CZECH UNIVERSITY OF LIFE SCIENCES IN PRAGUE

FACULTY OF ENVIRONMENTAL SCIENCES

Department of Ecology

Postglacial colonization of black alder (*Alnus glutinosa*) and grey alder (*Alnus incana*) in Europe

Postglaciální kolonizace olše lepkavé (*Alnus glutinosa*) a olše šedé (*Alnus incana*) v Evropě

PhD Thesis

AUTHOR: Ing. Alena Havrdová (Drašnarová)

SUPERVISOR: doc. Mgr. Bohumil Mandák, Ph.D.

CONSULTANT: Ing. Jan Douda, Ph.D.

2015

DECLARATION

I thereby declare that I wrote the PhD thesis by myself using the results of my own work or collaborative work with my colleagues and with help of other publication resources which are properly cited.

In Prague……………..

....

PODĚKOVÁNÍ (ACKNOWLEDGEMENTS IN CZECH)

Chtěla bych poděkovat školiteli Bohoušovi Mandákovi a konzultantovi Honzovi Doudovi za příležitost pracovat na projektu zabývajícím se postglaciální historií olší, a tím možnost poznat spoustu nového a zajímavého. Velké díky patří i dalším členům týmu, Věře Hadincové, Janě Doudové, Karolovi Krakovi, Petrovi Zákravskému a Petrovi Vítovi. Celému týmu děkuji za spoustu cenných rad, připomínek a konstruktivních diskuzí, bez kterých by tato práce nevznikla. Dále bych ráda poděkovala Martině Hadincové a Evě Ničové za pomoc v genetické laboratoři. Také děkuji celé své rodině, že mě během mého studia podporovali.

CONTENT

1. INTRODUCTION **5**

Quaternary as the key period for European biota 5

Different theories about survival of tree species during the LPG period in Europe 6

Classical southern refugium theory 6

Concept of cryptic refugia 8

Northern refugia are rather mystic than cryptic 13

What do molecular data tell us about postglacial history of European woody species? 16

Key results from phylogeographic studies 16

Problems of phylogeographic studies 18

Future directions of phylogeographic studies 19

2. BLACK ALDER (*ALNUS GLUTINOSA*) AND GREY ALDER (*ALNUS INCANA*) **20**

Why Alnus glutinosa and A. incana are suitable species for new study? 20

Previous knowledge about A. glutinosa and A. incana 21

Study species 21

Upper Pleistocene and Holocene *Alnus* history 23

Aims, questions and hypothesis of dissertation thesis 24

References 27

3. SCIENTIFIC PAPERS **36**

Paper I: Migration patterns of subgenus Alnus based on papaeoecological records 37

Paper II: SSR markers for Alnus glutinosa and Alnus incana 51

Paper III: Cytotypes of Alnus glutinosa in Europe 60

Paper IV: Postglacial history of Alnus incana 103

Paper V: Postglacial history of Alnus glutinosa 138

4. PRINCIPAL CONCLUSIONS **172**

Key results about migration patterns of subgenus Alnus in Europe from palaeoecological records 172

Key results about the distribution and origins of cytotypes of Alnus glutinosa 173

Key results about postglacial history of Alnus incana in Europe 174

Key results about postglacial history of Alnus glutinosa in Europe 176

New insights into postglacial history of European woody species 178

Position of refugia 178

Migration routes 180

Character of migration 181

References 183

5. SOUHRN (SUMMARY IN CZECH) **188**

Nejdůležitější výsledky o charakteru migrace podrodu Alnus v Evropě na základě paleoekologických dat 188

Nejdůležitější výsledky o rozšíření a původu cytotypů druhu Alnus glutinosa 189

Nejdůležitější výsledky o postglaciální historii druhu Alnus incana v Evropě 189

Nejdůležitější výsledky o postglaciální historii druhu Alnus glutinosa v Evropě 190
6. SUPPORTING INFORMATION **191**
7. LIST OF PUBLICATIONS **242**

1. INTRODUCTION

Quaternary as the key period for European biota

Current species distribution in Europe was mainly influenced by massive climatic and environmental changes during **the Quaternary** period (Hofreiter & Stewart 2009). Alteration of glacials and interglacials caused contractions and expansions of species ranges (Hewitt 1996; Taberlet *et al.* 1998). Last Ice Age, especially the Late Pleniglacial (LPG), was an interval of the most extreme glacial conditions and, by extension, the maximum contraction of tree populations (Tzedakis *et al.* 2013). This interval, which lasted from 15 000 to 24 000 cal. yr BP (Tzedakis *et al.* 2013), certainly had major effect on current distribution of European species (e.g. Bennett *et al.* 1991; Taberlet 1998; Hewitt 1999). It effects both temperate and boreal tree species, but resulting in different distributional changes. While temperate species draw away to southern refugia during glaciation and subsequently have increased their ranges after warming (Palmé & Vendramin 2002; Petit *et al.* 2002; Grivet & Petit 2003; Tzedakis *et al.* 2013), boreal species were probably distributed more widely during the LPG (Tarasov *et al.* 2000; Palmé *et al.* 2003 a, b), and could survive also in northern parts of Europe but the most cold-tolerant of them have retreated in postglacial times (Höhn *et al.* 2009).

During the LPG period the climate was strongly continental with cold winters and hot and dry summers (except the mountain regions). **Temperatures were 8°C colder on average** in Europe (Lomosino *et al.* 2006). As a consequence, north of Europe was covered by the extensive **ice sheet** and the most of mountain ranges had ice caps. The large amount of water bound in glaciers resulted in **lowered sea-levels by 130 m** on average compared to the current situation (Lambeck 2004) and formation of **land bridges** in several parts of Europe (Hewitt 2000), e.g. British Isles being connected to the European mainland, Corsica to Sardinia and the Apennine Peninsula to the Balkan Peninsula. In addition, **atmospheric CO_2 concentration** was more than 35% lower than today decreasing plant productivity and water-use efficiency and resulting in formation of sparser vegetation (Cowling & Sykes 1999; Ward *et al.* 2005). The strong **LPG wind** speeds might have also enhanced community patchiness restricting them to the less exposed areas (Leroy & Arpe 2007). Without any doubt, this had to lead to large scale changes in species distributions, when some species went **extinct** over large

parts of their distribution range, some **migrated** to new locations, some survived in **refugia** and then expanded to north after glacier retrieve (Webb & Bartlein 1992; Comes & Kadereit 1998; Hewitt 1999, 2000).

Understanding which environmental conditions can species tolerate and finding location with their long-term stable survival is of particular importance for conservation and management of genetic resources as well as for prediction of future distribution of species under climate change (Provan & Bennett 2008; Stewart *et al.* 2010; Keppel *et al.* 2012). Despite huge amount of studies, there are still many doubts about species survival during the LPG and species distributional changes after subsequent warming. Specifically, detailed directions of migration routes and existence of cryptic northern refugia are purely clarified.

Different theories concerning survival of tree species during the LPG period in Europe

During last few decades, different theories concerning survival of tree species during the LPG period in Europe were proposed (Huntley & Birks 1983; Bennett *et al.* 1991; Frenzel *et al.* 1992; Willis *et al.* 2000; Stewart & Lister 2001; Parducci *et al.* 2012; Tzedakis *et al.* 2013).

- **Classical southern refugium theory**

First concept known as **"tree-less tundra model"** or the classical **"southern refugium theory"** was based on palaeoecological records and first genetic studies (Huntley & Birks 1983; Bennett *et al.* 1991; Frenzel *et al.* 1992; Hewitt 1996, 1999; Taberlet *et al.* 1998). It described Europe during LPG as landscape with **inhospitable north and favourable south**. According to this concept north was covered by ice, plain of permafrost and tundra vegetation. It was followed by cold steppe dominated by grasses (Poaceae) and wormwoods (*Artemisia* spp.) which occurred in central parts from ca. 50° and changed into Mediterranean and Sub-Mediterranean steppe of mountain regions covering southern peninsulas (Frenzel *et al.* 1992, Fig. 1). Most of tree species survived unfavourable times in small isolated populations in southern parts of Europe, specifically in lowlands and mid-altitude sites in the southern mountains (Bennett *et al.* 1991; Taberlet *et al.* 1998; Médail & Diadema 2009). There are only few indices that restricted boreal populations (e. g. *Salix* spp. and *Picea abies*) might

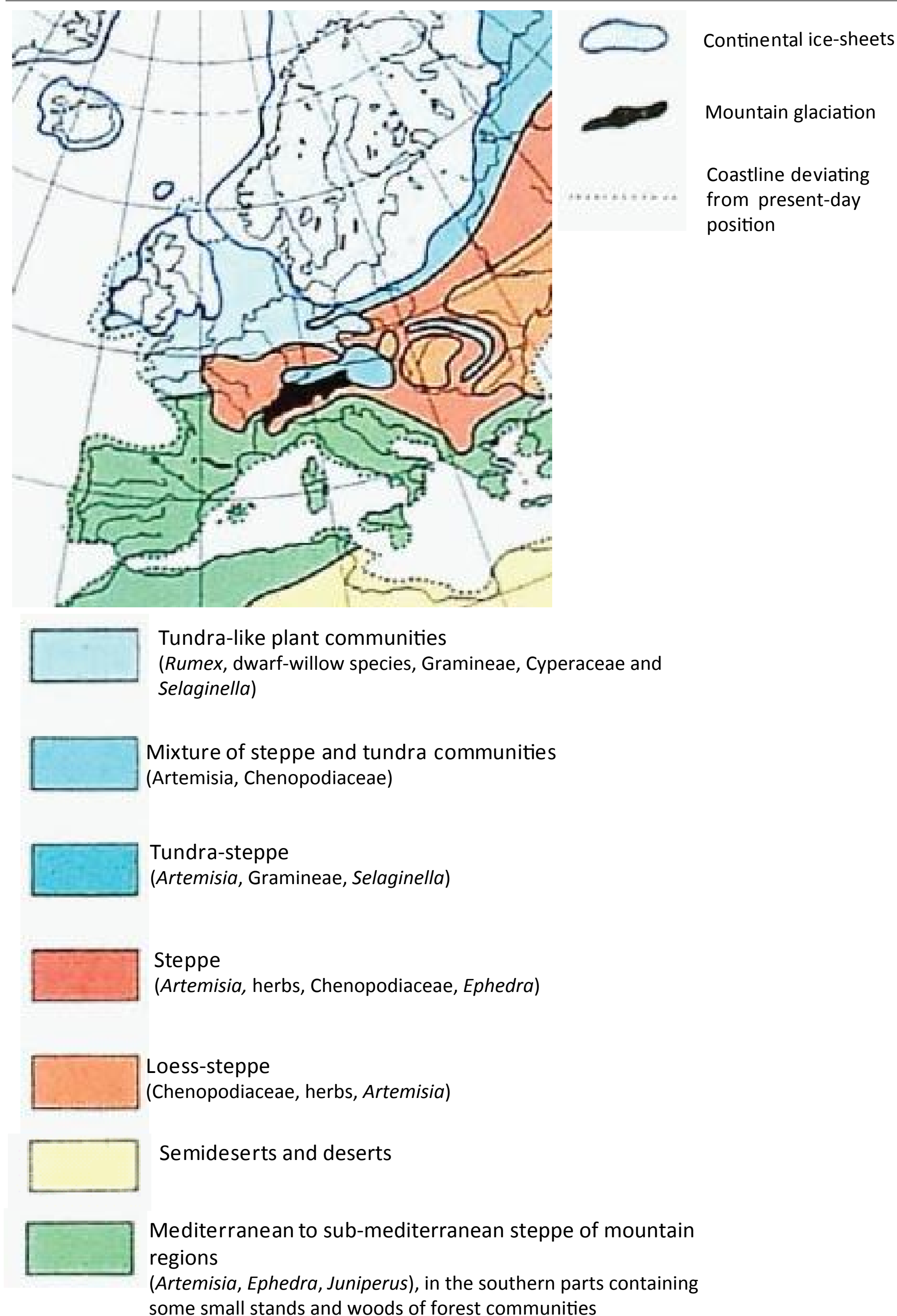


Fig. 1: Reconstruction of vegetation of maximum cooling of the last glaciation (about 20 000 to 18 000 yr BP) supporting southern refugium theory. Only in southern parts (marked with green) trees could survive during LPG. Adapted from Frenzel (1992).

have also survived farther north (Huntley & Birks 1983; Bennett *et al.* 1991; Taberlet *et al.* 1998).

From this perspective, only south of Europe provided topography and climatic conditions suitable for long term survival of many tree species through the climatic cycles (Hewitt 1999; Tzedakis *et al.* 2002). Even this model is theoretically sound and correct in many aspects, many discrepancies and evidences against this theory have

started to emerge soon and the nature and location of the LPG refugia for European tree species have been fascinating topic so far.

- **Concept of cryptic refugia**

C**oncept of "cryptic refugia"** is the theory that southern refugia for animal and plant taxa were supplemented by cryptic refugia in northern parts (Stewart & Lister 2001). Study by Stewart & Lister (2001) have started up to now lasting research about the nature and distribution of cryptic northern refugia and their role in shaping today´s biota (e.g. Petit *et al.* 2003; Willis & Van Andel 2004; Birks & Willis 2008; Binney *et al.* 2009; Hofreiter & Stewart 2009; Tzedakis *et al.* 2013; De Lafontaine *et al.* 2013, 2014). If species had northerly refugial populations, they may have achieved postglacil expansion from local, isolated populations, rather than via long-distance dispersal. Despite the fact that many direct as well as indirect evidences have pointed to the existence of northern refugia for wide variety of species, their importance for glacial survival and postglacial migration is not still completely clarified.

As early as the end of 19th century, British biologist Clements Reid tried to calculate how long it would take oak trees to colonize Britain once the glaciers left (Reid 1899) and recognized the problem today known as **"Reid´s paradox of rapid plant migration"**. Simply put, excessively high and not adequate postglacial migration rates suggested by fossil pollen data are not consistent with predictions based on life history and dispersal data for many tree species pointed to postglacial spread from previously undetected northern refugia (Clark *et al.* 1998; McLachtan *et al.* 2005).

Another important fact which highly influenced the survival of species during the LPG in the close proximity of glacier was the extent and continuity of the ice sheet. It was found that **ice sheet was thinner, less continuous and disintegrated earlier than conventionally mentioned** and that tree species may have colonized and grown on thin debris accumulations along the retreating ice immediately after warming (Paus *et al.* 2006). Moreover, there is an evidence **of local ice-free conditions in northern Scandinavia during the LPG** (Moller *et al.* 1992). Together with the fact that climatic conditions in the proximity of ice sheet could not be so harsh, **species distribution modelling** has showed that many **boreal tree species** can tolerate conditions simulated for the full-glacial environment due to their climatic hardiness and phenotypic plasticity (Nikolov & Helmisaari 1992).

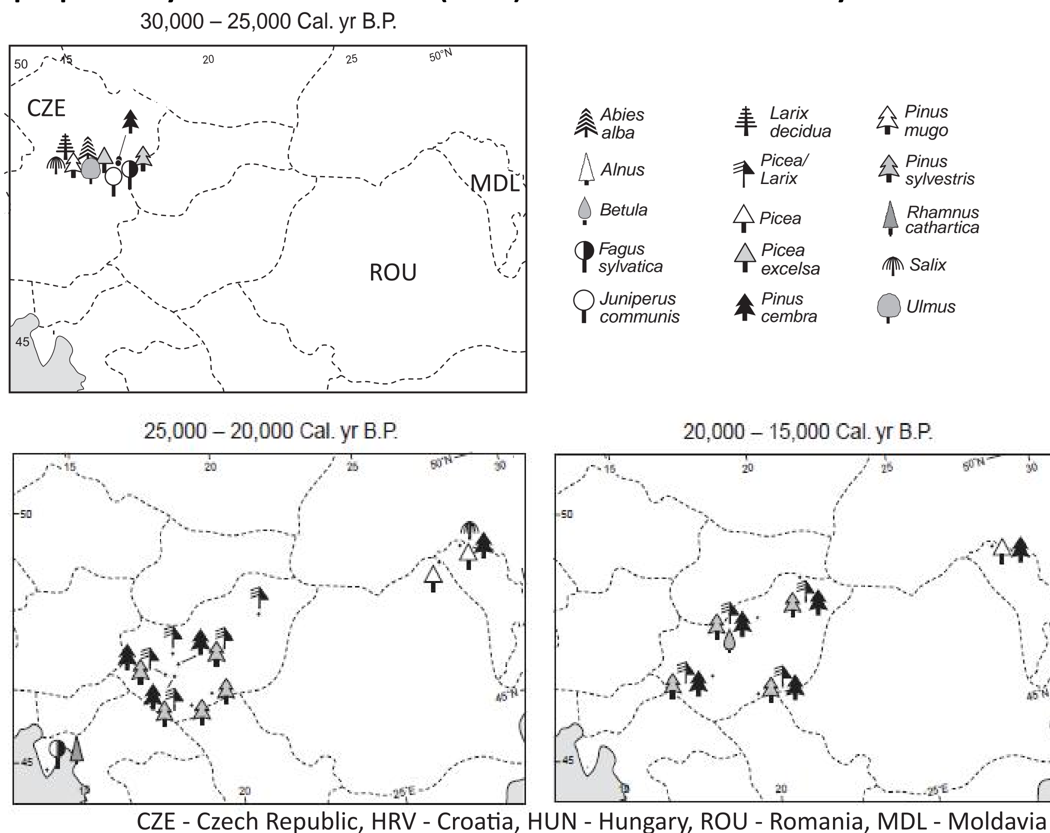
Apart from above mentioned indirected evidences, there are plenty **direct evidences from fossil and genetic analysis**. Willis & Van Andel (2004) reviewed the **macrofossil charcoal evidence** and considered it alongside another fossil, genetic and palaeoclimatic evidence and concluded that **coniferous as well as some broadleaf trees were continuously present throughout the last full-glacial interval in central and eastern Europe** (Fig. 2A). Despite the fact that it is necesary to consider their results with causion because most of the records preceeded the LPG, we still can find evidences suporting the LPG survival of trees in several places.

