tree of haplotypes includes four main genetic lineages (GL) (Figs 1 and 4). NETWORK shows the same structure as the tree (see Appendix S3). GL1 is the most basal and includes two sister groups of haplotypes of *J. communis* var. *saxatilis* – specific to Central Asia (Altai, Sayan, Shoria, Tien Shan) and the Himalayas respectively. The rest of the tree splits into the GL2 group including haplotypes H1 and H2 of *J. communis* var. *depressa* from North America (north-west Alaska), and a large clade that consists of GL3, GL4 and a few divergent haplotypes.

GL3 is the largest lineage containing haplotypes specific to different varieties of *J. communis* – var. *communis*, *saxatilis* and *oblonga* – and to different geographical regions of the major part of Northern Eurasia – Europe, Ural and most of Siberia, and also the central part of North Caucasus. GL3 is defined by the presence of a large insert in the *16S* fragment, suggesting that GL3 is monophyletic. Many haplotypes and their clades within G3 have a regional distribution. For example, the clade including haplotypes 22, 23, 24 is specific to the Urals, the clade consisting of haplotypes 52 and 72 to the central part of North Caucasus, and haplotypes 17, 18, 20, 23 to northern Europe and the Polar Urals.

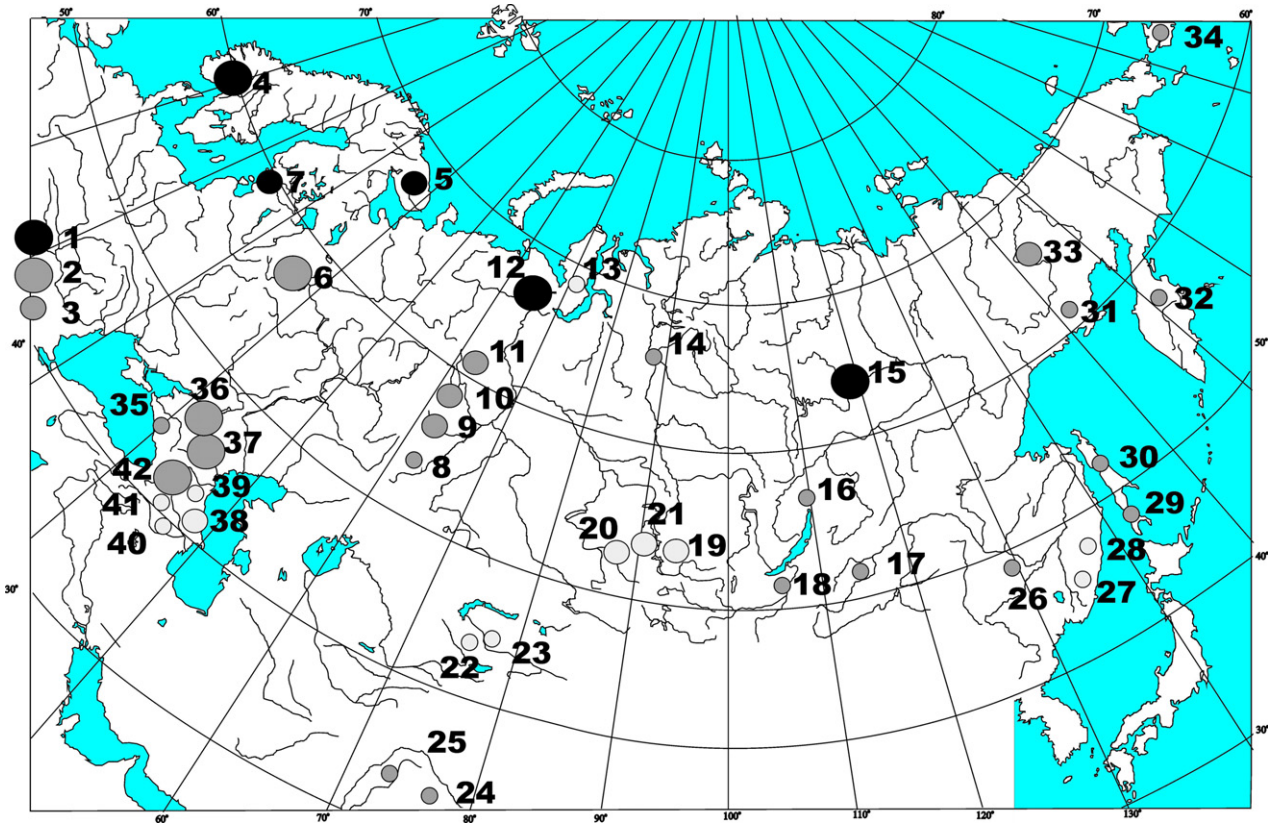

Figure 2 Intra-population diversity H_e and π based on 84 chloroplast DNA haplotypes in populations of *Juniperus communis*. The colour and the size refer to the values of H_e and π, respectively. H_e – white circles – 0–0.49, grey circles – 0.55–0.83, black circles – 0.85–0.96. π – small circles: 0–0.028, median circles – 0.031–0.051, large circles – 0.056–0.18.