Firstly, Willis & Van Andel (2004) concluded that **Central Europe during the LPG was covered by a cold-forest steppe** dominated by coniferous trees with deciduous trees in more protected area. Charcoal of *Abies alba*, *Larix decidua*, *Picea excelsa*, *Pinus sylvestris*, *Pinus cembra*, *Pinus mugo*, *Salix* sp., *Ulmus* sp., *Juniperus communis* and *Fagus sylvatica* were found in the occupation layer situated upon loess in Dolní Věstonice and dated throught the period from 32 260 ± 590 to 23 820 ± 380 cal. yr BP. Survival of trees in Central Europe during the LPG was also supported by pollen analysis from Bulhary (Rybníček & Rybníčková 1991) in the Czech Republic. The pollen diagram from Bulhary indicates a coniferous forest containing *Pinus sylvestris*, *P. cembra*, *Picea*, *Larix*, *J. communis* and *Betula* during the LPG period. There is also evidence for the scattered presence of temperate deciduous trees including *Ulmus*, *Acer*, *Corylus*, *Quercus* and *Tilia*. Plant macrofossils obtained from these peat deposits include leaves, seeds and wood of *Betula* cf. *pubescens* and *Salix* sp. (Rybníček & Rybníčková 1991). Coniferous forest probably extended to the east across the Carpathians to the northeastern Slovakia where the site Šafárka shows a diverse pollen assemblage of boreal and thermophilous trees with a direct date on *Larix* cone dated to the LPG period (Jankovská & Pokorný 2008). This view is strongly supported by recent malacostratigraphic investigation undertaken in central Slovakia by Ložek (2006) where snail assemblages consists of a peculiar mixture of cold- and warm-loving elements. It indicates that at the southern foot of the Western Carpathians mountain range a woodland zone persisted during the last full-glacial period (Ložek 2006).

Second area with survival of tree species during the LPG covering southeastern Europe (Willis & Van Andel 2004)**.** Charcoals of *Fagus sylvatica* and

Rhamnus cathartica, found in the cave in Croatia, were dated to 25 260 ± 240 cal yr BP and 25 890 ± 380 cal yr BP, respectivelly.

(A) Cryptic northern refugia in Central and Eastern Europe proposed by Willis & van Andel (2004) based on charcoal analysis



(B) Cryptic northern refugia in SW France proposed by De Lafontaine *et al.* (2013, 2014) based on microsatellites and charcoal macrofossils

Continuous presence of charcoal macrofossils between 15 900 and 14 800 and between 13 000 and 12 700 cal yr BP in The Landes de Gascogne region

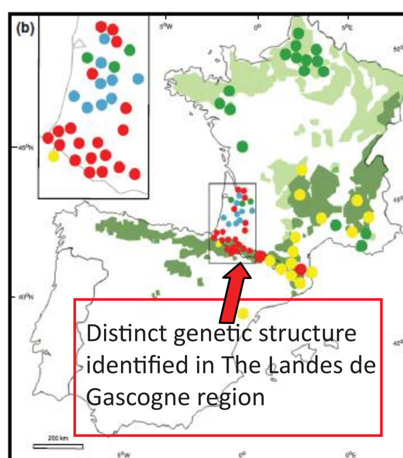
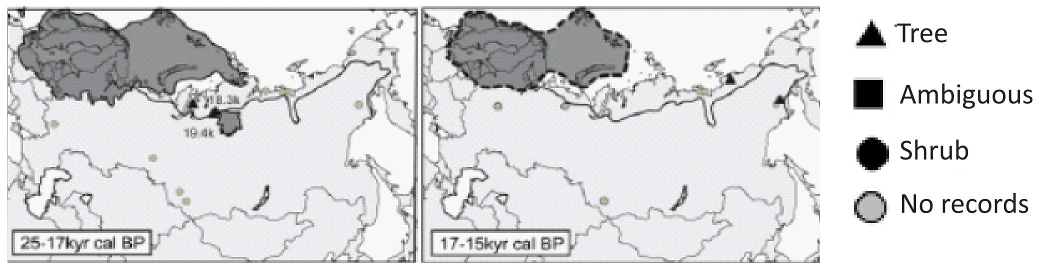


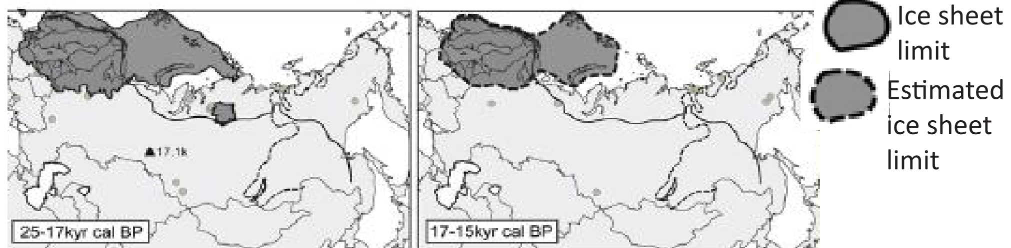
Fig. 2: Key evidences proposing concept of cryptic northern refugia for European tree species. Modified from Willis & Van Andel (2004), Binney *et al.* (2009), Parducci *et al.* (2012) and De Lafontaine *et al.* (2013, 2014).

(C) Cryptic northern refugia proposed by Binney *et al.* (2009) based on macrofosill database

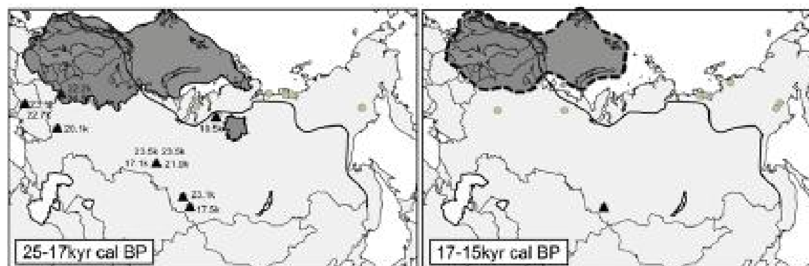
Larix



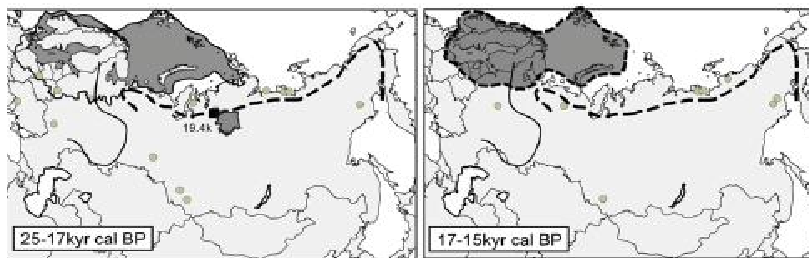
Pinus



Picea



Alnus



Betula

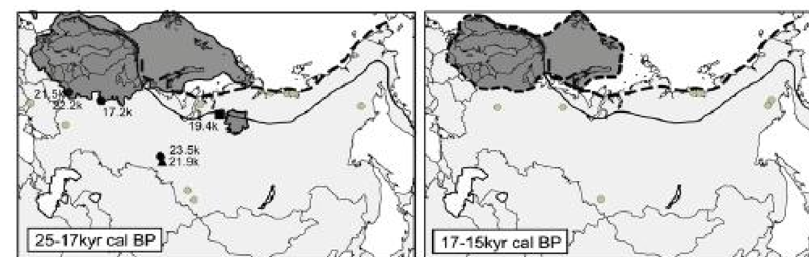


Fig. 2 *(continued)***:** Key evidences proposing concept of cryptic northern refugia for European tree species.

(D) Cryptic northern refugia in Scandinavia proposed by Parducci *et al.* (2012) based on mtDNA and aDNA

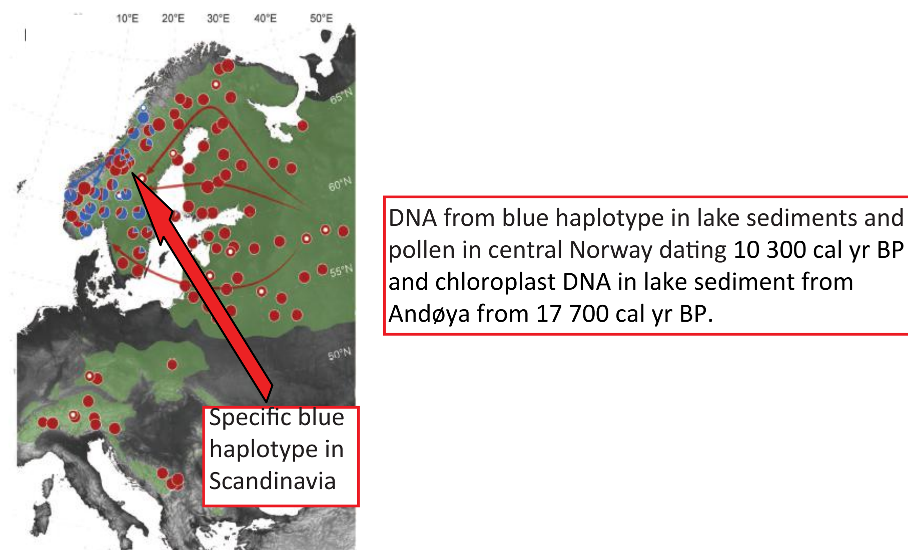


Fig. 2 *(continued)*: Key evidences proposing concept of cryptic northern refugia for European tree species.

In more eastern area, including Hungary, Romania and Moldavia, charcoals of *Picea*, *Larix*, *Pinus sylvestris*, *Pinus cembra*, *Betula* sp. and *Salix* sp. were dated to the period from 25 200 ± 280 cal yr BP to 18 520 ± 240 cal yr BP suggesting that the LPG landscape was probably a mixture of forest-steppe. Aditionally, pollen diagrams from Sarret in Hungary (Willis *et al.* 2000) with the LPG occurrence of *Pinus* and *Picea* and Steregoiu in NW Romania with the LPG occurrence of *Pinus*, *Betula*, *Juniperus* and *Salix* (Björkman *et al.* 2002, 2003) also indicate an open-forested vegetation during the LPG in these areas.

Recently, **the charcoal records have brought direct evidence that *Fagus sylvatica* was present in the Landes de Gascogne in southwestern France** between 15 900 and 14 800 cal yr BP (De Lafontaine *et al.* 2014). Previous genetic study also supported this fact (De Lafontaine *et al.* 2013). Although the dates are after the LPG period, the authors proposed that beech could survive the LPG in this region as well. The reason is that period known as Hendrich stadial-1 was colder and dryer than the LPG in this region (Kageyama *et al.* 2005). Given that Hendrich stadial-1 immediately followed the LPG, it would be surprising if beech was absent during the LPG and recolonized the area during harsher climatic conditions of Hendrich stadial-1 (Fig. 2B). In addition, there are also **molecular evidences about glacial refugia situated in Central Europe, Hungarian plains, Moldavia and Romania for *Pinus sylvestris*,**

Picea abies*, *Frangula alnus* and *Fagus sylvatica (Hampe *et al.* 2003; Cheddadi *et al.* 2006; Tollefsrud *et al.* 2008; De Lafontaine *et al.* 2013; see pages 16 – 18 for more information)

Valuable information about refugia for tree species in northern Europe was brought by Binney *et al.* (2009) who used a database of late-Quaternary plant macrofossil records for northern Eurasia (Fig. 2C). They concluded that **northern refugia most likely existed for some tree species, including *Picea*, *Larix*, *Pinus*, *Alnus* and *Betula*** and that some of them were situated in the close proximity of ice sheet. Majority of locations were situated in river valleys suggesting that trees persisted in cryptic refugia where growing conditions were locally more sheltered and/or moister than typical. It is in concordance with the results of Kullman (2008) who found **megafossil records of *Betula pubescens*, *Pinus sylvestris* and *Picea abies* in northern Scandinavia existed on early ice free mountain peaks (nunataks) during the Lateglacial**. He suggested that these tree species endured the glacial period at sites much closer to the receding Scandinavian ice front than previously postulated from pollen data (Huntley & Birks 1983). From palynological data, there is the long and well-dated sequence from Galich Lake in Russia, with abundant *Picea* pollen during the entire LPG (Velichko *et al.* 2001). Recently, molecular study based on mtDNA of current trees and ancient DNA from lake sediment of Parducci *et al.* (2012) proposed that ***Picea abies* survived in the ice-free refugia of Scandinavia during the last glaciation** (Fig. 2D for details).

- **Northern refugia are rather mystic than cryptic**

Surprisingly in the time of growing agreement about existence of northern refugia both for temperate and boreal tree species, the idea that for **temperate species these refugia are rather mystic than cryptic** have emerged (Tzedakis *et al.* 2013; Fig. 3). Tzedakis *et al.* (2013) critically evaluated evidence comes from plant macrofossils, pollen records, genetic data and potential glacial tree distribution. What emerges is the absence of temperate trees north of 45°N and a west-east asymmetry in boreal tree distribution, with treeless Western Europe north of 46°N, while restricted boreal populations persisted in Eastern Europe up to 49°N, and higher latitudes east of the Fennoscandian ice-sheet (Tzedakis *et al.* 2013).

Concerning the study of Willis & Van Andel (2004) study of Tzedakis *et al.* (2013), apart from the fact that most of charcoal records precede the LPG, mentioned

that the majority of the charcoal pieces have not been dated directly, but by association with dates of the cultural layers within with they have been found. The problem is that charcoals could be remobilized from post-depositional reworking or intrusive movement during archaeological excavation or there is also possibility that assemblages contained material of dispare ages. After reexamination considering only LPG direct dates on identified charcoal, only sites in Moldavia showed the continued presence of trees through-out the LPG with 17 dates on *Picea* and *Pinus cembra* charcoal. Also the results of De Lafontaine *et al.* (2013, 2014), who proposed existence of microrefugia in the western France based on charcoals and molecular analysis, was under critism (Huntley 2014). Huntley (2014) argued that Heindrich stadial-1 could not be the coldest period in the region and that the clustering of the populations differed depending on methods applied to molecular data.

A recurring theme in recent literature is that pollen analysis is not well suited to detecting glacial tree refugia because of decreased pollen productivity and methodological difficulties in interpreting low pollen percentages (Willis & Van Andel 2004, Birks & Willis 2008). Tzedakis *et al.* (2013) mentioned that the pollen profile from the site Bulhary (Rybníčková & Rybníček 1991) was dated only near the top of the section and it shows that the record predates the LPG. The similar problem is with the site Šafárka (Jankovská & Pokorný 2008) where the only reliable indicator for LPG tree presence is a direct date on *Larix* cone.

Regarding genetic data, their interpretation is usually equivocal according to Tzedakis *et al.* (2013) because several demographic factors (such as type of colonization or admixture among colonizing lineages) may complicate the interpretation of genetic signals and different genetic markers may have different information content. They suggested that only genetic variation in *Picea abies* and *Frangula alnus* pointed to the existence of northern refugia. In the case of *Picea abies* refugia were situated in the Carpathians and Russia (Heuertz *et al.* 2006; Tollefsrud *et al.* 2008, 2009) and southeastern refugia were proposed for *Frangula alnus* (Hampe *et al.* 2003). Existence of refugia for *Picea abies* in Scandinavia proposed by Parducci *et al.* (2012) based on locally endemic derived mtDNA haplotype found exclusively in Scandinavia is unclear and its origin could be pre-LPG, post-LPG mutation or post-LPG immigration. Moreover, this study has been questioned on the grounds of contamination or reworking of ancient DNA and the need for more discriminating modern genetic data to reject alternative explanations (Birks *et al.* 2012,Vorren *et al.*

2013). Therefore, its results should be further tested with additional genetic data. Tzedakis *et al.* (2013) concluded that the rest of phylogeographical studies has not

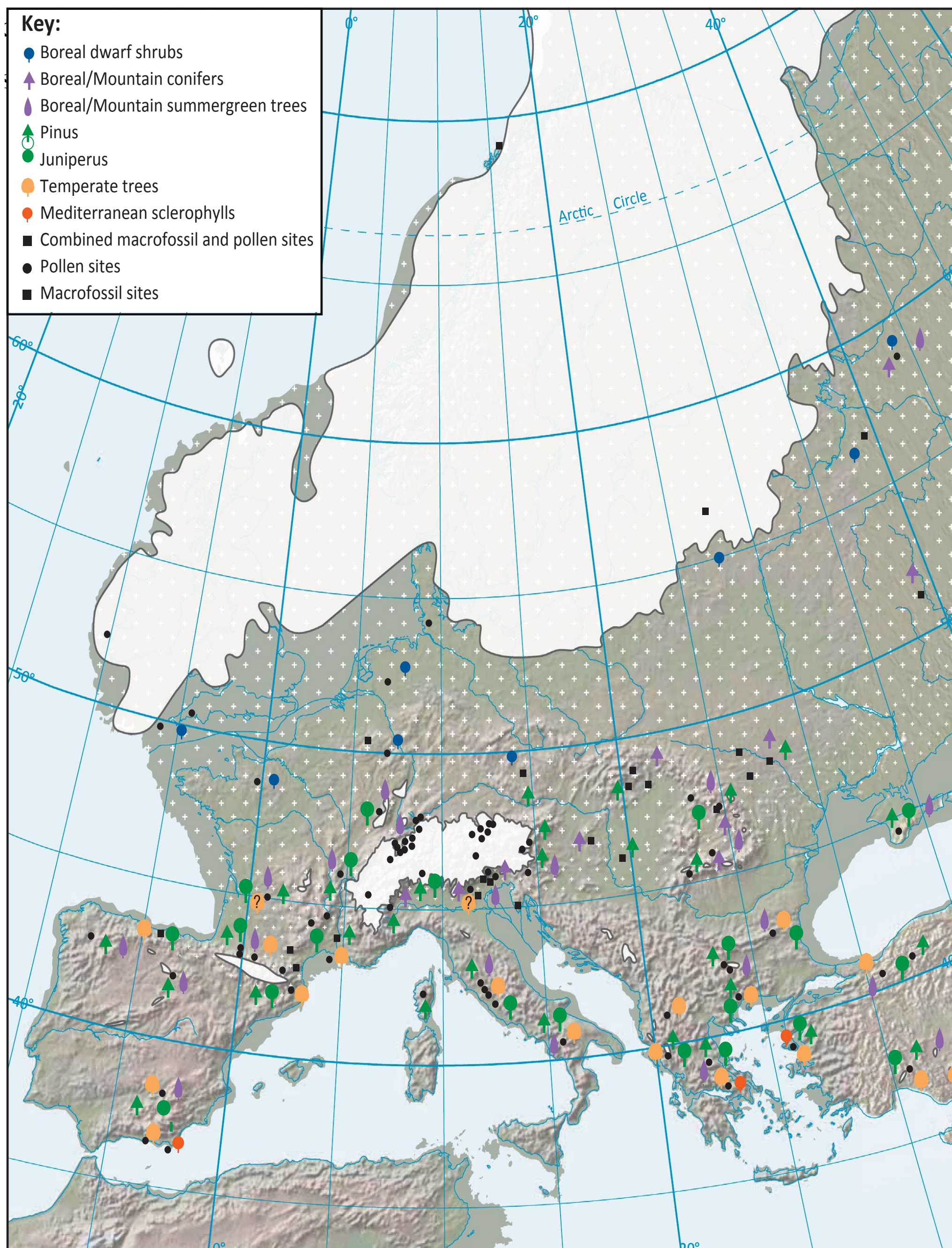


Fig. 3: Distribution of macrofossil and pollen sites and inferred tree presence supporting that northern refugia are rather mystic than cryptic. Question marks over vegetation symbols denote uncertainty. Adapted from Tzedakis (2013).

What do molecular data tell us about postglacial history of European woody species?

Theories about tree survival during LPG period in Europe mentioned in previous chapter were based mainly on fossil evidences. Nowadays, we have plenty of information from molecular data which gave rise to new branch of biogeography – **phylogeography** – that has recently been developed (Avise *et al.* 1987). It has brought new insights into postglacial history and detected refugia and migration routes by analysing geographical distribution of genetic lineages across the range of a species.

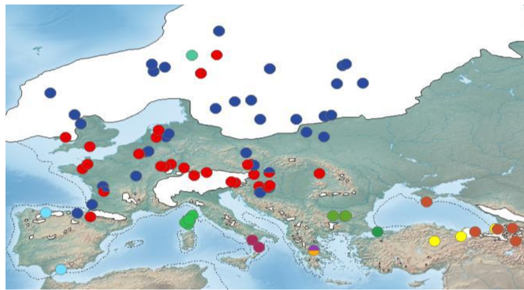
- **Key results from phylogeographic studies**

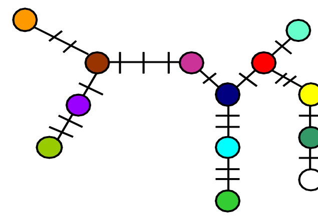
Phylogeographic studies have shown that **many private highly divergent haplotypes of European trees are usually harboured in southern Mediterranean populations,** e.g. *Alnus glutinosa*, *Carpinus betulus* and *Corylus avellana* (King & Ferris 1998; Palmé & Vendramin 2002; Grivet & Petit 2003). It pointed to long term survival in these areas which in most cases did not contribute to the colonization of the northern parts of Europe and remained trapped there. As exceptions, Mediterranean haplotypes of *Quercus* sp. and *Fraxinus excelsior* expanded into central and northern Europe (Petit *et al.* 2002; Heuertz *et al.* 2004a). Moreover, refugia of European trees were also found in Anatolia and North Africa (e.g. *Alnus glutinosa)*, that also did not contribute to the postglacial colonization of Europe (King & Ferris 1998; Lepais *et al.* 2013).

After ice sheet retreat, the most common **migration route proposed for many temperate trees** (Fig. 4), e.g. *Alnus glutinosa*, *Carpinus betulus*, *Frangula alnus*, *Malus sylvestris* and *Populus nigra*, **was the postglacial colonization from the Balkan Peninsula** (King & Ferris 1998; Grivet & Petit 2003; Hampe *et al.* 2003; Cottrell *et al.* 2005; Cornille *et al.* 2013). **Other effective refugia were situated to the Dinaric Alps or foothills of the Alps** (e.g. *Fagus sylvatica*; Magri *et al.* 2006). In few cases, **multiple expansions** from all three southern peninsulas was proposed for *Fraxinus excelsior* (Heuertz *et al.* 2004a) and *Quercus* sp. (Petit *et al.* 2002).

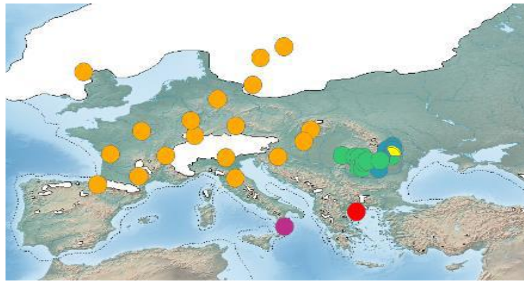
From genetic data there is nearly **no evidence pointing to existence of northern refugia for temperate trees.** Few exceptions were proposed, **(i)** existence of microrefugia in southwestern France for *Fagus sylvatica* based on charcoal and microsatellite analysis (De Lafontaine *et al.* 2013, 2014, Fig. 2B) and **(ii)** Hungarian refugium for *Frangula alnus*, the area with presence of four different cpDNA haplotypes (Hampe *et al.* 2003).

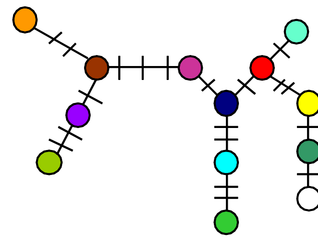
A. *Alnus glutinosa*



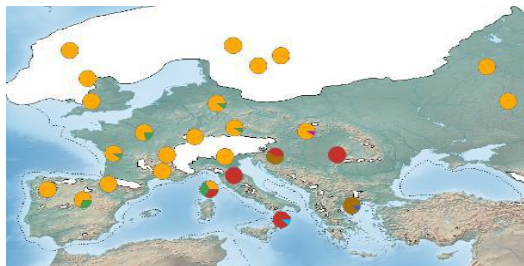


B. *Carpinus betulus*





C. *Corylus avellana*



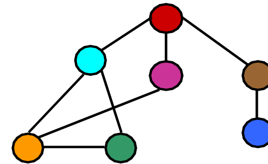
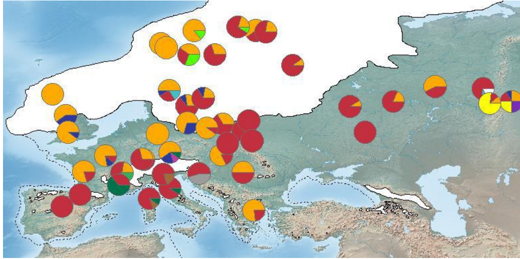


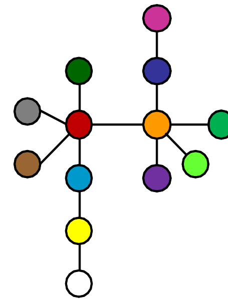
Fig. 4: Phylogeographical patterns of temperate trees **A.** *Alnus glutinosa*, **B.** *Carpinus betulus* and **C.** *Corylus avellana* with decreasing genetic diversity from south to north. Haplotypes networks next to the maps are based on PCR-RFLP cpDNA for *A. glutinosa* and *C. betulus* and cpDNA SSRs for *C. avellana* (the lines in the network indicate mutational steps). Modified from King & Ferris (1998); Palmé & Vendramin (2002) and Grivet & Petit (2003).

On the other hand, **refugia in northern areas such as northern Carpathians, Russian plains and central Europe** were commonly proposed **for boreal tree species** such as *Pinus sylvestris*, *Picea abies*, *Salix caprea* and *Betula pendula* (Palmé *et al.* 2003a, b; Cheddadi *et al.* 2006; Maliouchenko *et al.* 2007; Tollefsrud *et al.* 2008). Populations of *Salix caprea* and *Betula pendula* showed high level of genetic diversity and the lack or weak genetic structure north of the Alp (Fig. 5) suggesting that these tree species **could survive in northern areas** fragmented into several isolated populations. It should also be stressed evidence concerning **refugia for *Picea abies* close the ice sheet in Scandinavia** (Fig. 2D), when proximity to the Atlantic Ocean may have ameliorated the harsh conditions (Parducci *et al.* 2012). Parducci *et al.* (2012) showed the presence of a rare mitochondrial DNA haplotype of Norway spruce that appears unique to Scandinavia. They further extracted DNA from this haplotype

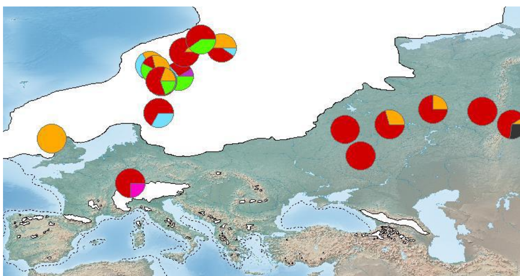
in lake sediments and pollen in central Norway dating ~ 10 300 years BP and chloroplast DNA in lake sediments adjacent to the ice-free Andøya refugium in northwestern Norway as early as ~ 17 700 years BP (Parducci *et al.* 2012).

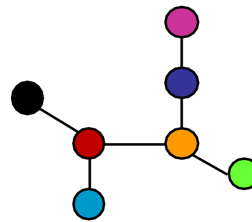
A. *Betula pendula*



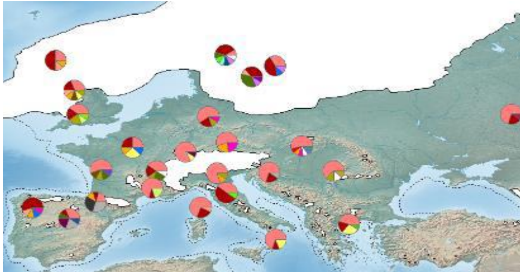


B. *Betula pubescens*





C. *Salix caprea*



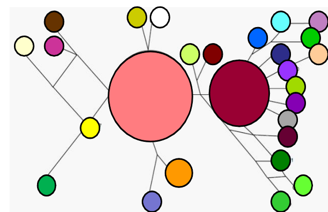


Fig. 5: Phylogeographical patterns of boreal trees **A.** *Betula pendula*, **B.** *Betula pubescens* and **C.** *Salix caprea* with high genetic diversity in northern parts. Haplotypes networks next to the maps are based on PCR-RFLP cpDNA (for *Salix caprea* the lengths of the branches are proportional to the number of changes between haplotypes and the sizes of circles reflect the frequency of the haplotypes). Modified from Palmé *et al.* (2003a; b) and Maliouchenko *et al.* (2007).

- **Problems of phylogeographic studies**

It is difficult to find general trends about postglacial colonization of European woody species because each species need **specific environmental conditions** and has **different dispersal ability**. Therefore their survival under unfavourable conditions and subsequent migrations could have individualistic character (Taberlet *et al.* 1998; Hewitt 2000).

Another problem is that **complexity of knowledge about individual woody species differs significantly**. Only few species (i.e. *Fagus sylvatica* and *Picea abies*) have been well studied throughout their entire distribution range with sufficient sampling intensity by combining approach of fossil data analysis and several genetic markers (Magri *et al.* 2006; Tollefsrud *et al.* 2009).

Moreover, majority of studies focused on post glacial migration in Europe have been based on **one molecular marker**, frequently **cpDNA or isozymes**. The problem is that these markers are **insufficiently variable** in many cases and we receive only coarse picture of population genetic structure which do not allow us to determine specific positions of refugia, migration routes and underlying processes.

- **Future directions of phylogeographic studies**

Despite previous phylogeographic studies have brought valuable knowledge about postglacial colonization of European tree species, they still have suffered from above mentioned limits. Even the most comprehensive studies focused on phylogeography of *Picea abies* (Latałowa & van der Knaap 2006; Tollefsrud *et al.* 2008, 2009; Parducci *et al.* 2012) have not brought complete picture about its history and there are still some mysteries which need to be proofed in future studies. One of them is existence of northern refugia in Scandinavia proposed by Parducci *et al.* (2012), which was questioned and it was suggested that more sampling effort and genetic analysis is needed to confirmed this fantastic discovery (Birks *et al.* 2012).

It was shown that studies which **combined different methods**, ideally review the fossil data and analyse sufficient number of samples across the whole distribution range by more molecular markers, can shed a new light on geographical position of glacial refugia and postglacial migration routes (Cheddadi *et al.* 2006; Magri *et al.* 2006). Moreover, **using more variable molecular markers** such as microsatellites can provide us more detailed picture of postglacial colonization (Heuertz *et al.* 2004b; Tollefsrud *et al.* 2009; Cornille *et al.* 2013). They provide finer scale resolution of historical dynamics, distinguish effective refugia from not effective ones and estimation of population parameters such as population size, growth, subdivision, admixture, patterns of gene flow and the timing of divergence (Knowles 2009; Hickerson *et al.* 2010).

2. BLACK ALDER (*ALNUS GLUTINOSA*) AND GREY ALDER (*ALNUS INCANA*): suitable species for new phylogeographic study

Why Alnus glutinosa and A. incana are suitable species for new study?

There are several reasons, why black alder (*Alnus glutinosa*) and grey alder (*Alnus incana*) are suitable species for new phylogeographic study:

(1) Studying these two species enable us to **compare postglacial histories of temperate tree** (*A. glutinosa*) **and boreal tree** (*A. incana*) which are expected to differ significantly. Refugia of more cold-tolerant *A. incana* could be present in northerly located areas than refugia of temperate *A. glutinosa*. Consequently, postglacial migration routes of *A. glutinosa* could be on longer distances while A. *incana* could expand also locally from northern refugia.

(2) These two species **cannot always be successfully distinguished based on pollen and macrofossil remains** (Huntley & Birks 1983). Thus, genetic studies may provide a more accurate picture about historical patterns of migration because it may help to reject some hypotheses postulated only on the basis of genus specific pollen and macrofossil data (King & Ferris 1998; Palmé *et al.* 2003b; Cheddadi *et al.* 2006). Likewise, genetic data may help to distinguish the effective refugia, which have contributed to the postglacial colonization from the non-effective ones, which were not important for postglacial migrations (Magri *et al.* 2006; Tollefsrud *et al.* 2008; Liepelt *et al.* 2009).

(3) Distributions of both alders have not been largely expanded by human reforestation; thus there are **plenty of non-planted stands** especially on waterlogged sites. These sites can serve as a good source for genetic analysis due to low human impact and therefore there is a low chance of dealing with plant material introduced by man.

(4) *Alnus* species are **keystones of alluvial and wetland habitats** (Douda *et al.* 2009; Douda 2010) distributed through the European forest zones from the northern treeline to the Mediterranean, therefore understanding the past history and postglacial migration pattern of *Alnus* populations may help understand the resistance and

resilience of wetland forest habitats to global climate change (Erwin 2008; Garssen *et al.* 2015).

(5) There are **substantial gaps in reconstructions of postglacial histories** of *A. glutinosa* and *A. incana* in Europe. It partly results from difficulties with species determination only from fossil records (Huntley & Birks 1983) and partly from the lack of molecular studies. Up until now, only one large-scale study dealing with postglacial history of *A. glutinosa* has been performed (King & Ferris 1998). Unfortunately, this study has not provided detailed information about migration routes, firstly because insufficiently variable cpDNA not suitable for precise detection of postglacial history was used and secondly, interpretations of results followed rather Huntley & Birks (1983) pollen maps than detected genetic patterns.

Previous knowledge about A. glutinosa and A. incana

- **Study species**

Two common tree species of *Alnus* grow natively in Europe (Jalas & Suominen 1976). **Black alder** (*Alnus glutinosa*) is considered a **temperate tree**. It commonly occurs in the lowlands and mountains across Europe except Scandinavia, where it is associated with a coastal oceanic climate in southern areas (Tallantire 1974) (Fig. 6).

Fig. 6: The current distribution of *A. glutinosa*. Modified from EUFORGEN 2009 (www.euforgen.org).

The cold-climate limitation also likely affects its distribution in high-elevation mountainous areas, where black alder populations are often absent. Scarce distributions are found in the Mediterranean region and in the arid Great Hungarian

plains, the Ukraine and the Russian steppe zone. Outside Europe, the distribution extends as far as western Siberia and the mountains of Turkey, Iran and North Africa (McVean 1953; Jalas & Suominen 1976). In Corsica and southern Italy, *A. glutinosa* grows sympatrically with *A. cordata* (Loisel.) Duby.

Grey alder (*Alnus incana*) is considered a **boreal and mountain tree**. Similar to Norway spruce (*Picea abies*), the range of *A. incana* is divided into a northern and a southern area, which meet in the Polish lowlands (Fig. 7). In northern Europe,

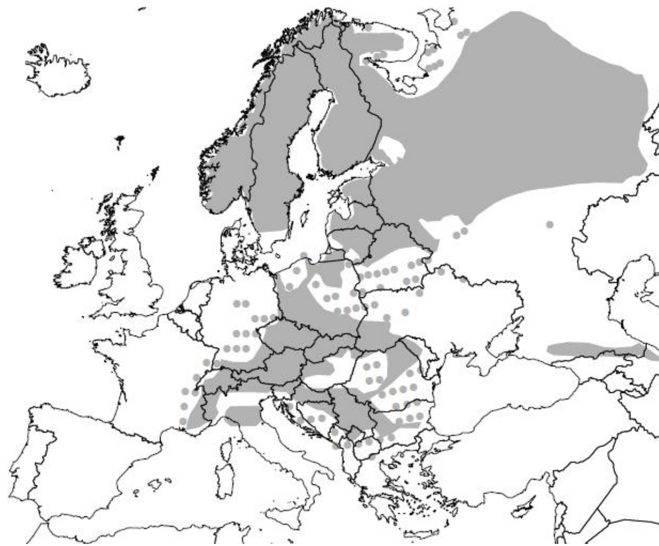


Fig. 7: The current distribution of *A. incana*. Modified from Schwabe (1985) and Hultén & Fries (1986).