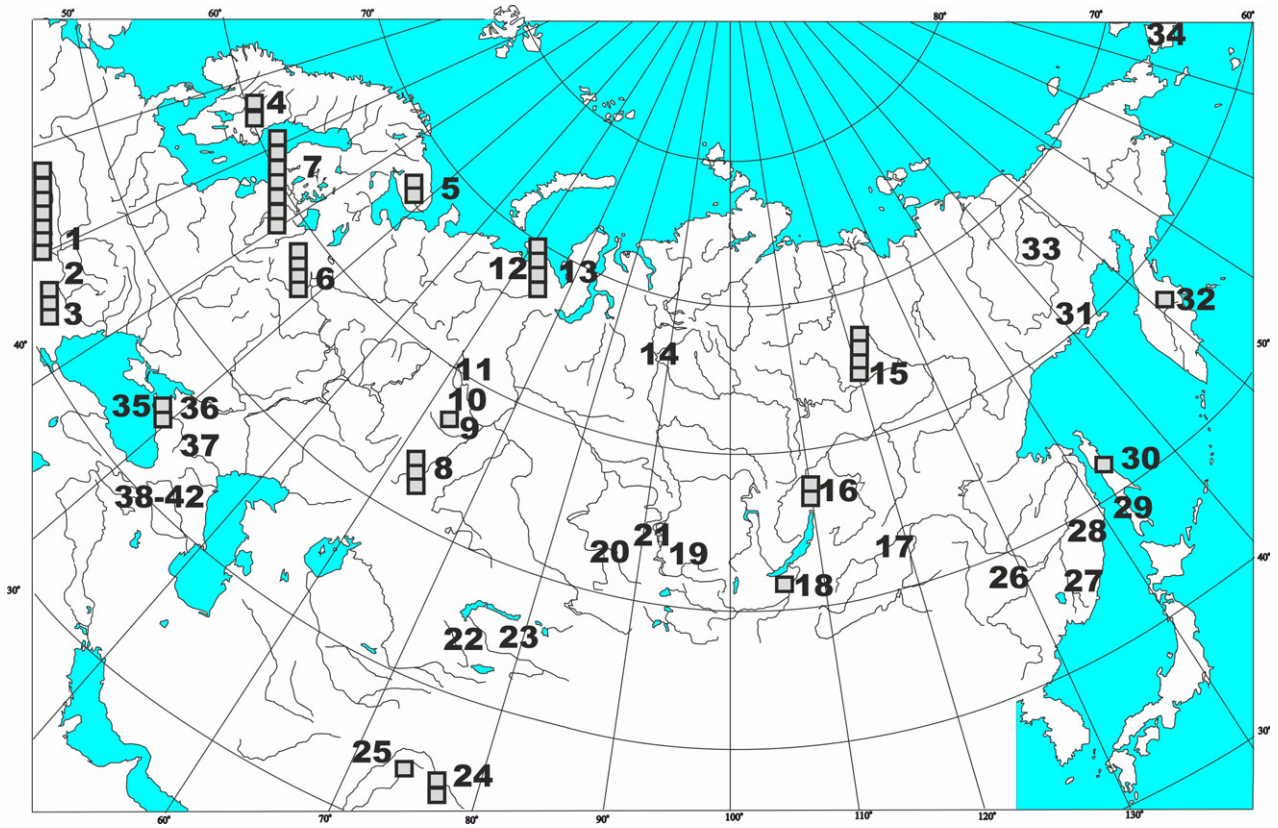

Figure 3 Distribution of private haplotypes in populations of *Juniperus communis*. Bar chart represents the number of private haplotypes in a given population.

GL4 splits into several haplogroups specific to the following geographical regions: 1 – north-east of Siberia including Kamchatka and south of Russian Far East and Sakhalin Island, 2 – eastern part of North Caucasus, 3 – the Alps (mainly the population of North Tyrol), plus a single haplotype H61 found in the South Urals.

The analysis of molecular variance (AMOVA) revealed significant differences across the total distribution of the species (F_{ST} = 0.76; P < 0.001).

The SAMOVA groups are congruent with those obtained from the Bayesian tree (Fig. 1 and 4). The F_{CT} value reached a plateau (0.737) under K = 6 (see Appendix S3). The large SAMOVA group III including the populations of Northern Eurasia and of the central part of North Caucasus is almost entirely composed of closely related haplotypes belonging to the GL3 lineage. Population differentiation of *J. communis* in these groups was markedly lower than that in the entire range (F_{ST} = 0.28; P < 0.001) in spite of the vast geographical range, but within-population haplotype diversity (H_e = 0.77) is higher than average (H_e = 0.60). The value for N_{ST} (0.204) was not higher than G_{ST} (0.191, P < 0.25).

SAMOVA groups Ia and Ib (Central Asia) correspond to GL1, group II (Alaska) corresponds to GL2, and group IVa (North East Asia and Russian Far East) and group IVb (east part of North Caucasus) to GL4. All of these groups are strongly differentiated: they have few haplotypes in common, and haplotype diversity is low: Central Asia – H_e = 0.48; Alaska – 0.55; North East Asia and Russian Far East – 0.50; Caucasus – 0.47.