A. incana continuously covers the east Baltic region and all of Scandinavia with a northern margin at latitudes greater than 70°N (Tallantire 1974; Jalas & Suominen 1976). In northern Scandinavia, the nominal subspecies grows sympatrically with *A. incana* subsp. *kolaensis* (Orlova) Á.Löve & D.Löve (Jalas & Suominen 1976). The distribution of grey alder continues eastwards across European Russia to western Siberia, which contrasts with its patchy mountain occurrence in the southern part of the range linked to the Alps, the northern Apennines, the Hercynian Mountains, the Carpathians, the Bulgarian Mountains, the Dinaric Alps, the Caucasus and Turkey (Jalas & Suominen 1976).

Alnus glutinosa and *A. incana* dominate in floodplain and swamp forests. These species are indifferent to soil nutrient conditions, except for extremely poor peat bogs. Seeds are dispersed effectively by water, while wind dispersal is commonly limited to the vicinity of the parent tree (McVean 1953). Under unfavourable environmental conditions, such as in cold climates, *A. incana* is able to survive and reproduce by

clonal growth (Kullman 1992). Compared with the relatively short-lived *A. incana* (c. 20–50 years), *A. glutinosa* is a long-lived tree (c. 100–120 years), although the age of reproduction is similar for the two species (i.e., 10–20 years) (McVean 1953; Tallantire 1974).

- **Upper Pleistocene and Holocene *Alnus* history**

(1) Paleoecological data

Pleistocene pollen and macrofossil data indicate **repeated population increases and decreases** of *Alnus* in Europe, reflecting climate oscillations between glacial periods and interglacials, particularly noticeable in the Middle and Upper Pleistocene (Wijmstra 1969; West 1980; Tzedakis 1994). The majority of Upper Pleistocene pollen profiles support a common presence of *Alnus* in the Eemian interglacial (Marine Isotope Stage, MIS; 125–115 kyr BP) throughout Europe, i.e. in Denmark, Greece, Italy and France (Andersen 1965; Wijmstra 1969; De Beaulieu & Reille 1984, 1992; Tzedakis 1994; Follieri *et al.* 1998; Müller 2000; Reille *et al.* 2000) and its disappearance after the start of the last glacial period.

Huntley & Birks (1983) assumed that the main **source refugia for the *Alnus* expansion after the LGM lay in the eastern Alps, the Carpathians and the Ukrainian lowlands**. Other LGM refugia were located in Corsica, western France, northern Spain and northwestern Russia. The authors supposed that the Holocene migration of *Alnus* likely began somewhere in eastern Europe and continued by the northward expansion of *A. glutinosa* and *A. incana* to the Baltic and Scandinavia and by the westward expansion of *A. glutinosa* along the southern shore of the North Sea as far as the British Isles (Huntley & Birks 1983).

(2) Molecular data

King & Ferris (1998), who accepted Huntley and Birks' (1983) east-west migration pattern in their genetic study, revealed **13 cpDNA haplotypes of *A. glutinosa*, mainly associated with southern European peninsulas** (Fig. 8). They suggested that **two of these haplotypes colonized northern and central Europe from LPG refugia located in the Carpathians**, specifically the northward expansions to the Baltic and Scandinavia and the westward expansions of along the southern shore of the North Sea as far as the British Isles as was previously suggested by fossil data (Tallantire 1974; Huntley & Birks 1983). It was proposed that the delayed migration of *A. glutinosa* to

the Western Europe probably resulted from distant glacial refugia (Huntley & Birks 1983).

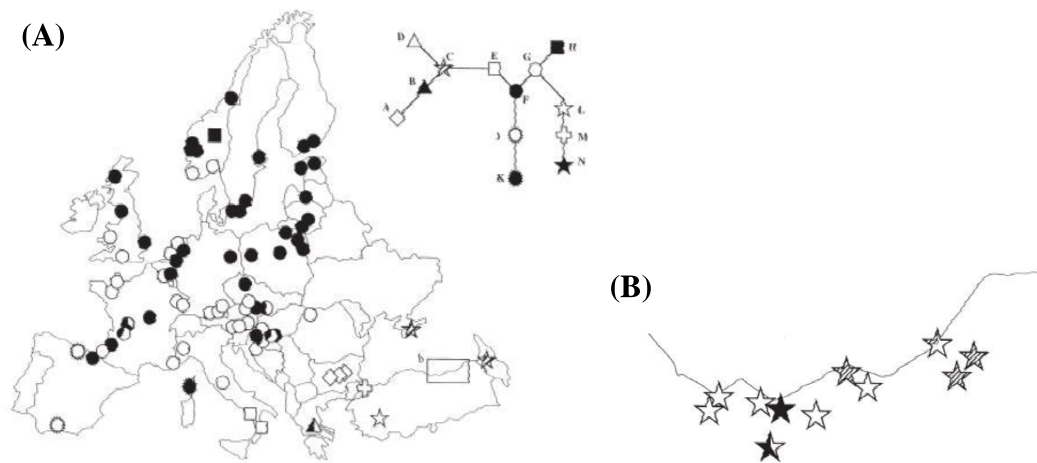


Fig. 8: Geographic distribution of 13 cpDNA haplotypes indentified in *Alnus glutinosa* (King & Ferris 1998). (A) European distribution. A simplified version of the minimum-spanning network tree is included. (B) Detail of northeast Turkey. For polymorphic populations shading is proportional to haplotype proportion. Adapted from King & Ferris (1998).

Lepais *et al.* (2013) revealed demographic history of *A. glutinosa* at the rear-edge of the species' distribution in North Africa (Fig. 9) where alder populations inhabit barely accessible landscape in small pockets of locally wet habitat isolated from the main distribution area that stands across continental Europe. Analysis of microsatellite markers shows that **Moroccan populations were estimated to be ancient relicts** characterised by a high genetic diversity, a strong genetic distinctiveness and tetraploid origin. In contrary, **Tunisian and Algerian populations were diploids more closely related to European populations** (Lepais *et al.* 2013).

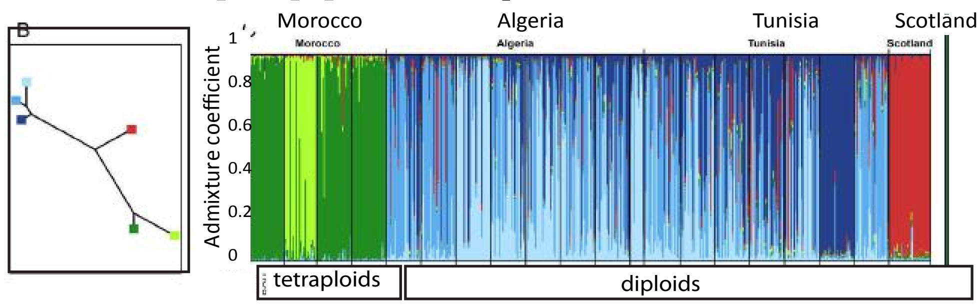


Fig. 9: Regional genetic structure of *A. glutinosa* in North Africa and Scotland as inferred by Structure (Lepais *et al.* 2013). Neighbour joining tree computed using the genetic distance between clusters and histogram of individual assignment to clusters where each individual is represented by a thin vertical bar partitioned into several coloured segments proportionally to its membership of a given cluster. Adapted from Lepais *et al.* 2013.

Aims, questions and hypothesis of dissertation thesis

My dissertation thesis aims to combine information from fossil records and genetic analysis to improve knowledge on locations of glacial refugia and postglacial colonization routes of *Alnus glutinosa* and *Alnus incana*. To find more details, our molecular study will be based on extensive population sampling, will test hypotheses coming from comprehensive fossil analysis and will combine two genetic markers, i.e. chloroplast DNA and microsatellites. Maternally inherited cpDNA will give us coarse picture of individual haplotype distribution all over Europe and recombining genetic marker, i.e. microsatellites, will provide us with precise information concerning population-genetic structure, gene flow among populations, the degree of inbreeding and isolation by distance. Based on the results, important questions concerning the postglacial colonization and changes in genetic pattern of populations during the last glacial cycle in Europe will be able to be answered. **The thesis will shed light up on**: **(i)** refugia of *A. glutinosa* and *A. incana* important for postglacial colonization of Europe, **(ii)** possible existence of northerly located refugia in the proximity of glacier, **(iii)** postglacial migration routes of both *Alnus* species and **(iv)** character of colonization, i.e. form of dispersal, gene flow among populations and loss of genetic variation during migration.

Using published data (Tallantire 1974; Huntley & Birks 1983; Chambers & Elliott 1989; Hewitt 1996; King & Ferris 1998; Cruzan & Templeton 2000; Stewart & Lister 2001; Willis & Van Andel 2004) I formulated a set of specific questions and hypotheses that can be directly tested by empirical data.

Q1: Did *A. glutinosa* and *A. incana* differ in geographic position of their source refugia for postglacial colonization? Did they colonize Europe from single source area or from multiple refugia? More cold-tolerant *A. incana* could be present in the central and eastern European mountain areas and north-west Russia lowlands, whereas southern European peninsulas and lowlands with large rivers in south-west Russia and Ukraine could be the main effective refugia for *A. glutinosa* (**H1**).

Q2: Did refugia of alders occur in the proximity of glaciers during the LPG; if so, which role did they play during colonization of Europe? Several northern areas such as southeast Great Britain, central Germany, Czech Republic and north-west Russia could be refugia of alders during the LPG (**H2**).

Q3: Do *A. glutinosa* and *A. incana* differ in their genetic population subdivision? We hypothesize that *A. glutinosa* as the more sensitive species to unfavourable climate will be characterized by higher genetic differentiation than *A. incana* due to its retirement to more isolated southern refugia during the last glaciation (**H3**).

Q4: What role did the long-distance dispersal events play during colonization of alder species? The coarse patch patterns in genetic subdivision will be apparent, if the rare long-distance dispersal events were important during the colonization process (**H4**).

Q5: Does genetic variation change with distance to source area of Holocene range expansion? A decrease of genetic variation along with an increasing distance from glacial refugia is expected due to the process of genetic drift following the expansion (**H5**)

References

Andersen S (1965) Interglacialer og interstadialer i Danmarks Kvartaer. *Meddelelser fra Dansk Geologisk Forening*, **15**, 496–506.

Avise JC, Arnold J, Ball RM, Bermingham E, LAmb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

De Beaulieu J-L, Reille M (1984) A long Upper Pleistocene pollen record from Les Echets, near Lyon, France. *Boreas*, **13**, 111–132.

De Beaulieu J-L, Reille M (1992) The last climatic cycle at La Grande Pile (Vosges, France) a new pollen profile. *Quaternary Science Reviews*, **11**, 431–438.

Bennett KD, Tzedakis PC, Willis K (1991) Quaternary refugia of north European trees. *Journal of Biogeography*, **18**, 103–115.

Binney HA, Willlis KJ, Edwards ME, Bhagwat SA, Anderson PM, Andreev AA, Blaauw M, Damblon F, Haesaerts P, Kienast F, Kremenetski KV, Krivonogov SK, Lozhkin AV, MacDonald GM, Novenko EY, Oksanen P, Sapelko TV, Väliranta M, Vazhenina L (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews*, **28**, 2445–2464.

Birks HH, Giesecke T, Hewitt GM, Tzedakis PC, Bakke J, Birks HJB (2012) Comment on "Glacial survival of boreal trees in northern Scandinavia". *Science*, **338**, 742.

Birks HJB, Willis KJ (2008) Alpines, trees, and refugia in Europe. *Plant Ecology & Diversity*, **1**, 147–160.

Björkman L, Feurdean A, Cinthio K, Wohlfart B, Possnert G (2002) Late Glacial and early Holocene woodland development in the Gutaiului Mountains, NW Romania. *Quaternary Science Reviews*, **21**, 1039–1059.

Björkman L, Feurdean A, Wohlfart B (2003) Late-Glacial and Holocene forest dynamics at Steregoiu in the Gutaiului Mountains, Northwest Romania. *Review of Palaeobotany and Palynology*, **124**, 79–111.

Clark JS, Fastie C, Hurtt G, Jackson ST, Johnson C, King GA, Lewis M, Lynch J, Pacala S, Prentice C, Schupp EW, Webb III T, Wyckoff P (1998) Reid's paradox of rapid plant migration. *BioScience*, **48**, 13–24.

Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432–438.

Cornille A, Giraud T, Bellard C, Tellier A, Le Cam B, Smulders MJM, Kleinschmit J, Roldan-Ruiz I, Gladieux P (2013) Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Molecular Ecology*, **22**, 2249–2263.

Cottrell JE, Krystufek V, Tabbener HE, Milner AD, Connolly T, Sing L, Fluch S, Burg K, Lefevre F, Achard P, Bordács S, Gebhardt K, Vornam B, Smulders MJM, Vanden Broeck AH, Van SLycken J, Storme V, Boerjan W, Castiglione S, Fossati T, Alba N, Agúndez D, Maestro C, Notivol E, Bovenschen J (2005) Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA. *Forest Ecology and Management*, **206**, 71–90.

Cowling SA, Sykes MT (1999) Physiological significance of low atmospheric CO_2 for plant–climate interactions. *Quaternary Research*, **52**, 237–242.

Cruzan M, Templeton A (2000) Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution*, **15**, 491–496.

Douda J (2010) The role of landscape configuration in plant composition of floodplain forests across different physiographic areas. *Journal of Vegetation Science*, **21**, 1110–1124.

Douda J, Čejková A, Douda K, Kochánková J (2009) Development of alder carr after the abandonment of wet grasslands during the last 70 years. *Annals of Forest Science*, **66**, 1–13.

Erwin KL (2008) Wetlands and global climate change: the role of wetland restoration in a changing world. *Wetlands Ecology and Management*, **17**, 71–84.

Follieri M, Giardini M, Magri D, Sadori L (1998) Palynostratigraphy of the last glacial period in the volcanic region of central Italy. *Quaternary International*, **47**, 3–20.

Frenzel B, Pécsi M, Velichko A (1992) *Atlas of paleoclimates and paleoenvironments of the Northern hemisphere. Late Pleistocene - Holecene* (B Frenzel, M Pécsi, A Velichko, Eds). Geographical Research Institute, Hungarian Academy of Sciences, Budapest.

Garssen A, Baattrup-Pedersen A, Voesenek LACJ, Verhoeven JTA, Soons MB (2015) Riparian plant community responses to increased flooding: a meta-analysis. *Global Change Biology*.

Grivet D, Petit R (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics*, **4**, 47–56.

Hampe A, Arroyo J, Jordano P, Petit RJ (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, **12**, 3415–3426.

Heuertz M, Fineschi S, Anzidei, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG (2004a) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology*, **13**, 3437–3452.

Heuertz M, Hausman J-F, Hardy OJ, Vendramin GG, Frascaria-Lacoste N, Vekemans X (2004b) Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution*, **58**, 976–988.

Heuertz M, De Paoli E, Källman T, Larson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics*, **174**, 2095–2105.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

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

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.

Hofreiter M, Stewart J (2009) Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology*, **19**, 584–594.

Höhn M, Gugerli F, Abran P, Bisztray G, Buonamici A, Cseke K, Hufnagel L, Quintela-Sabarís C, Sebastiani F, Vedramin GG (2009) Variation in the chloroplast DNA of Swiss stone pine (Pinus cembra L.) reflects contrasting post-glacial history of populations from the Carpathians and the Alps. *Journal of Biogeography*, **36**, 1798–1806.

Huntley B. (2014) Extreme temporal interpolation of sparse data is not a sufficient basis to substantiate a claim to have uncovered Pleistocene forest microrefugia. *New Phytologist*, **204**, 447–449.

Huntley B, Birks H (1983) *An atlas of past and present pollen maps for Europe: 0–13000 years ago*. Cambridge University Press., Cambridge.

Chambers F, Elliott L (1989) Spread and expansion of *Alnus* Mill. in the British isles: timing , agencies and possible vectors. *Journal of Biogeography*, **16**, 541–550.

Cheddadi R, Vendramin GG, Litt T, Francois L, Kageyama M, Lorentz S, Laurent J-M, De Beaulieu J-L, Sadori S, Jost A, Lunt D (2006) Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, **15**, 271–282.

Jalas J, Suominen J (1976) *Atlas Florae Europaeae. Distribution of vascular plants in Europe. 3. Salicaceae to Balanophoraceae.* The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinky.

Jankovská V, Pokorný P (2008) Forest vegetation of the last full-glacial period in the Western Carpathians (Slovakia and Czech Republic). *Preslia*, **80**, 307–324.

Kageyama M, Combourieu Nebout N, Sepulchre P, Peyron O, Krinner G, Ramstein G, Cazet J-P (2005) The Last Glacial Maximum and Heinrich Event 1 in terms of climate and vegetation around the Alboran Sea: a preliminary model-data comparison. *Comptes Rendus Geosciences,* **337**, 983–992.

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, **21**, 393–404.

King A, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Knowles LL (2009) Statistical Phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.

Kullman L (1992) The ecological status of gray alder (*Alnus incana* (L) Moench) in the upper sub-alpine birch forest of the Central Scandes. *New Phytologist*, **120**, 445–451.

Kullman L (2008) Early postglacial appearance of tree species in northern Scandinavia: review and perspective. *Quaternary Science Reviews*, **27**, 2467–2472.

De Lafontaine G, Amasifuen Guerra CA, Ducousso A, Petit RJ (2014) Cryptic no more: soil macrofossils uncover Pleistocene forest microrefugia within a periglacial desert. *The New Phytologist*, **204**, 715–729.

De Lafontaine G, Ducousso A, Lefèvre S, Magnanou E, Petit RJ (2013) Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology*, **22**, 4397–4412.

Lambeck K (2004) Sea-level change through the last glacial cycle: geophysical, glaciological and palaeogeographic consequences. *Comptes Rendus Geoscience*, **336**, 677–689.

Latałowa M, van der Knaap WO (2006) Late Quaternary expansion of Norway spruce *Picea abies* (L.) Karst. in Europe according to pollen data. *Quaternary Science Reviews*, **25**, 2780–2805.

Lepais O, Muller SD, Ben Saad-Limam S, Benslama M, Rhazi L, Belouahem-Abed D, Daoud-Bouattour A, Gammar AM, Ghrabi-Gammar Z, Bacles CFE (2013) High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PloS one*, **8**, e75029.

Leroy SAG, Arpe K (2007) Glacial refugia for summer-green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulations. *Journal of Biogeography*, **34**, 2115–2128.

Liepelt S, Cheddadi R, de Beaulieu J-L, Fady B, Gömöry D, Hussendörfer E, Konnert M, Litt T, Longauer R, Terhürne-Berson R, Ziegenhagen B (2009) Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) — A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139–149.

Lomosino M, Riddle B, Brown J, Whittaker R (2006) *Biogeography*. Sinauer Associates, Sunderland, MA.

Ložek V (2006) Last Glacial paleoenvironments of theWestern Carpathians in the light of fossil malacofauna. *Anthropozoikum*, 26, 73–84.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latalowa M, Litt T, Paule L, Route JM, Tantau I, Van de Knaap WO, Petit RJ, De Beaulieu J-L (2006) A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *The New Phytologist*, **171**, 199–221.

Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, 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.

McLachtan J, Clark J, Manos P (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

McVean D (1953) Account of *Alnus glutinosa* (L.) Gaertn. for the Biological Flora of the British Isles. *Journal of Ecology*, **41**, 447–466.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.

Moller J, Danielsen T, Fjalstad A (1992) Late weichselian glacial maximum on Andoya, North Norway. *Boreas*, **21**, 1–13.

Müller U (2000) A Late-Pleistocene pollen sequence from the Jammertal, south-western Germany with particular reference to location and altitude as factors determining Eemian forest composition. *Vegetation History and Archaeobotany*, **9**, 152–131.

Nikolov N, Helmisaari H (1992) Silvics of the circumpolar boreal forest tree species. In: *A systems analysis of the global boreal forest* (eds Shugart H, Leemans R, Bonan G), pp. 13–85. Cambridge University Press, Cambridge, UK.

Palmé AE, Semerikov V, Lascoux M (2003a) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003b) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology*, **12**, 201–212.

Palmé AE, Vendramin GG (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology*, **11**, 1769–1779.

Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y, Edwards ME, Andersen K, Rasmussen M, Boessenkool S, Coissac E, Brochmann Ch, Taberlet P, Houmark-Nielsen M, Larsen NK, Orlando L, Gilbert MTP, Kjaer KH, Alsos IG, Willerslev E (2012) Glacial survival of boreal trees in northern Scandinavia. *Science*, **335**, 1083–1086.

Paus A, Velle G, Larsen J, Nesje A, Lie Ø (2006) Lateglacial nunataks in central Scandinavia: Biostratigraphical evidence for ice thickness from Lake Flåfattjønn, Tynset, Norway. *Quaternary Science Reviews*, **25**, 1228–1246.

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

Petit J, Brewer S, Cheddadi R, Coart E, Cottrell J, Csaikl UM, Van Dam B, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JA, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Popescu F, Slade D, Tabbener H, De Vries SGM, Ziegenhagen B, De Beauliue J-L, Kremer A (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**, 564–571.

Reid C (1899) *The origin of the British flora*. Dulau, London, United Kingdom.

Reille M, de Beaulieu J-L, Svobodova H, Andrieu-Ponel V, Goeury C (2000) Long, Pollen analytical biostratigraphy of the last five climatic cycles from a continental sequence from the Velay region (Massif Central, France). *Journal of Quaternary Science*, **15**.

Rybníčková E, Rybníček K (1991) The environment of the Pavlovian: palaeoecological results from Bulhary, South Moravia. In: *Palaeovegetational*

development in Europe, Proc. Pan-European Palaeobotanical Conference (ed. Kovar-Eder J), pp. 73–79. Museum of Natural History, Wien, Austria.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, **16**, 608–613.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of The Royal Society. Biological sciences*, **277**, 661–671.

Svenning J, Normand S, Skov F (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, **31**, 316–326.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tallantire (1974) The palaeohistory of the grey alder (*Alnus incana* (L.) Moench.) and black alder (*A. glutinosa* (L.) Gaertn.) in Fennoscandia. *New Phytologist*, **73**, 529–546.

Tarasov PE, Volkova VS, Webb III T, Guiot J, Andreev AA, Bezusko LG, Bezusko TV, Bykova GV, Dorofeyuk NI, Kvavadze EV, Osipova IM, Panova NK, Sevasrtyanov DV (2000) Last glacial maximum biones reconstructed from pollen and plant macrofossil data from Northern Eurasia. *Journal of Biogeography*, **27**, 609–620.

Tollefsrud MM, Kissling R, Gugerli F *et al.* (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.

Tollefsrud MM, Sønstebø JH, Brochmann C, Johnsen O, Skroppa T, Vendramin GG (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity*, **102**, 549–562.

Tzedakis P (1994) Hierarchical biostratigraphical classification of long pollen sequences. *Journal of Quaternary Science*, **9**, 257–259.

Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696–704.

Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC (2002) Buffered tree population changes in a quaternary refugium: evolutionary implications. *Science*, **297**, 2044–2047.

Velichko AA, Kremenetsky KV, Negendank J, Mingram J, Borisova OK, Gribchenko YuN, Zelikson EM, Klimanov A, Novenko EYu, Priumova LG, Pisareva VV, Razumovsky LV, Timireva SN (2001) Late Quaternary palaeogeography of the North-East Europe (based on the complex study of the Galich Lake sediments). *Izvestiya Akademii Nauk, Seriya Geograficheskaya,* **3**, 42–54.

Vorren TO, Vorren K-D, Aasheim O, Dahlgren KIT, Forwick M, Hassel K (2013) Palaeoenvironment in northern Norway between 22.2 and 14.5 cal. Ka BP. *Boreas*, **42,** 876–895.

Ward JK, Harris JM, Cerling TE, Wiedenhoeft A, Lott MJ, Dearing M-D, Coltrain JB, Ehleringer J (2005) Carbon starvation in glacial trees recovered from the La Brea tar pits, southern California. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 690–694.

Webb T, Bartlein P (1992) Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics*, **23**, 141–173.

West R (1980) Pleistocene forest history in East Anglia. *New Phytologist*, **85**, 571–622.

Wijmstra T (1969) Palynology of the first 30 metres of a 120 m deep section in Northern Greece. *Acta Botanica Neerlandica*, **18**, 511–527.

Willis K, Van Andel T (2004) Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, **23**, 2369–2387.

Willis KJ, Rudner E, Sümegi P (2000) The full-glacial forests of Central and Southeastern Europe. *Quaternary Research*, **53**, 203–213.

3. SCIENTIFIC PAPERS

Paper I

Douda J, Doudová J, **Drašnarová A**, Kuneš P, Hadincová V, Krak K, Zákravský P, Mandák B (2014) Migration patterns of subgenus *Alnus* in Europe since the Last Glacial Maximum: a systematic review. *PloS one*, **9**, e88709.

Paper II

Drašnarová A, Krak K, Vít P, Doudová J, Douda J, Hadincová V, Zákravský P, Mandák B (2014) Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genetics & Genomes*, **10**, 865-873.

Paper III

Mandák B, Vít P, Krak K, Trávníček P, **Havrdová A**, Hadincová V, Zákravský P, Jarolímová V, Bacles CFE, Douda J. (in press) Putative glacial refugia inferred from the geographic distribution of *Alnus glutinosa* cytotypes in Europe. *Annals of Botany.*

Paper IV

Mandák B, **Havrdová A**, Krak K, Hadincová V, Vít P, Zákravský P, Douda J (submitted, *New Phytologist*) Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation.

Paper V

Havrdová A, Douda J, Krak K, Vít P, Hadincová V, Zákravský P, Mandák B (in press) Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture. *Molecular Ecology*.

For supporting information to these papers see pages 191–241.

OPEN ACCESS Freely available online

PLOS | ONE

Migration Patterns of Subgenus *Alnus* in Europe since the Last Glacial Maximum: A Systematic Review

Jan Douda[1]*, Jana Doudová[2], Alena Drašnarová[1,2], Petr Kuneš[3], Věroslava Hadincová[2], Karol Krak[2], Petr Zákravský[2], Bohumil Mandák[1,2]

1 Department of Ecology, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic, 2 Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice, Czech Republic, 3 Department of Botany, Faculty of Science, Charles University in Prague, Prague, Czech Republic

Abstract

Background/Aims: Recently, new palaeoecological records supported by molecular analyses and palaeodistributional modelling have provided more comprehensive insights into plant behaviour during the last Quaternary cycle. We reviewed the migration history of species of subgenus *Alnus* during the last 50,000 years in Europe with a focus on (1) a general revision of *Alnus* history since the Last Glacial Maximum (LGM), (2) evidence of northern refugia of *Alnus* populations during the LGM and (3) the specific history of *Alnus* in particular European regions.

Methodology: We determined changes in *Alnus* distribution on the basis of 811 and 68 radiocarbon-dated pollen and macrofossil sites, respectively. We compiled data from the European Pollen Database, the Czech Quaternary Palynological Database, the Eurasian Macrofossil Database and additional literature. Pollen percentage thresholds indicating expansions or retreats were used to describe patterns of past *Alnus* occurrence.

Principal Findings: An expansion of *Alnus* during the Late Glacial and early Holocene periods supports the presence of alders during the LGM in southern peninsulas and northerly areas in western Europe, the foothills of the Alps, the Carpathians and northeastern Europe. After glaciers withdrew, the ice-free area of Europe was likely colonized from several regional refugia; the deglaciated area of Scandinavia was likely colonized from a single refugium in northeastern Europe. In the more northerly parts of Europe, we found a scale-dependent pattern of *Alnus* expansion characterised by a synchronous increase of *Alnus* within individual regions, though with regional differences in the times of the expansion. In southern peninsulas, the Alps and the Carpathians, by contrast, it seems that *Alnus* expanded differently at individual sites rather than synchronously in whole regions.

Conclusions: Our synthesis supports the idea that northern LGM populations were important sources of postglacial *Alnus* expansion. The delayed *Alnus* expansion apparent in some regions was likely a result of environmental limitations.

Citation: Douda J, Doudová J, Drašnarová A, Kuneš P, Hadincová V, et al. (2014) Migration Patterns of Subgenus *Alnus* in Europe since the Last Glacial Maximum: A Systematic Review. PLoS ONE 9(2): e88709. doi:10.1371/journal.pone.0088709

Editor: Jerome Chave, Centre National de la Recherche Scientifique, France

Received September 29, 2013; **Accepted** January 9, 2014; **Published** February 21, 2014

Copyright: © 2014 Douda et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by grant no. P504/11/0402 from the Grant Agency of the Czech Republic (http://www.gacr.cz/en/) and grant CIGA no. 20124201 from the Czech University of Life Sciences Prague (http://ga.czu.cz/) and as part of the long-term research development project no. RVO 67985939 (http://www.msmt.cz/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: douda@fzp.czu.cz

Introduction

The recent distribution of species in the Northern Hemisphere has been significantly influenced by processes occurring in the last Quaternary cycle, during the last glacial period and subsequent Holocene warming [1], [2]. The 'classic' paradigm states that during the Last Glacial Maximum (LGM, i.e., from 26.5 to 19 to 20 thousand years before present (kyr BP) [3]), temperate plant species, particularly climate-sensitive trees, were harboured in low-latitude refugia. In Europe, southern peninsulas (i.e., Iberian, Italian and Balkan) served as refugial areas for many species [4], [5].

Recently, new palaeoecological records supported by molecular analyses and palaeodistributional modelling have provided more comprehensive insights into plant behaviour during the last Quaternary cycle [6], [7]. In eastern Europe, more northerly distributions of many temperate and boreal plants during the last glacial period have been confirmed, although fossil records directly from the LGM are scarce. The eastern Alps, northern Dinaric Alps, the Carpathians and the Pannonian region probably served as northern refugia for many temperate tree species, namely *Abies alba*, *Carpinus betulus*, *Fagus sylvatica*, *Taxus baccata* and *Ulmus* [6], [8], [9], [10]. Open taiga and hemiboreal forests dominated by *Larix*, *Pinus*, *Picea* and *Betula* likely occurred in the northern Carpathians, Belarus and the northwestern Russian plains [11], [12], [13].

The early postglacial expansion of trees in northern areas thus need not reflect migration from southern regions but may be the result of the population growth and expansion of small tree populations persisting in scattered refugia relatively close to the margin of the ice sheet [14], [15], [16]. Climatic warming has been determined as the most important driver initiating the expansion of trees [1]. However, regional differences in climatic

and environmental conditions recorded for the Late Glacial and early Holocene periods could have resulted in nontrivial species-specific and regionally dependent patterns of expansion [17], [18], [19].

Taxonomic Status

The genus *Alnus* Mill. belongs to the family Betulaceae [20], [21]. The oldest macrofossil records assigned to *Alnus* have been reported from the middle Eocene, but fossil pollen grains of *Alnus* from the Late Cretaceous have also been found [22]. The genus *Alnus* comprises about 29 to 35 species of monoecious trees and shrubs distributed throughout the Northern Hemisphere and along the Andes in South America [23]. In Europe, three species of subgenus *Alnus* (i.e. *Alnus cordata*, *A. glutinosa* and *A. incana*) and one species of subgenus *Alnobetula* (*A. alnobetula* (Ehrh.) K. Koch) occur [24]. It has been estimated using molecular methods that the subgenera *Alnus* and *Alnobetula* diverged in the Eocene, 48.6 million years (Myr) BP [25]. *A. cordata* separated from the *A. glutinosa-incana* complex in the Oligocene (22.9 Myr BP), and *A. glutinosa* and *A. incana* diverged in the Pliocene (7.9 Myr BP) [25].

Upper Pleistocene and Holocene *Alnus* History

Pleistocene pollen and macrofossil data indicate repeated population increases and decreases of *Alnus* in Europe, reflecting climate oscillations between glacial periods and interglacials, particularly noticeable in the Middle and Upper Pleistocene [5], [26], [27]. The majority of Upper Pleistocene pollen profiles support a common presence of *Alnus* in the Eemian interglacial (Marine Isotope Stage, MIS 5e–5d) throughout Europe (i.e., 125–115 kyr BP) and its disappearance after the start of the last glacial period [i.e., Hollerup (DK) – [28]; Tenaghi Philipon (GR) – [26]; Valle di Castiglione (IT) – [29]; Les Echets (FR), La Grande Pile (FR) – [30], [31]; Praclaux Crater (FR), Ribains (FR) – [32]; Ioannina (GR) – [5]; Jammertal (DE) – [33]].

In their classic study, Huntley and Birks [1] assumed that the main source refugia for the *Alnus* expansion after the LGM lay in the eastern Alps, the Carpathians and the Ukrainian lowlands. Other LGM refugia were located in Corsica, western France, northern Spain and northwestern Russia. The authors supposed that the Holocene migration of *Alnus* likely began somewhere in eastern Europe and continued by the northward expansion of *A. glutinosa* (L.) Gaertn. and *A. incana* (L.) Moench to the Baltic and Scandinavia and by the westward expansion of *A. glutinosa* along the southern shore of the North Sea as far as the British Isles [1].

A large-scale genetic survey that included Europe and Turkey and focused on the postglacial history of *Alnus glutinosa* accepted Huntley and Birks' migration patterns [34]. King and Ferris [34] revealed 13 cpDNA haplotypes of *A. glutinosa* mainly associated with southern European peninsulas. They suggested that two of these haplotypes colonized northern and temperate Europe from LGM refugia located in the Carpathians. While the first haplotype expanded primarily into western Europe, the second mainly colonized northern Europe [34]. However, the presence of only two haplotypes in the northern part of Europe limits a more detailed determination of *A. glutinosa* migration patterns. Surprisingly, no follow-up study focusing on postglacial migration pattern of *A. glutinosa* has since been published. No phylogeographic study has been performed for *A. incana* thus far, either.

Since the seminal studies of Huntley and Birks [1] and King and Ferris [34] were published, many palaeoecological studies have presented new knowledge about the history of *Alnus* in Europe during the last glacial period and Holocene. We reviewed the migration history of species of subgenus *Alnus*, including *A. glutinosa* and *A. incana*, during the last 50,000 years in Europe based on large numbers of pollen records and macrofossil remains. Pollen of different species of subgenus *Alnus* (further collectively referred to as "*Alnus*") is indistinguishable in palaeoecological studies, but alder species can be identified based on macrofossil remains. In particular, we focused on (1) a general revision of *Alnus* history since the LGM, (2) evidence of northern refugia of *Alnus* during the LGM and (3) the specific history of *Alnus* in particular European regions.

Alnus species are keystones of alluvial and wetland habitats [35], [36] distributed through the European forest zones from the northern treeline to the Mediterranean. Understanding their last glacial occurrence and postglacial migration pattern may shed light upon the resistance and resilience of wetland forest habitats in the course of global climate change. The results of our study allow us to propose guidelines for the sampling design and interpretation of a future detailed phylogeographic and population-genetic survey of *Alnus* species in Europe.

Materials and Methods

Study Species

Two common tree species of *Alnus* grow natively in Europe [24]. Black alder (*A. glutinosa*) is considered a temperate tree. It commonly occurs in the lowlands and mountains across Europe except Scandinavia, where it is associated with a coastal oceanic climate in southern areas [37] (Figure 1). The cold-climate limitation also likely affects its distribution in high-elevation mountainous areas, where black alder populations are often absent. Scarce distributions are found in the Mediterranean region and in the arid Great Hungarian plains, the Ukraine and the Russian steppe zone. Outside Europe, the distribution extends as far as western Siberia and the mountains of Turkey, Iran and North Africa [24], [38]. In Corsica and southern Italy, *A. glutinosa* grows sympatrically with *A. cordata* (Loisel.) Duby.