Inference of demographic processes

A mismatch distribution analysis was conducted for the major lines obtained with MRBAYES. Results of Fu's F_s tests and the raggedness statistic of the mismatch distribution analysis are provided in Table 1. In GL1, GL2 and GL4, the sum of squared deviations (SSD) between the observed mismatch distribution and the expectation under the expansion model and the raggedness index values are rather high and suggest a stable population size. For GL4 goodness-of-fit test reject a sudden expansion model. Only in GL3, the highly significant and large negative Fu's F_s indicates a deviation from neutrality, and supports population expansion.

Species distribution modeling

The environmental niche modelling for *J. communis* at the LGM and the LIG produced putative geographical ranges which were similar to the current distribution (Fig. 5). The regions with suitable climatic conditions (marked as green) occupy a large proportion of the current range. The most favourable areas appeared in Europe, the Urals and the eastern part of the range – Russian Far East, North East Asia as

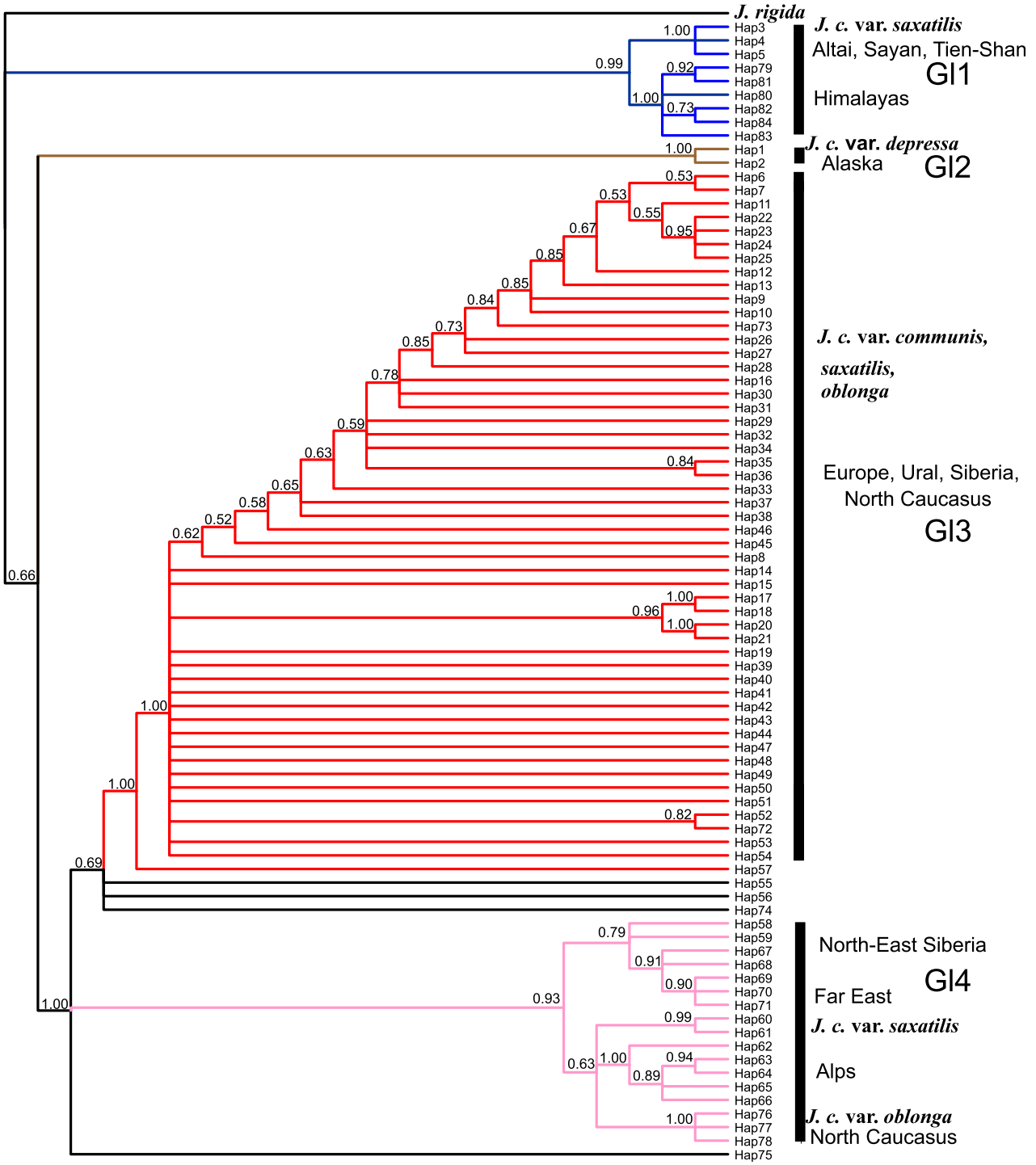

Figure 4 The Bayesian tree phylogeny with genetic lines (GLs) based on the 84 cpDNA haplotypes of *Juniperus communis* in Eurasia with *Juniperus rigida* as outgroup. Four major clades GL1 – GL4 are marked with blue, brown, red and pink correspondingly.

well as the mountainous region of Central Asia (Tien-Shan, Himalayas).

DISCUSSION

Our results suggest that *J. communis* has a complex demographic history with multiple post-glacial colonization events in different parts of its range. This study revealed a remarkably high level of population differentiation with F_{ST} = 0.76 in the total data set and high genetic diversity within populations of *J. communis*, which is expected for such a widespread species. Additionally, F_{ST}, with a value of 0.28, was much lower in the 'northern group'. Genetic structure at chloroplast and nuclear DNA are fairly congruent

Table 1 Results of mismatch distribution analysis and tests of demographic expansion of main genetic lineages of *Juniperus communis* in Eurasia.

GL	Tau	SSD	HRI	*P*-value	F_s	*P*-value
GL1	19.36	0.13	0.26	0.16	10.40	0.98
GL2	0.85	0.034	0.31	0.095	1.20	0.66
GL3	4.21	0.0044	0.010	0.82	−24.85	0.000
GL4	10.63	0.075	0.18	0.000	0.087	0.57

GL – genetic lineage; SSD – sum of squared deviation under expansion model; HRI – Harpending's raggedness index; *Fs* – Fu's *Fs* test statistic.

(Hantemirova *et al.*, 2012) and are generally compatible with a model where southern populations had a limited contribution to northern ones from which they strongly diverged, suggesting an ancient split and limited gene flow thereafter. The pattern is therefore strikingly similar to that observed previously in *Larix* (Semerikov *et al.*, 2013). Our data also show that the pattern observed today does not solely reflect the LGM, but also more ancient demographic events. In our data and according to the biogeographical study of Mao *et al.* (2010), the central Asian cpDNA lineage GL1 is very distinct and well differentiated from the others, so that it likely represents old, fragmented, relic populations that did not contribute to the recolonization of the north. Perhaps somewhat more unexpectedly the Caucasus populations contributed in a very minor way to the last wave of post-glacial recolonization given the post-glacial refugia status that has generally been ascribed to Caucasus for other species, for example, in oaks (Dumolin-Lapegue *et al.*, 1997), in Eurasian shrew (Dubey *et al.*, 2006) and in lynx (Rueness *et al.*, 2014). On the contrary, in juniper the Caucasus likely was not the source of post-glacial recolonization but rather the recipient of immigrants from the north as in hedgehog (Seddon *et al.*, 2002) or Scots pine (V.L. Semerikov, unpublished). The isolation of populations from the eastern part of North Caucasus from the rest of the species range is likely