Grey alder (*Alnus incana*) is considered a boreal and mountain tree. Similar to Norway spruce (*Picea abies*), the range of *A. incana* is divided into a northern and a southern area, which meet in the Polish lowlands. In northern Europe, *A. incana* continuously covers the east Baltic region and all of Scandinavia with a northern margin at latitudes greater than 70°N [24], [37]. In northern Scandinavia, the nominal subspecies grows sympatrically with *A. incana* subsp. *kolaensis* (Orlova) Á. Löve & D. Löve [24]. The distribution of grey alder continues eastwards across European Russia to western Siberia, which contrasts with its patchy mountain occurrence in the southern part of the range linked to the Alps, the northern Apennines, the Hercynian Mountains, the Carpathians, the Bulgarian Mountains, the Dinaric Alps, the Caucasus and Turkey [24].

Alnus glutinosa and *A. incana* dominate in floodplain and swamp forests. These species are indifferent to soil nutrient conditions, except for extremely poor peat bogs. Seeds are dispersed effectively by water, while wind dispersal is commonly limited to the vicinity of the parent tree [38]. Under unfavourable environmental conditions, such as in cold climates, *A. incana* is able to survive and reproduce by clonal growth [39]. Compared with the relatively short-lived *A. incana* (c. 20–50 years), *A. glutinosa* is a long-lived tree (c. 100–120 years), although the age of reproduction is similar for the two species (i.e., 10–20 years) [37], [38].

Pollen Data

This systematic review follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement as a guide [40] (see Checklist S1). We compiled freely available data

Figure 1. Pollen sites (dots) and European regions (dotted lines) included in this study. a, Iberian region; b, Italian region; c, Balkan region; d, the Carpathians; e, the Alps; f, Baltic and northeastern European plains; g, Scandinavia; h, Hercynian Mountains and Massif Central; i, western European plain; j, British Isles. Bold dashed and dashed-dotted lines show the northern boundary of *Alnus glutinosa* in Scandinavia and the western range boundary of *A. incana* in Western Europe, respectively [24].
doi:10.1371/journal.pone.0088709.g001

from the European Pollen Database (EPD, http://europeanpollendatabase.net, [41]), the Czech Quaternary Palynological Database (PALYCZ, http://botany.natur.cuni.cz/palycz, [42]) and additional literature (Figure 1). The search for additional literature was performed in September 2011 in Web of Science and augmented by Google Scholar. The search included combinations and derivations of the following terms: radiocarbon dates, pollen, wood remains, macrofossils, glacial, vegetation, LGM, Holocene and Europe. To guarantee the chronological accuracy of changes in *Alnus* distribution, we used only pollen data with radiocarbon dating. In total, we used 553 and 258 pollen profiles from databases and the literature, respectively. The list of original publications and sites available in September 2011 is provided in Tables S1 and S3.

Age-depth models were constructed and radiocarbon dates were calibrated (cal.) for all profiles in the EPD [43], PALYCZ and publications using the CLAM code [44] in R [45]. The age-depth models were constructed using smoothing-spline fitting with a default smoothing factor of 0.3 or linear interpolation with preferences for a smoothing spline. Possible errors in the pollen diagram chronology were minimised in several ways. We excluded parts of the chronology outside the marginal ^{14}C dates. Additionally, to determine the oldest unquestionable time of expansion, we checked parts of the pollen diagrams indicating the start of the *Alnus* expansion (i.e., ≥2.5% pollen threshold) to determine whether i) the nearest radiocarbon date is closer than 2,000 years to time of the expansion, ii) there is no presence of reworked pollen or iii) the expansion does not start at the end of the previous 1,000-year interval. Reworking was assumed when an isolated pollen spectrum with *Alnus* ≥2.5% was recorded or when the basal spectra of *Alnus* ≥2.5% were followed by a steep decrease in pollen.

To describe the temporal patterns of *Alnus* occurrence at particular sites, we recorded the pollen percentage of *Alnus* at 1,000-year intervals in the time period from the present to 26 cal. kyr BP (i.e., the start of the LGM) and at 5,000-year intervals in the time period preceding the LGM. The average percentage of pollen of *Alnus* at the site in each time interval (1,000 and 5,000 year) was calculated by dividing the *Alnus* pollen count by the total pollen sum in each sample after excluding aquatic species, cryptogam spores and indeterminable pollen. We excluded pollen of subgenus *Alnobetula* from the total *Alnus* pollen count. Due to their specific pollen morphology, pollen grains of species of subgenus *Alnobetula* are identified and counted separately in

palaeoecological studies [46], [47]. We chose 5,000-year intervals before 26 kyr BP because the pollen records were fragmentary and the rare radiocarbon dates do not sufficiently cover the pollen profiles. The total pollen sum was calculated in most literature sources in the same way, allowing us to determine past *Alnus* pollen value for each time period by simple visual inspection of pollen diagrams.

Because alders are high pollen producers, they are generally overrepresented in pollen diagrams [1]. Moreover, *Alnus glutinosa* and *A. incana* often dominate in swamps, at lake and stream shores and at the margins of peat bogs in close vicinity to sample sites [48]. To record the regional presence of *Alnus* from pollen diagrams, several thresholds ranging from 0.5 to 8% have been used in the literature [49], with greater agreement for 2–3% [1], [50], [51], [52]. We used the 2.5% threshold suggested for *Alnus* in a recent study comparing modern pollen data with European tree species distribution [49]. This 2.5% threshold corresponds to the presence of *Alnus* within approximately 50 km of a pollen site [49]. We also incorporated the threshold of 0.5% as an indicator of possible scarce regional occurrence despite the risk of contamination by long-distance pollen transport. Lisitsyna et al. [49] still found strong agreement between pollen presence defined by the 0.5% threshold and the regional occurrence of *Alnus*. Pollen values greater than the 10% threshold are assumed to correspond to the occurrence of an *Alnus*-dominated forest at the site [1], [37], [51]. In summary, four percentage categories were used to describe the patterns of past *Alnus* occurrence, where less than 0.5% indicates the regional absence of a species, 0.5–2.5% may be the result of long-distance pollen transport but could also capture the presence of relatively small populations in the region, 2.5–10% indicates a species' presence within the region, and values greater than 10% indicate the local presence of a species at the site. The description of *Alnus* distribution in the results and discussion section is based on ≥2.5% pollen records to eliminate possible misinterpretation based on the 0.5% threshold.

Macrofossil Data

To obtain macrofossil evidence (e.g., cones, fruits, male catkins, twigs, wood pieces), we used free data available from the Eurasian Macrofossil Database (NEMD, http://oxlel.zoo.ox.ac.uk/reference-collection, [13]) and additional published records. In total, we used macrofossil data from 14 sites in the database and 54 sites in the literature (Tables S2 and S3). Macrofossils of *Alnus glutinosa* and *A. incana* were determined at 38 and 15 sites, respectively. Macrofossil records were assigned according to 1,000- or 5,000-year pollen intervals based on constructed age-depth models (see Pollen data chapter). We interpreted only the *Alnus* presence, as it is problematic to evaluate data regarding the absence or abundance of macrofossils [9].

Pollen and Macrofossil Maps

The pollen and macrofossil maps indicate *Alnus* occurrence at particular time periods during the last 50,000 years. We merged 5,000- and 1,000-yr intervals with a limited number of records to logical periods of the last glacial period; 50–26 cal. kyr BP includes the period preceding the LGM, 26–20 cal. kyr BP the period of the LGM and 20–15 cal. kyr BP the period after the LGM, also known as the Oldest Dryas. The macrofossil remains and maximum pollen thresholds recorded during the merged periods were plotted in maps. We also marked changes in the pollen percentages between the time periods, indicating the expansion, stability or decrease of *Alnus*. The term "*Alnus*" indicates macrofossils that were not assigned to individual species in original studies whereas the names "*Alnus glutinosa*" and "*A. incana*" refer to those that were.

Regional Differences in Late Glacial and Holocene History

To determine the specific postglacial history of *Alnus* in individual European regions, we delimited 10 regions based on different environmental conditions in the last glacial period and the Holocene (Figure 1). The Iberian, Italian and Balkan regions include areas considered southern LGM refugia of trees (Figure 1, regions a–c). The Baltic and northeastern European plains, Scandinavia and the British Isles are regions that were largely covered by the Scandinavian ice sheet during the LGM (Figure 1, regions f, g, j). The Carpathians and Alps covered areas of potential LGM refugia for some temperate and many boreal trees (Figure 1, regions d, e). The Hercynian Mountains, the Massif Central and highlands located to the north of the Alps were mostly ice-free regions (Figure 1, region h). Ice-free lowland areas of the Western European plain were influenced by the oceanic climate (Figure 1, region i). We determined the proportion of pollen sites in each region and time period that reached the 0.5%, 2.5% and 10% thresholds. Only time intervals with more than 10 sites available in particular regions were considered in the analysis. The region of the Great Hungarian plains was excluded from all analyses because fewer than 10 pollen sites had been found there.

Results

Pre-LGM *Alnus* Distribution (50–26 cal. kyr BP)

In southern Europe, *Alnus* exceeds the 2.5% pollen threshold in the Pyrenees Mountains [53] and at several Italian sites [54], [55], [56], [57] (Figure 2A). Other pollen records exceeding 2.5% have been obtained from northwestern France [58] and the western Russian plains [59]. In western Russia and Belarus, the occurrence of *A. glutinosa* and *A. incana* is supported by macrofossil remains [60], [61], [62]. *Alnus* macrofossil records are present along the northern border of the Pannonian lowlands in the Czech Republic and the northeastern foothills of the Carpathians in Romania [10] (Figure 2A).

LGM *Alnus* Distribution (26–20 cal. kyr BP; Figure 2B)

At the LGM, the *Alnus* pollen values decrease in Italy and France (Figure 2B). The only 2.5%-threshold pollen evidence for *Alnus* occurrence in southern-European peninsulas was detected in the Pyrenees Mountains [53]. Further north in Europe, *Alnus* pollen values exceed the 2.5% threshold at two sites in the Bodmin moor in Cornwall [63] and in the Timan Ridge in Arctic Russia [64] (Figure 2B).

Late Glacial *Alnus* Distribution (20–12 cal. kyr BP; Figure 2C, 2D, 3A and 3B)

Between 20 and 15 cal. kyr BP, the 2.5%-threshold pollen evidence of *Alnus* continues in southern England and Arctic Russia (Figure 2C). In southern Europe, only one new 2.5%-threshold pollen record has emerged in the Rila Mountains in Bulgaria [65]. Macrofossil remains of *Alnus* occur in the southwestern foothills of the Alps [66], the Thracian plain in Bulgaria [67] and southern Lithuania [61] (Figure 2C).

Between 15 and 12 cal. kyr BP, several pollen sites exceed the 2.5% *Alnus* threshold in the southwestern and western parts of the Alps (Figs 2D, 3A and 3B). Moreover, macrofossil remains of *A. glutinosa* occur there [66] (Figure 3B). South of the Alps, *Alnus* pollen increases and reaches more than 2.5% in Corsica [68] (Figure 3A), the northern Apennines [69] (Figure 3B) and central Italy [70] (Figure 3B). In the Carpathians, *Alnus* pollen records

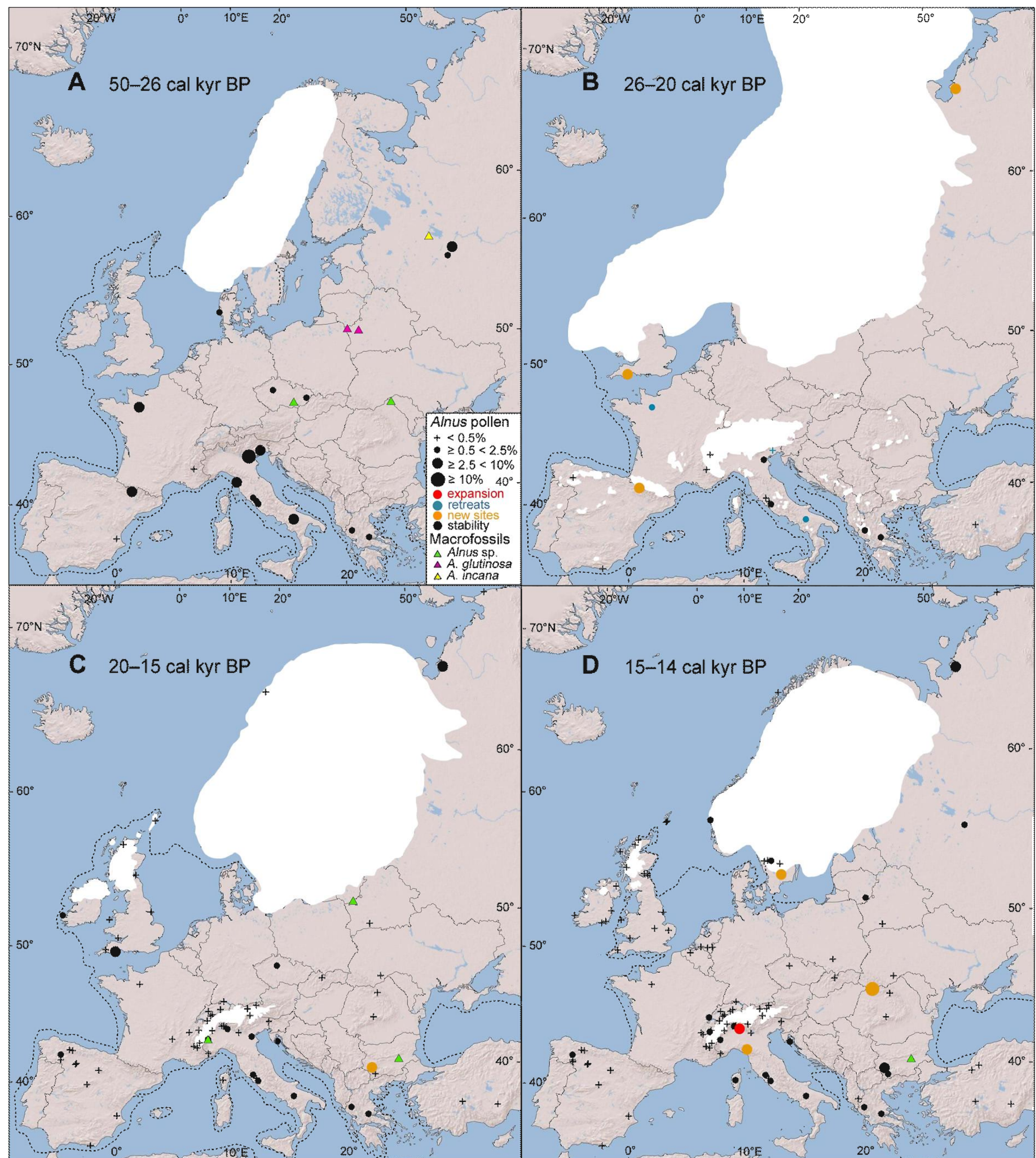


Figure 2. Last glacial period distribution (50–14 cal. kyr BP) of *Alnus* pollen sites. According to four classes of the percentage of *Alnus* pollen and macrofossil remains. The dot colour indicates changes compared with the previous period: red, expansion, *Alnus* pollen <2.5% in preceding period; blue, retreat, *Alnus* pollen ≥2.5% in preceding period; orange, new pollen sites of *Alnus* pollen ≥2.5%; black, stability; the course of deglaciation (white) and changes in coastline (dotted lines).
doi:10.1371/journal.pone.0088709.g002

exceeding 2.5% are present in the Gutaiului Mountains in northwestern Romania [71] (Figure 2D). Sites with evidence of more than 2.5% of *Alnus* pollen emerge in southern Scandinavia [72] (Figure 2D), Estonia [73], [74] (Figure 3A) and northwestern and western Russia [75], [76] (Figure 3A). Macrofossil remains occur in Poland [77] (Figure 3A), Lithuania [61], [78] (Figure 3A and 3B) and Belarus [62] (Figure 3B).