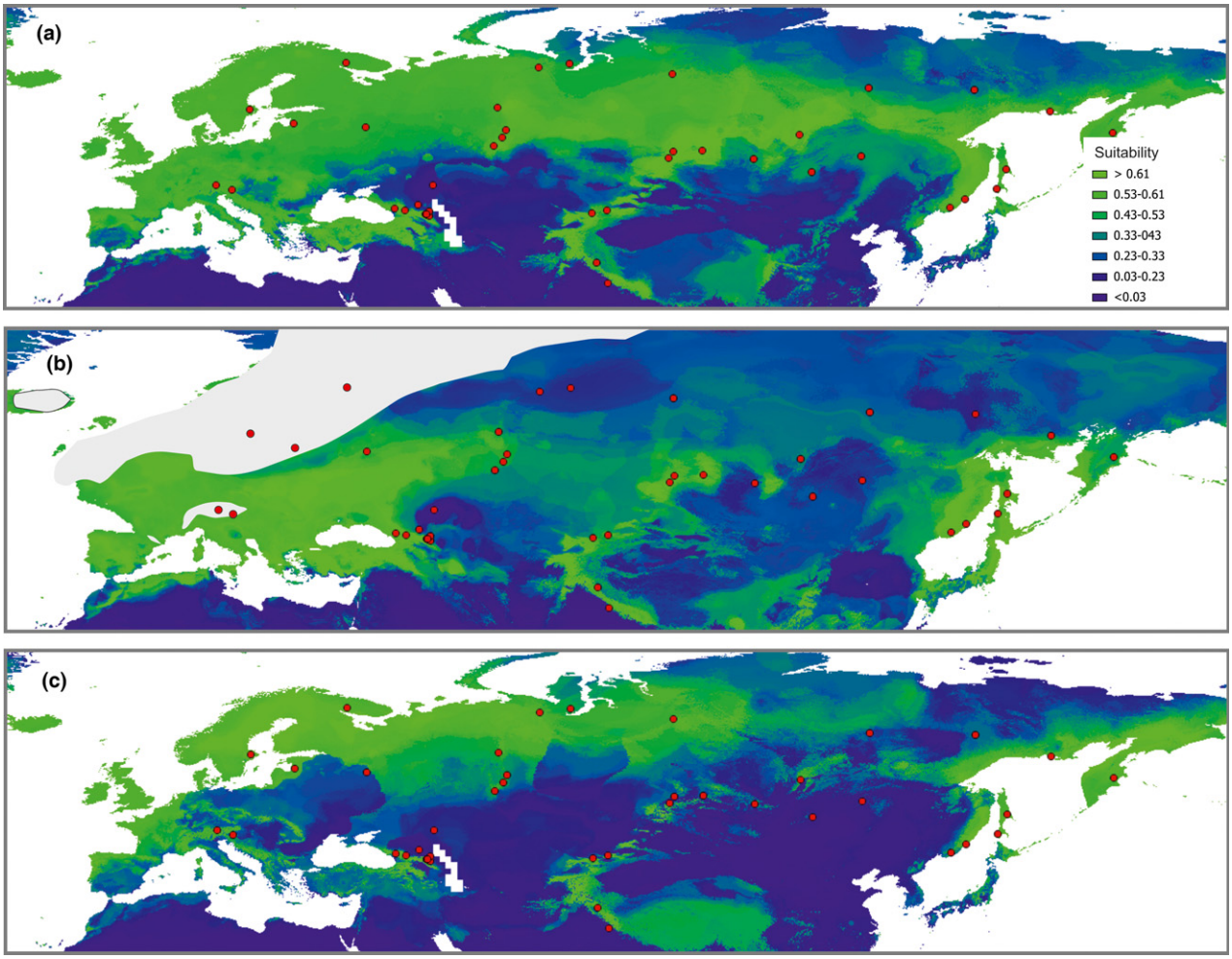

Figure 5 Distribution models *J. communis* for (a) the present, (b) the Last Glacial Maximum (*c.* 21,000 years before present) and (c) interglacial period (*c.* 120,000 years BP). Green colour indicates the most suitable modelled habitat, whereas dark blue indicates a low suitability. Sampled populations of *J. communis* are represented by red circles. The limits of the northern Europe – Barents sea ice sheet (after Hughes *et al.*, 2016) and Alpine and Pyrenean glaciers (after Ehlers & Gibbard, 2004) at the LGM are also indicated by light grey.

due to the topography of the region and the arid climate that prevailed during long periods. Populations of the more humid central part of North Caucasus had more possibilities for expansion and experienced more gene flow from group GL3, accumulating European haplotypes (although there are no shared haplotypes with European populations, suggesting an absence of recent contact). Less surprisingly, and in line with previous studies (Polezhaeva *et al.*, 2010; Hantemirova *et al.*, 2012), populations from Beringia and the Russian Far East also constituted a separate group. Adams & Pandey (2003) studied only two juniper samples from one population of this region, namely, from Kamchatka. This population was also very distinct from European and Siberian populations. For many species Beringia (including north-east of Asia) was an ancient refugium isolated from the rest of the territory by the Verkhoyansky mountain ridge (Eidesen *et al.*, 2013). Macrofossils and pollen data suggest that the current boreal vegetation probably survived in isolated refugia located in Beringia (Abbott & Brochmann, 2003; Brubaker *et al.*, 2005).

While other groups did not significantly depart from the standard neutral model, in the 'northern' group Fu's F_s statistics indicated strong deviation from neutrality or recent population expansion. As we noted above, GL3 is defined by the presence of a large insert in the *16*S fragment suggesting that GL3 is monophyletic. Taken together, the low differentiation of the 'northern group' (F_{ST} = 28%), significant F_s and monophyly of GL3 suggest a relatively recent expansion of northern populations from a single source. Given the high diversity of the Alpine populations and considering that Southern European mountains were a major refugium for some boreal species [e.g. Norway spruce (Lagercrantz & Ryman, 1990)], we speculate that the Alpine region was this source. At the same time, ecological niche modeling shows that survival of common juniper was possible during LGM over the northern part of the current range (Fig. 5), suggesting that a hypothetical initial 'out-of-Alps' migration, if at all, took place before the LGM. *Juniperus communis* was likely able to survive during the cold episodes of the Pleistocene at high latitudes in many periglacial microrefugia, spreading out and admixing during warm intervals over and over again. Apart from the Urals, populations in the north of Europe (Sweden, Estonia, Kola Peninsula) and Asia (Yakutia) that show high genetic diversity and many rare haplotypes could also have served as cryptic (micro-)refugia. The presence of trees at high latitudes during the early Holocene is supported by the macrofossils and stomata in the fossil record from the area east of the margin of the Scandinavian ice sheet and west of the Ural Mountains during the early Holocene (Välirantä *et al.*, 2011) and in the north of West Siberia during the LGM (Binney *et al.*, 2009; Kosintsev *et al.*, 2012). Broad distribution of cold-resistant trees and shrubs during the LGM is also suggested by the abundance of large herbivores (mammoth, rhinoceros, horse, reindeer, bison, etc.), the diet of which included branches (Ukraintseva, 1993; Boeskorov *et al.*, 2011; Kosintsev *et al.*, 2012). At the same time these animals could suppress the development of dense forest stands, promoting instead shrubs including juniper (Michalczyk *et al.*, 2010). Juniper is not exceptional in this respect, and high genetic diversity in northern Europe was also observed for *Picea abies* (Tollefsrud *et al.*, 2008), *Dryas octopetala* (Skrede *et al.*, 2006), *Vaccinium uliginosum* and *Vaccinium vitis-idaea* (Eidesen *et al.*, 2007, 2013). A unique mtDNA haplotype in *P. abies* populations of NW Norway also suggested survival in LGM at high latitudes (Parducci *et al.*, 2012). A refugium in Yakutia is supported by fossil evidence of *Dryas* from the middle to late Weichselian (Kienast *et al.*, 2001, 2005).

The population expansion observed in *J. communis* was also detected in the wood lemming (Fedorov *et al.*, 2008), the brown bear (Korsten *et al.*, 2009), Siberian larch (Semerikov *et al.*, 2013), black alder (Havrdova *et al.*, 2015) and the Siberian salamander (Malyarchuk *et al.*, 2013). In many of these cases the age of population expansion, when it was estimated, exceeds the age of the LGM (in the wood lemming 125,000 years, in the Siberian salamander before 250,000 years, and in Siberian larch 500,000 years). What these cases also have in common is a clearly subdivided population, with secondary centres of distribution – refugia associated with more recent glaciation. In addition to one (or more) groups of 'young' populations in the northern part of the range, in most species there were also populations characterized by deeply divergent genetic lines. These populations are located mainly on the southern periphery of the range. Like the southern populations of *J. communis* they did not participate in the re-colonization of the northern regions. Thus, in summary, in most of these cases, the 'rear edge' hypothesis is only partly consistent with the observed distribution of genetic variation.

CONCLUSIONS

In summary, chloroplast DNA diversity in *J. communis* revealed a pattern similar to that previously observed in other boreal species with large Eurasian ranges: several deeply divergent groups of populations located in southern regions and a very large group covering most of Europe and Siberia. The latter was recently colonized after the last glacial from cryptic refugia in its midst, with no or limited contribution from the southern refugia, notably in the Alps. In general, it is important to keep in mind that the patterns that are observed today are not just the result from the action of a single glacial maximum, but the combined effect of several cycles of retraction and expansion of species (see Tsuda *et al.*, 2016). In Eurasia, the documented persistence of microrefugia within the areas affected by glaciations certainly makes the story more complex than in the classical southern refugia paradigm that dominated the early days of phylogeographical studies. At the same time, the rapid development of whole genome tools, ancient DNA studies and a new generation of fossil maps should allow us to start to get a better apportionment of the contributions of the different refugia

as well as of the time frame under which the different waves of recolonization took place.

ACKNOWLEDGEMENTS

This work was financially supported by the Russian Foundation for Basic Research, grant 13-04-01793. We are grateful to A. Belyaev, A. Belyankin, A. Berkutenko, O. Chernyagina, M. Danilov, O. Dymshakova, I. Katz, D. Krivenko, V. Mazepa, A. Nikolaev, E. Nikolin, E. Pimenova, M. Polezhaeva, G. Radjabov, A. Surkov, B. Toktoraliev, V. Fedorov, E. Filippov, for help with material collections. The authors thank the members of the Centre for Himalayan scientific research, the Saint-Petersburg Association of Scientists and Scholars, and personally B.K. Gannibal for assistance with field data collection from the Himalayas.

REFERENCES

Abbott, R.J. & Brochmann, C. (2003) History and evolution of the arctic flora: in the footsteps of Eric Hultén. *Molecular Ecology*, **11**, 299–313.

Adams, R.P. (2014) *Junipers of the world: the genus Juniperus*, 4th edn. Trafford Publishing, Vancouver, B.C., 422 p.

Adams, R.P. & Pandey, R.N. (2003) Analysis of *Juniperus communis* and its varieties based on DNA fingerprinting. *Biochemical Systematics and Ecology*, **31**, 1271–1278.

Bandelt, H.J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenics. *Molecular Biology and Evolution*, **16**, 37–48.

Binney, H.A., Willis, K.J., Edwards, M.E., Bhagwat, S.A., Anderson, P.M., Andreev, A.A., Blaauw, M., Damblon, F., Haesaerts, P. & Kienast, F. (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews*, **28**, 2445–2464.

Boeskorov, G.G., Lazarev, P.A., Sher, A.V., Davydov, S.P., Bakulina, N.T., Shchelchkova, M.V., Binladen, J., Willerslev, E., Buigues, B. & Tikhonov, A.N. (2011) Woolly rhino discovery in the lower Kolyma River. *Quaternary Science Reviews*, **30**, 2262–2272.

Boratynski, A., Lewandowski, A., Boratinska, K., Monserrat, J.M. & Roma, A. (2009) High level of genetic differentiation of *Juniperus phoenicea* (Cupressaceae) in Mediterranean region: geographic implications. *Plant Systematics and Evolution*, **277**, 163–172.

Brubaker, L.B., Anderson, P.M. & Edwards, M.E. (2005) Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data. *Journal of Biogeography*, **32**, 833–848.

Devey, M.E., Bell, J.S., Smith, D.N., Neal, D.B. & Motan, G.F. (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers. *Theoretical and Applied Genetics*, **92**, 673–679.

Dubey, S., Zaitsev, M., Cosson, J.F., Abdukadier, A. & Vogel, P. (2006) Pliocene and Pleistocene diversification and multiple refugia in a Eurasian shrew (*Crocidura suaveolens* group). *Molecular Phylogenetics and Evolution*, **38**, 635–647.

Dumolin-Lapegue, S., Demesure, B., Fineschi, S., LeCorre, V. & Petit, R.J. (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics*, **146**, 1475–1487.

Dupanloup, I., Schneider, S. & Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Ehlers, J. & Gibbard, P.L. (2004) *Quaternary glaciations – extent and chronology. Part 1: Europe*. Elsevier, Amsterdam.

Eidesen, P.B., Alsos, I.G., Popp, M., Stensrud, Ø., Suda, J. & Brochman, C. (2007) Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species. *Molecular Ecology*, **16**, 3902–3925.

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**, 898–910.