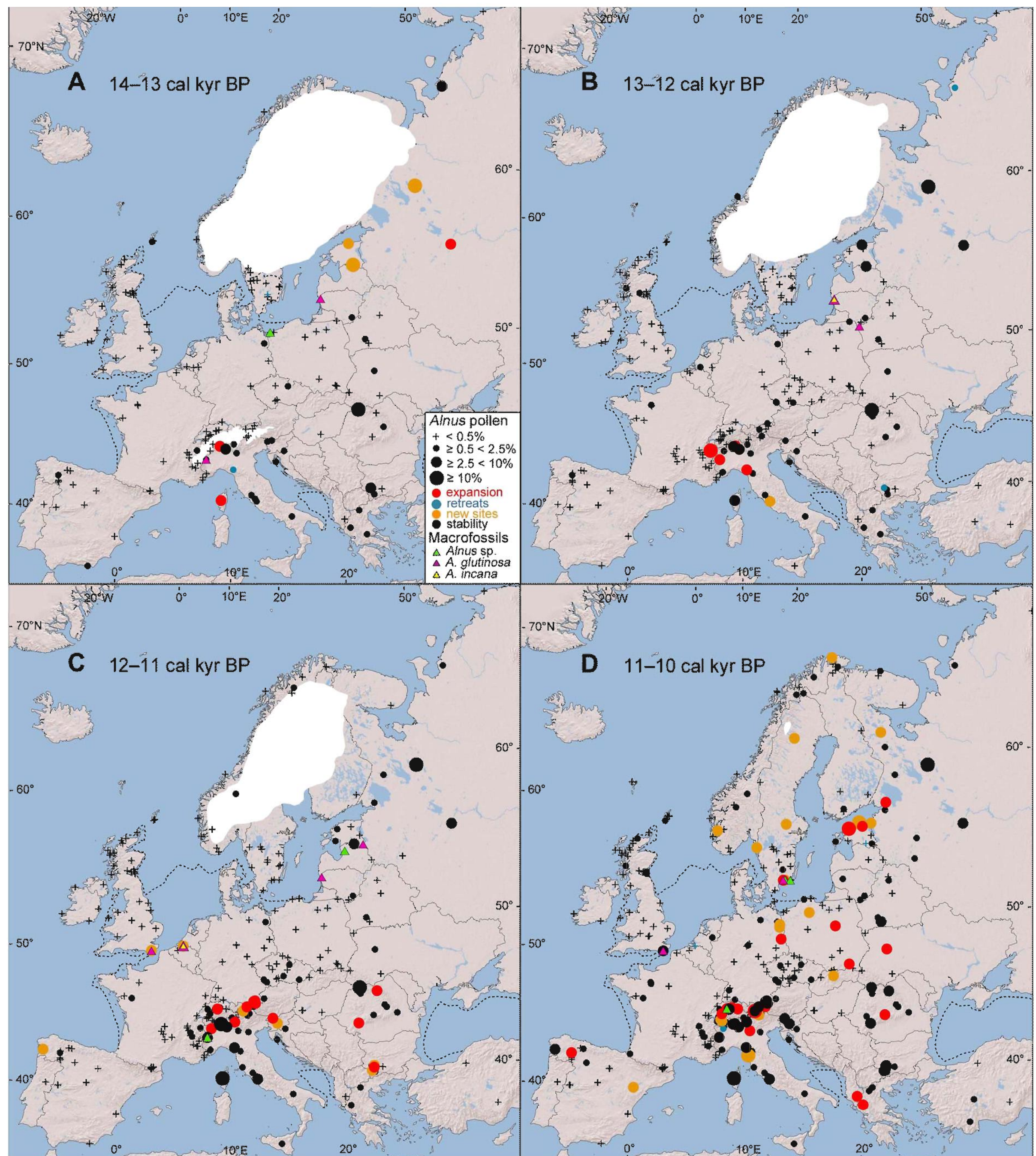


Figure 3. Late Glacial and early Holocene distribution (14–10 cal. kyr BP) of *Alnus* pollen sites. According to four classes of the percentage of *Alnus* pollen and macrofossil remains; for details, see Figure 2.
doi:10.1371/journal.pone.0088709.g003

Holocene *Alnus* Distribution (12–0 cal. kyr BP; Figure 3C, 3D and 4, Figure S1 and S2)

At the beginning of the Holocene (i.e., 12–11 cal. kyr BP, Figure 3C), a continual increase in the number of sites with at least 2.5% *Alnus* pollen is apparent across the Alps, with the exception of the western areas. In western Europe, macrofossils of *A. glutinosa* and *A. incana* are present at the Kreekrak site in southwestern Netherlands [79] and *A. glutinosa* in Pannel Bridge, East Sussex [80]. Several pollen sites exceed the 2.5% pollen threshold in the Romanian Carpathians [81] and the Rila and Pirin Mountains in Bulgaria [82], [83], [84]. The first piece of evidence since the

LGM of more than 2.5% of *Alnus* pollen has been recorded in the Iberian peninsula [85] (Figure 3C).

Between 11 and 10 cal. kyr BP (Figure 3D), many sites reach at least 2.5% of *Alnus* pollen in a large area of the Polish lowland, the northern Carpathians and Scandinavia, including its northern part [86], [87]. An increase of sites exceeding the 2.5% pollen threshold is also evident in the Iberian and the Balkan peninsula (Figure 3D).

Between 10 and 9 cal. kyr BP, the majority of localities in the Carpathians and the Baltic region, including southern Scandinavia, exceed the 2.5% *Alnus* pollen threshold (Figure 4A). Macrofossil remains of *A. glutinosa* occur in the northern border of its recent distribution in central Sweden [14]. By contrast, few sites with more than 2.5% *Alnus* pollen are present in a large zone running from the Bohemian Massif and the northern foothills of the Alps through the Massif Central and the French Alps to western Europe and the British Isles (Figure 4A).

Between 9 and 8 cal. kyr BP, the increase of sites with more than 2.5% of *Alnus* pollen is apparent over the British Isles, the northern foothills of the Alps, the Bohemian Massif, northern Scandinavia and likely in the western European plain (Figure 4B). During the next two millennia (i.e., 8–6 cal. kyr BP), many sites with 2.5% *Alnus* evidence emerge in the French Alps, northern Scotland, Ireland and all southern peninsulas (Figure 4C and 4D). Finally, the number of sites exceeding 2.5% of *Alnus* pollen increases in the Massif Central and the remaining unoccupied areas of France between 7 and 6 cal. kyr BP (Figure 4D).

During the period between 6 and 0 cal. kyr BP, a decrease in the number of sites with more than 2.5% *Alnus* pollen is present in large areas of Europe, likely except in the southern peninsulas and the Carpathians (Figure 5; Figure S1 and S2). After 6 cal. kyr BP, *Alnus* enters a period of retreat in northern Scandinavia and continues southward up to the present (Figure 5G; Figures S1 and S2). In other regions, a decrease is apparent during approximately the last three millennia. A relatively strong decrease appears in the Alps (Figure 5E), Hercynian Mountains (Figure 5H), the western European plain (Figure 5I) and the British Isles (Figure 5J) whereas a weak decrease is apparent in the Baltic region (Figure 5F).

Regional Differences at the Beginning of the *Alnus* Expansion

In the southern peninsulas, the Alps and the Carpathians, there is an increase in the number of sites exceeding the 2.5% pollen threshold beginning in the Late Glacial period and increasing gradually during most of the Holocene (Figure 5A–5E). In more northerly regions, the number of sites with more than 2.5% pollen evidence rises abruptly after the beginning of the Holocene. Specifically, an increase in the number of sites in the Baltic region (Figure 5F) and Scandinavia (Figure 5G) starts between 11 and 10 cal. kyr BP and over three thousand years reaches more than 80% of occupied sites. In Hercynian Mountains (Figure 5H), the western European plain (Figure 5I) and the British Isles (Figure 5J), the expansion starts between 10 and 9 cal. kyr BP, and 80% of sites are occupied after four thousand years.

Discussion

Northern LGM Refugia

For the Last Glacial Maximum, there are two records with more than 2.5% *Alnus* pollen (Figure 6) from the periglacial landscape of the Scandinavian ice sheet in southern England [63] and Arctic Russia [64], but they are likely influenced by wind pollen transport from more distant sites. This is indicated by a low concentration of *Alnus* pollen and the presence of steppe taxa in pollen profiles [64], [88].

Because of the absence of reliable records from the LGM, we used pollen sites and macrofossils from the Late Glacial and early Holocene periods as indicators of possible *Alnus* LGM refugial areas (Figure 6). These sources indicate the presence of *Alnus* during the LGM in western Europe, the northern foothills of the Alps, the Romanian Carpathians and a large area of northeastern Europe (Figure 6). Evidence of more than 2.5% pollen from sites located in northeastern Europe from the Late Glacial period are generally interpreted as a reworking of earlier climatically favourable periods or long-distance dispersal [72], [73], [74], [75], [76] but macrofossil remains found in Poland, Belarus, Lithuania and Latvia support the occurrence of *Alnus* in this area (Figure 6).

The ability of alder trees to tolerate the climatic conditions of the LGM in northern areas has been supported in several ways. Kullman [39] showed a high tolerance of *Alnus incana* to cold climates by assessing its regeneration patterns in a subalpine forest of central Sweden. He suggested that *A. incana* could have survived the last glacial period in northern areas because it has high vegetative survivability far above its generative limit. Palaeodistributional modelling based on the climatic tolerance of trees has suggested the possible existence of *A. incana* in the proximity of the ice sheet, including southern England, northern France, Central Europe, the northern Carpathians and Belarus, but this modelling has also suggested that *A. incana* was absent from the northwestern Russian plains [7]. The northern occurrence of *A. glutinosa* reached France and the northern foothills of the Alps, but the species was absent from the northern Carpathians, Belarus and the northwestern Russian plains [7]. The survival of *Alnus* species in the North throughout the LGM might be supported by their occurrence in floodplains, which were moister and more sheltered sites than the typical dry habitats of the surrounding uplands with the occurrence of permafrost [13].

Southern LGM Refugia

Surprisingly, the 2.5% threshold does not support the Mediterranean peninsulas as LGM refugial areas for *Alnus* with the sole exception of the Pyrenees [53]. This finding contradicts the conclusions of a phylogeographic study on *A. glutinosa* that detected specific cpDNA haplotypes for particular southern peninsulas [34]. A recent population-genetic study of Lepais et al. [89], supported also by pollen data [90], [91] highlights the behaviour of rear-edge stable populations of *A. glutinosa* in North Africa. They found that tetraploid *A. glutinosa* populations in Morocco have diverged for a long-time without contribution of gene flow of Algerian or Tunisian diploid populations. This supports to the idea that *Alnus* survived in the Mediterranean area at mesoclimatically favourable sites (e.g., in foothill valleys) in sparse and isolated populations, which are generally hard to detect by pollen analyses [4], which possibly explains the low percentage of *Alnus* pollen.

Holocene *Alnus* Expansion in Northern Regions

The expansion of *Alnus* began in the Baltic region and Scandinavia between 11 and 10 cal. kyr BP (Figure 3D). The absence of *Alnus* evidence in most of central and northwestern Europe indicates that populations in northeastern Europe were predominant sources for the colonisation of Scandinavia. The delayed expansion of *Alnus* in the British Isles between 10 and 8 cal. kyr (Figure 4A and 4B) appears to have originated in a western European refugium [92], [93], [94], [95] rather than in eastern Europe, as suggested by Huntley and Birks [1]. However, eastern populations, which colonised the Baltic states and

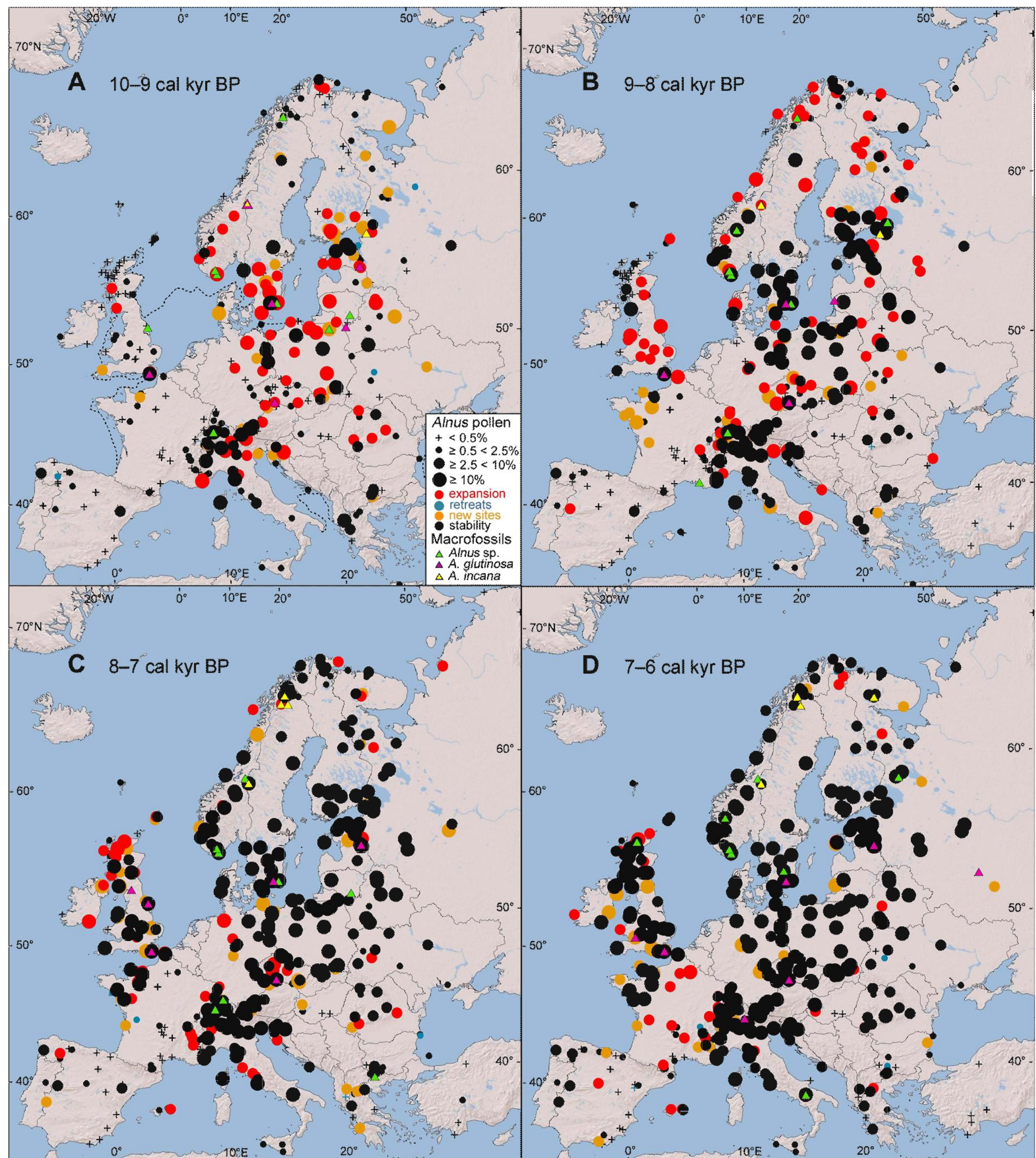


Figure 4. Holocene distribution (10–6 cal. kyr BP) of *Alnus* pollen sites. According to four classes of the percentage of *Alnus* pollen and macrofossil remains; for details, see Figure 2.
doi:10.1371/journal.pone.0088709.g004

Scandinavia, could have spread southwest and mixed with western populations [93]. Synchronously with the rise of *Alnus* in the British Isles, alders expanded in Hercynian Mountains, but it is impossible to tell whether Baltic, Carpathian, Alpine or local alder populations contributed to this expansion (Figure 4A and 4B). Source populations are also unknown for the *Alnus* expansion in the Massif Central and the remaining unoccupied area of France between 7 and 6 cal. kyr BP (Figure 4D).

Scale-dependent pattern of alnus expansion. In northern areas, the *Alnus* expansion shows a scale-dependent pattern characterised by a synchronous increase of *Alnus* within individual regions, but with regional differences in the times of the expansion.

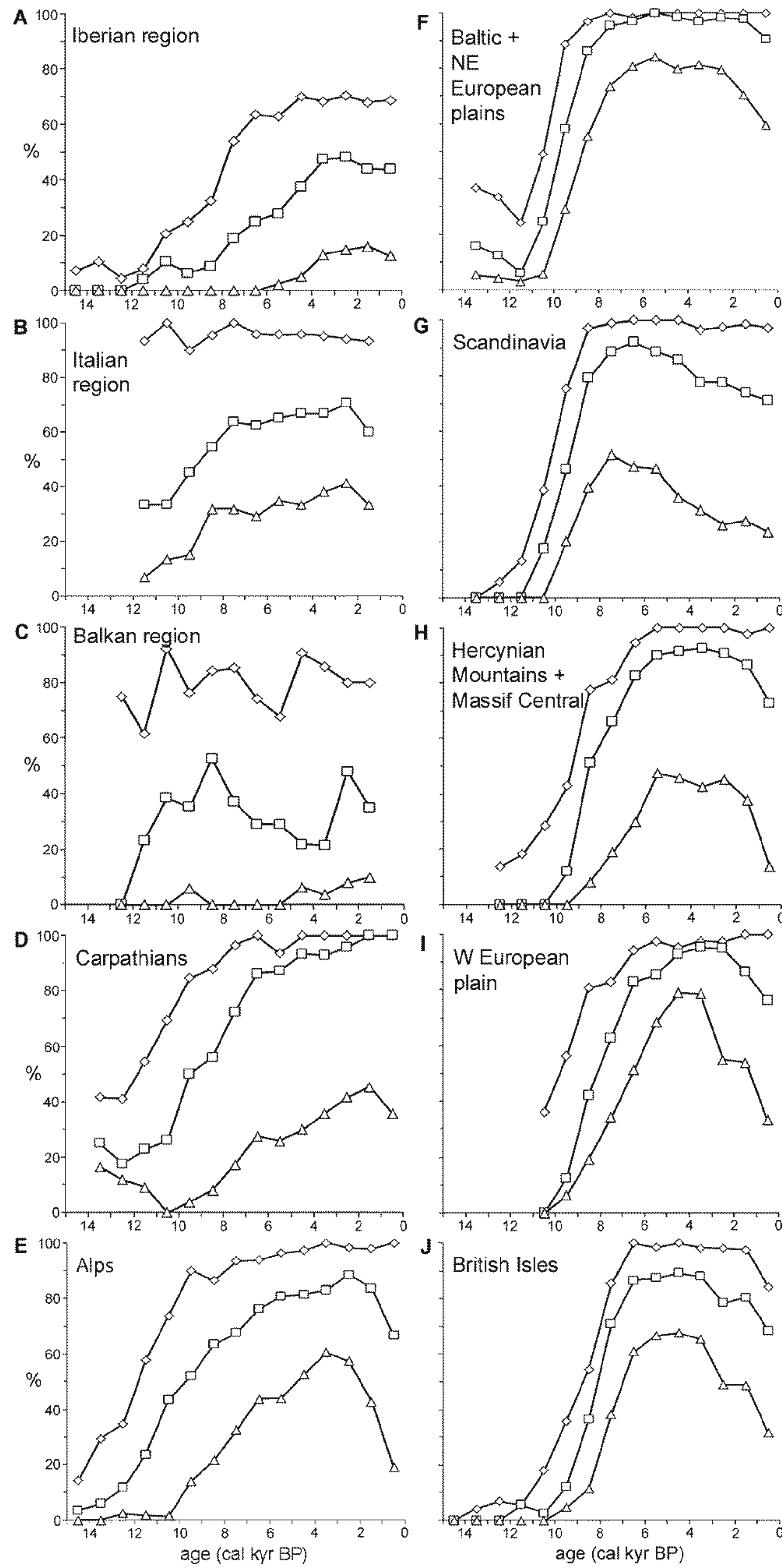

A
Iberian region
B
Italian region
C
Balkan region
D
Carpathians
E
Alps
F
Baltic + NE European plains
G
Scandinavia
H
Hercynian Mountains + Massif Central
I
W European plain
J
British Isles
%
age (cal kyr BP)

Figure 5. Regional proportions of *Alnus* pollen sites during the Late Glacial and Holocene periods. Pollen thresholds: 0.5% (diamonds), 2.5% (squares) and 10% (triangles). Only time intervals with more than 10 sites available in particular regions were considered.
doi:10.1371/journal.pone.0088709.g005

At the scale of hundreds to a thousand kilometres within individual regions, we recorded little or no directional pattern in the *Alnus* expansion, i.e., sites with *Alnus* evidence initially occurred across the whole region, and then the number of sites increased equally. We recorded this pattern in all northern regions, including the deglaciated area of Scandinavia, corroborating the descriptions of Bennett and Birks [95] for the British Isles and Giesecke et al. [18] for the Baltic area. Such a general absence of spatial coherence of the *Alnus* expansion within large areas seems to be very specific in comparison with the generally observed "stepping stone" character of expansions commonly recorded for other European trees [95]. This pattern suggests that the delayed *Alnus* expansion apparent in some regions was likely a result of environmental limitations rather than the effect of slow colonization.