Excoffier, L. & Lischer, H. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes – application to human mitochondrial-DNA restriction data. *Genetics*, **131**, 479–491.

Farjon, A.A. (2001) *World checklist and bibliography of conifers. England*, 2nd edn. The Royal Botanic Gardens, Kew.

Fedorov, V.B., Goropashnaya, A.V., Boeskorov, G.G. & Cook, J.A. (2008) Comparative phylogeography and demographic history of the wood lemming *(Myopus schisticolor):* implications for late Quaternary history of the taiga species in Eurasia. *Molecular Ecology*, **17**, 598–610.

Filipowicz, N., Piotrowski, A., Ochocka, R. & Aszemborska, M. (2006) The phytochemical and genetic survey of common and dwarf juniper (*Juniperus communis* and *Juniperus nana*) identifies chemical races and close taxonomic identity of the species. *Planta Medica*, **72**, 850–853.

Guo, Y.P., Chen, C.Y., Zhou, D.W. & Liu, J.Q. (2010) Allopatric divergence and regional range expansion of Juniperus sabina in China. *Journal of Systematics and Evolution*, **48**, 153–160.

Hall, T.A. (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series*, **41**, 95–98.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.

Hantemirova, E.V., Berkutenko, A.N. & Semerikov, V.L. (2012) Systematics and gene geography of *Juniperus communis* L. inferred from isoenzyme data. *Russian Journal of Genetics*, **48**, 920–926.

Havrdova, A., Douda, J., Krak, K., Vit, P., Hadincova, V., Zakravsky, P. & Mandak, B. (2015) Higher genetic

diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture. *Molecular Ecology*, **24**, 4759–4777.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 965–978.

Huelsenbeck, J.P. & Ronquist, F. (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754–755.

Hughes, A.L.C., Gyllencreutz, R., Lohne, O.S., Mangerud, J. & Svendsen, J.I. (2016) The last Eurasian ice sheets - a chronological database and time-slice reconstruction, DATED-1. *Boreas*, **45**, 1–45.

Kienast, F., Siegert, C., Dereviagin, A. & Mai, D.H. (2001) Climatic implications of Late Quaternary plant macrofossil assemblages from the Taymyr Peninsula, Siberia. *Global and Planetary Change*, **31**, 265–281.

Kienast, F., Schirrmeister, L., Siegert, C. & Tarasov, P. (2005) Palaeobotanical evidence for warm summers in the East Siberian Arctic during the last cold stage. *Quaternary Research*, **63**, 283–300.

Kondo, T., Sumura, Y., Kawahara, T. & Okamura, M. (1998) Inheritance of chloroplast and mitochondrial DNA in interspecific hybrids of *Chamaecyparis* spp. *Breeding Science*, **78**, 177–179.

Korsten, M., Ho, S.Y.W., Davison, J., Pahn, B., Vulla, E., Roht, M., Tumanov, I.L., Kojola, I., Anderson-Lilley, Z. & Ozolins, J. (2009) Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Molecular Ecology*, **18**, 1963–1979.

Kosintsev, P.A., Lapteva, E.G., Korona, O.M. & Zanina, O.G. (2012) Living environments and diet of the Mongochen mammoth, Gydan Peninsula, Russia. *Quaternary International*, **276–277**, 253–268.

Lagercrantz, U. & Ryman, N. (1990) Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. *Evolution*, **44**, 38–53.

Lascoux, M., Palme′, A.E., Cheddadi, R. & Latta, R. (2004) Impact of the Ice Ages on the genetic structure of trees and shrubs. *Philosophical Transaction of the Royal Society B: Biological Sciences*, **359**, 197–207.

Li, Z., Zhang, Q., Liu, J. & Källman, T. (2011) The Pleistocene demography of an alpine juniper of the Qinghai-Tibetian Plateau: *tabula rasa*, cryptic refugia or a something else? *Journal of Biogeography*, **38**, 31–43.

Maliouchenko, O., Palme, A.E., Buonamici, A., Vendramin, G.G. & Lascoux, M. (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. *Journal of Biogeography*, **34**, 1601–1610.

Malyarchuk, B., Derenko, M. & Denisova, G. (2013) Phylogeny and genetic history of the Siberian salamander (*Salamandrella keyserlingii*, Dybowski, 1870) inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **67**, 348–357.

Mao, K., Hao, G., Liu, J., Adams, R. & Milne, R. (2010) Diversification and biogeography of *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental dispersals. *New Phytologist*, **188**, 254–272.

Michalczyk, I.M., Opgenoorth, L., Luecke, Y., Huck, S. & Ziegenhagen, B. (2010) Genetic support for periglacial survival of *Juniperus communis* L. in Central Europe. *The Holocene*, **20**, 887–894.

Mogensen, H.L. (1996) The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany*, **83**, 383–404.

Neale, D.B., Marshall, K.A. & Sederoff, R.R. (1989) Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. *Proceedings of the National Academy of Sciences USA*, **86**, 9347–9349.

Neale, D.B., Marshall, K.A. & Harry, D.E. (1991) Paternal inheritance of chloroplast and mitochondrial DNA in incense-cedar (*Calocedrus decurrens* (Torr.) Florin). *Canadian Journal of Forest Research*, **21**, 717–720.

Opgenoorth, L., Vendramin, G.G., Mao, K., Miehe, G., Miehe, S., Liepelt, S., Liu, J. & Ziegenhagen, B. (2010) Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. *New Phytologist*, **185**, 332–342.

Otto-Bliesner, B.L., Marsha, S.J., Overpeck, J.T., Miller, G.H., Hu, A.X. & Mem, C.L.I.P. (2006) Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science*, **311**, 1751–1753.

Palme, A.E., Semerikov, V. & Lascoux, M. (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.

Parducci, L., Jorgensen, T., Tollefsrud, M.M. *et al.* (2012) Glacial survival of boreal trees in Northern Scandinavia. *Science*, **335**, 1083–1086.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B., Palmé, A., Martín, J.P., Rendell, S. & Vendramin, G.G. (2003) Glacial refugia: hotspot but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Polezhaeva, M.A., Lascoux, M. & Semerikov, V.L. (2010) Cytoplasmic DNA variation and biogeography of Larix Mill. in Northeast Asia. *Molecular Ecology*, **19**, 1239–1252.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Prost, S., Guralnick, R.P., Waltari, E., Fedorov, V.B., Kuzmina, E., Smirnov, N., Kolfschoten, T., Hofreiter, M. &

Vrieling, K. (2013) Losing ground: past history and future fate of Arctic small mammals in a changing climate. *Global Change Biology*, **19**, 1854–1864.

Rogest, A.R. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rueness, E.K., Naidenko, S., Trosvik, P. & Stenseth, N.C. (2014) Large-scale genetic structuring of a widely distributed carnivore – the Eurasian Lynx (*Lynx lynx*). *PLoS ONE*, **9**, e93675.

Seddon, J.M., Santucci, F., Reeve, N. & Hewitt, G.M. (2002) Caucasus Mountains divide postulated postglacial colonization routes in the white-breasted hedgehog, *Erinaceus concolor*. *Journal of Evolutionary Biology*, **15**, 463–467.

Semerikov, V.L., Semerikova, S.A., Polezhaeva, M.A., Kosintsev, P.A. & Lascoux, M. (2013) Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): a range-wide analysis of cytoplasmic markers. *Molecular Ecology*, **22**, 4958–4971.

Skrede, I., Eidesen, P.B., Portela, R.P. & Brochmann, C. (2006) Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Molecular Ecology*, **15**, 1827–1840.

Sokolov, S.I., Svjaseva, O.A. & Kubli, V.A. (1977) *Distribution ranges of trees and shrubs in the USSR, V.1*. Academy of Sciences of the USSR, Moscow (in Russian).

Taberlet, P.T., Geilly, L., Patou, G. & Bouvet, J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Thomas, P.A., El-Barghathi, M. & Polwart, A. (2007) Biological Flora of the British Isles: *Juniperus communis* L. *Journal of Ecology*, **95**, 1404–1440.

Tollefsrud, M.M., Kissling, R., Gugerli, F., Johnsen, Ø., Skrøppa, T., Cheddadi, R., Van Der Knaap, W.O., Latałowa, M., Terhürne-Berson, R., Litt, T., Geburek, T., Brochmann, C. & Sperisen, C. (2008) Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology*, **17**, 4134–4150.

Tsuda, Y., Chen, J., Stocks, M., Källman, M., Sønstebø, J.H., Parducci, L., Semerikov, V.L., Sperisen, C., Politov, D., Ronkainen, T., Väliranta, M., Vendramin, G.G., Tollefsrud, M.M. & Lascoux, M. (2016) The extent and meaning of hybridization and introgression between Siberian spruce (*Picea obovata*) and Norway spruce (*P. abies*): cryptic refugia as stepping stones to the west? *Molecular Ecology*, **25**, 2773–2789.

Ukraintseva, V.V. (1993) *Vegetation cover and environment of the "Mammoth Epoch" in Siberia*. The Mammoth site of Hot Springs South Dakota, Hot Springs, South Dakota.

Välirantä, M., Kaakinen, A., Kuhry, P., Kultti, S., Salonen, J.S. & Seppä, H. (2011) Scattered late- glacial and early Holocene tree populations as dispersal nuclei for forest development in north- eastern European Russia. *Journal of Biogeography*, **38**, 922–932.

Vargas, P. (2003) Molecular evidence for multiple diversification patterns of alpine plants in the Mediterranean Europe. *Taxon*, **52**, 463–476.

Vogelstein, B. & Gillespie, D. (1979) Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences USA*, **76**, 615.

Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.

Willis, K.J. & Van Andel, T.H. (2004) Trees or not trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, **23**, 2269–2287.

Zhang, Q., Chiang, T.Y., George, M., Liu, J.Q. & Abbott, R.J. (2005) Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. *Molecular Ecology*, **14**, 3513–3524.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 *Juniperus communis* populations.
Appendix S2 Haplotype distribution in juniper populations.
Appendix S3 Haplotype network and F_{CT} plot.

BIOSKETCH

Elena V. Hantemirova is a researcher in the fields of population genetics of plants, phylogeography and geobotany.

Author contributions: E.H., V.S. and B.H. conceived the study, and analysed the data; S.K., B.H. and A.M. collected data; E.H., B.H., V.S. and M.L. wrote the paper.

Editor: Peter Linder