Giesecke et al. [18] suggested that the climate is an important factor affecting regional differences in the expansion of *Alnus*. Global warming is generally assumed to be a trigger of the rapid *Alnus* expansion that began at the turn of the Late Glacial and Holocene periods [1]. However, an arid climate in some regions could have limited the onset of the *Alnus* expansion. The ecological requirements of *Alnus* and their recent distribution indicate that *Alnus* occurrence significantly declines in areas with an arid climate [37]. *Alnus glutinosa* is currently absent from large, arid areas of the Hungarian, Romanian and Ukrainian lowlands and the Iberian peninsula (http://euforgen.org). Increased oceanicity and rising sea levels after the separation of the British Isles from the continent possibly drove the *Alnus* expansion at approximately 9 cal. kyr BP in the British Isles, as suggested by Godwin [96] and Chambers and Elliott [94]. Similarly, the early *Alnus* expansion in the Baltic area could be accelerated by the large area of the Ancylus Lake (i.e., the Baltic sea).

Alnus Expansion in Southern Peninsulas, the Alps and the Carpathians

In southern regions, *Alnus* began its expansion in the Late Glacial and early Holocene periods. It seems that *Alnus* expanded

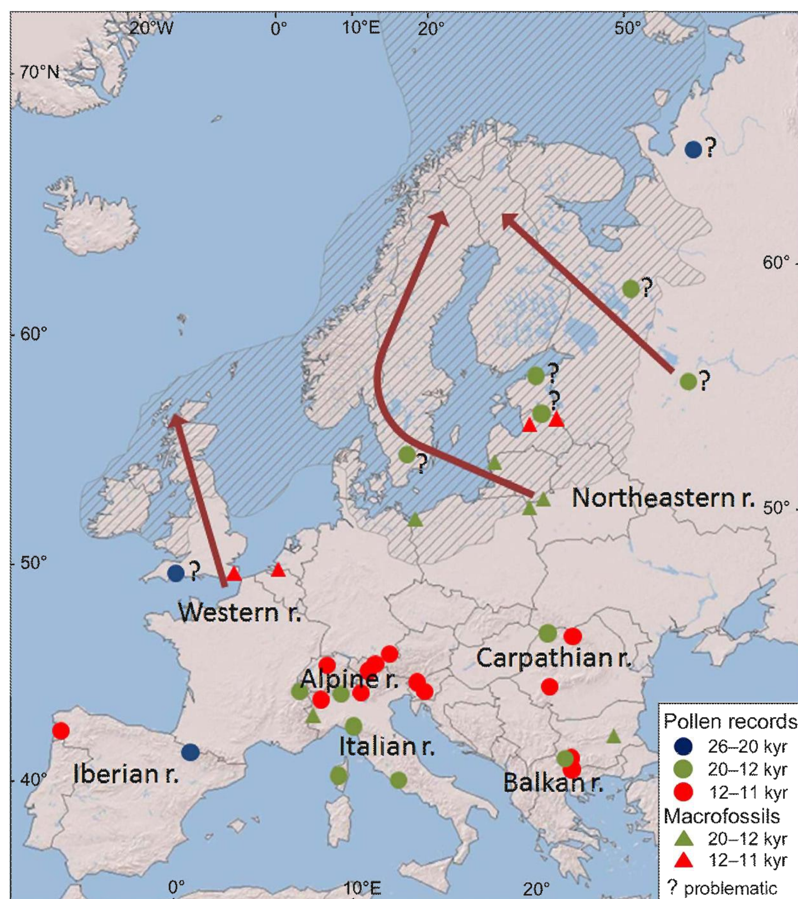


Figure 6. Putative Last Glacial Maximum refugia and directions of postglacial *Alnus* migration. The triangles and dots indicate macrofossil and pollen (≥2.5%) records from the LGM (blue), Late Glacial (green) and early Holocene (red). Arrows indicate directions of *Alnus* migration after northern deglaciation; question marks show problematic pollen records – possible reworking or long-distance pollen dispersal; hatching indicates the maximal extent of the ice sheet during the LGM.
doi:10.1371/journal.pone.0088709.g006

at individual sites rather than synchronously in whole regions. We assume that the arid climate of the Mediterranean, which was temporarily and spatially variable during the Holocene [97], possibly limited the establishment of new populations and locally caused population decreases. Similarly, a harsh, unstable mountain climate [98], [99] possibly drove a relatively slow expansion in the Alps and Carpathians.

During the second part of the Holocene, between 6 and 0 cal. kyr BP, *Alnus* retreats took place in most regions of Europe (Figure 5). In Scandinavia, the northward-southward direction of its population decrease is positively correlated with climate cooling and ombrogenous peat formation, which are likely the main factors initiating this process [37], [39]. Human activity in floodplains resulting in deforestation could be an additional factor contributing to thinning [100].

Species-specific History of *Alnus glutinosa* and *A. incana* Based on Macrofossils

Differences in the LGM refugia of *Alnus glutinosa* and *A. incana* could have significantly affected the time of the *Alnus* expansion in particular regions. For example, the earlier *Alnus* expansion in the Baltic area could be an expansion of the more cold-tolerant *A. incana*. Available macrofossils, however, do not support such differences between *A. glutinosa* and *A. incana*, although results may be influenced by their relative scarcity. It seems that both *Alnus* species colonised Scandinavia from the area of the northeastern refugium. Macrofossil evidence of *A. incana* from Netherlands also supports its Late Glacial occurrence in western Europe, i.e., outside its recent range [79] (Figure 3C).

Drawbacks of the Approach

Different factors may influence the proportion of *Alnus* pollen at individual sites and potentially underestimate or overestimate *Alnus* occurrence in the past. The recorded pollen proportion of species depends on pollen production and dispersal of other species in the vegetation [49]. It has been shown that the occurrence of trees (e.g. *Betula*) in areas with low pollen production, such as borders of tundra and taiga, may be overestimated in comparison to forest zones [49]. Recent studies have shown that the size of sedimentary basins, including bogs and lakes, importantly influences the source area of pollen coming from surrounding vegetation [101], [102]. Small sedimentary basins reflect the composition of surrounding vegetation at the expense of regional vegetation patterns and, thus, may underestimate regional species occurrence [101], [102].

One important factor influencing the representation of *Alnus* pollen is its dispersal ability. *Alnus* has small and light pollen grains (fall speed 0.021 ms^{-1}, according to Eisenhut [103]) effectively dispersed by wind over large distances. Studies have shown that *Alnus* pollen may occur in quite high relative quantities (4%) in remote areas thousands of kilometres from its closest occurrence in the vegetation [104]. An unstable climate and strong winds in the last glacial period likely facilitated long-distance dispersal of *Alnus* pollen, biasing pollen records in generally treeless landscapes with low pollen production.

The level of taxonomic resolution of the pollen spectra may bias interpretations when considering occurrences of *Alnus glutinosa* and *A. incana*. May and Lacourse [105] pointed out problems with the identification of three species, *A. rubra* (analogous to *A. glutinosa*), *A. incana* and *A. alnobetula*, in pollen spectra based on a dataset from North America. They concluded that if all three species were present in the vegetation, it would be statistically impossible to determine their pollen at the species level. This makes it difficult to distinguish *A. alnobetula* pollen from the other two and complicates the interpretation of pollen records. In southern Italy and Corsica, *Alnus* pollen records may also include pollen grains of *A. cordata*, which grows there sympatrically with *A. glutinosa* in alluvial habitats. Similarly, we cannot fully exclude the presence of pollen transported over long distances belonging to other species of subgenus *Alnus* such as *A. djavanshirii* Zare, *A. dolichocarpa* Zare, *A. orientalis* Decne. and *A. subcordata* C. A. Mey, all recently growing in the Eastern Mediterranean area and Iran [106], [107]. Despite the above-mentioned facts, the accordance of macrofossils with the pollen records confirms the robustness of the relative pollen data used in this study.

Comparison with Huntley and Birks, and King and Ferris

Using a much larger pollen dataset, we broadly confirmed LGM refugial areas and the general pattern of the postglacial expansion of *Alnus* as presented in the Huntley and Birks [1] "Pollen Maps", thus supporting the robustness and actuality of their work. The main differences between our study and the conclusions of Huntley and Birks [1] lie in the interpretation of the importance of northern LGM refugial areas for the *Alnus* expansion. Based on our dataset, the refugium in northeastern Europe appears to be more important for the *Alnus* expansion than was proposed by Huntley and Birks [1]. Huntley and Birks [1] mentioned this area only as a possible LGM refugium of *A. incana* subsp. *kolaensis*. We also support that the western refugium rather than eastern European one was the source for the expansion in the British Isles.

King and Ferris [34] have suggested the Carpathians as possible source areas for the expansion of *Alnus glutinosa* in the northern part of Europe. Our study also supports northeastern and western Europe. However, some conclusions of King and Ferris [34] seems to be based on the work of Huntley and Birks [1] rather than on molecular data. Only two largely distributed haplotypes, the first occurring across all northern parts of Europe and the second in the Alps, the Carpathians, western Europe and Scandinavia, were recorded by King and Ferris [34]. The presence of two weakly spatially structured haplotypes in the northern part of Europe may reflect the postglacial expansion of genotypes from the Carpathians [34] but may also correspond to the fragmentation of the continual *A. glutinosa* range during cold phases of the last glacial period. Similarly, some tree species most likely surviving the last glacial period in the northern part of Europe, such as *Betula pendula*, *B. pubescens*, *Populus tremula* and *Salix caprea*, exhibit a low level of phylogeographic structure [108], [109], [110]. To shed light on the last glacial period and Holocene history of *A. glutinosa* in northern Europe, future molecular studies should combine several approaches. For example, more variable chloroplast DNA markers [111] and microsatellites capable of determining the demographic history of *A. glutinosa* [112] in a particular region using approximate Bayesian computation [113] could be employed. A similar study is needed for *A. incana*, for which molecular studies are still missing.

Huntley and Birks [1] postulated two questions concerning the expansion pattern of *Alnus* in Europe. First, they asked why *Alnus* delayed its expansion north of the Alps. They hypothesised that this delay could have been caused by the occurrence of only cold-demanding *A. incana* and *A. alnobetula* in the Alpine LGM refugium. These species were unable to colonise the upland and lowland areas north of the Alps. This answer remains plausible, but the macrofossil finding of *A. glutinosa* in the southern foothills of the Alps in the Late Glacial period makes their interpretations less probable. Second, they posed a question about the importance of a western refugium for the *Alnus* expansion, which appears to be the source for the *Alnus* expansion in the British Isles in our study. However, only future phylogeographic studies can bring progress

towards answering the following additional questions: (i) Are there any distinctions among northern LGM refugial areas of *A. glutinosa* and *A. incana* that could influence regional differences at the beginning of the *Alnus* expansion? (ii) Was Scandinavia colonised only from the northeastern refugium, or were there other sources of colonisation located, for example, in western Europe? (iii) What is the origin of *A. incana* subsp. *kolaensis*, whose range has recently been limited to the north of Scandinavia? Within this context, the large area of northwestern Russia and the Baltic states appears to be crucial for future molecular sampling.

Supporting Information

Figure S1 Holocene distribution (6–2 cal. kyr BP) of *Alnus* pollen sites. According to four classes of percentage of *Alnus* pollen and macrofossil remains. The colour of dots indicates changes compared to the previous period; red, expansion, *Alnus* pollen < 2.5% in preceding period; blue, retreat, *Alnus* pollen ≥2.5% in preceding period; orange, new pollen sites of *Alnus* pollen ≥2.5%, respectively; black, stability; the course of deglaciation (white) and changes in coastline (dot lines).
(DOCX)

Figure S2 Holocene distribution (2–0 cal. kyr BP) of *Alnus* pollen sites. According to four classes of percentage of *Alnus* pollen and macrofossil remains; for details see Figure S1.
(DOCX)

Table S1 Location of the pollen sites from EPD, PALYCZ and the literature (Lit.).
(DOCX)

Table S2 Location of the macrofossil sites from NEMD and the literature (Lit.).
(DOCX)

Table S3 References of the pollen and macrofossil sites from EPD, PALYCZ, NEMD and the literature (Lit.).
(DOCX)

Checklist S1 PRISMA checklist.
(DOCX)

Acknowledgments

We wish to thank all contributors to EPD, NEMD and PALYCZ for making their data available and, thus, making this study possible. We would like to thank F. Krahulec and H. Svobodová-Svitavská for insightful comments on previous drafts of this manuscript and American Journal Experts and Frederic Rooks for improving the English of this paper.

Author Contributions

Conceived and designed the experiments: J Douda J Doudová AD BM. Analyzed the data: J Douda J Doudová AD PK VH PZ KK. Wrote the paper: J Douda J Doudová AD PK VH PZ KK BM.

References

1. Huntley B, Birks HJB (1983) An atlas of past and present pollen maps for Europe: 0–13000 years ago. Cambridge: Cambridge University Press. 688 p.
2. Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15: 4261–4293.
3. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, et al. (2009) The Last Glacial Maximum. Science 325: 710–714.
4. Bennett K, Tzedakis P, Willis K (1991) Quaternary refugia of north European trees. J Biogeogr 18: 103–115.
5. Tzedakis PC (1994) Hierarchical biostratigraphical classification of long pollen sequences. J Quaternary Sci 9: 257–259.
6. Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, et al. (2006) A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. New Phytol 171: 199–221.
7. Svenning J-Ch, Normand S, Kageyama M (2008) Glacial refugia of temperate trees in Europe: insights from species distribution modelling. J Ecol 96: 1117–1127.
8. Litynska-Zajac M (1995) Anthracological analysis. In: Hromada J, Kozlowski J, editors. Complex of upper Palaeolithic sites near Moravany, western Slovakia. Krakow: Jagellonian University Press. 74–79.
9. Willis KJ, Rudner E, Sumegi P (2000) The full-glacial forests of central and south-eastern Europe. Quaternary Res 53: 203–213.
10. Willis KJ, van Andel TH (2004) Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. Quat Sci Rev 23: 2369–2387.
11. Jankovská V, Pokorný P (2008) Forest vegetation of the last full-glacial period in the Western Carpathians (Slovakia and Czech Republic). Preslia 80: 307–324.
12. Kuneš P, Pelánková B, Chytrý M, Jankovská V, Pokorný P, et al. (2008) Interpretation of the last-glacial vegetation of eastern-central Europe using modern analogues from southern Siberia. J Biogeogr 35: 2223–2236.
13. Binney HA, Willis KJ, Edwards ME, Bhagwat SA, Anderson PM, et al. (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. Quat Sci Rev 28: 2445–2464.
14. Kullman L (1998) Non-analogous tree flora in the Scandes Mountains, Sweden, during the early Holocene - macrofossil evidence of rapid geographic spread and response to palaeoclimate. Boreas 27: 153–161.
15. Väliranta M, Kaakinen A, Kuhry P, Kultti S, Salonen JS, et al. (2011) Scattered late-glacial and early Holocene tree populations as dispersal nuclei for forest development in north-eastern European Russia. J Biogeogr 38: 922–932.
16. Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, et al. (2012) Glacial survival of boreal trees in northern Scandinavia. Science 335: 1083–1086.
17. Giesecke T, Hickler T, Kunkel T, Sykes MT, Bradshaw RHW (2007) Towards an understanding of the Holocene distribution of *Fagus sylvatica* L. J Biogeogr 34: 118–131.
18. Giesecke T, Bennett KD, Birks HJB, Bjune AE, Bozilova E, et al. (2011) The pace of Holocene vegetation change – testing for synchronous developments. Quat Sci Rev 30: 2805–2814.
19. Henne PD, Elkin CM, Reineking B, Bugmann H, Tinner W (2011) Did soil development limit spruce (*Picea abies*) expansion in the Central Alps during the Holocene? Testing a palaeobotanical hypothesis with a dynamic landscape model. J Biogeogr 38: 933–949.
20. Furlow JJ (1979) The systematics of the American species of *Alnus* (Betulaceae). Rhodora 81: 151–248.
21. Stevens PF (2001 onwards) Angiosperm Phylogeny Website. Version 12, July 2012 [and more or less continuously updated since]. Available: http://www.mobot.org/MOBOT/research/APweb/.
22. Chen ZD, Manchester SR, Sun HY (1999) Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology, and paleobotany. Am J Bot 86: 1168–1181.
23. Chen Z, Li J (2004) Phylogenetics and biogeography of *Alnus* (Betulaceae) inferred from sequences of nuclear ribosomal DNA ITS region. Int J Plant Sci 165: 325–335.
24. Jalas J, Suominen J (1976) Atlas Florae Europaeae. Distribution of vascular plants in Europe. 3. Salicaceae to Balanophoraceae. Helsinki: The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo. 128 p.
25. Rochet J, Moreau P-A, Manzi S, Gardes M (2011) Comparative phylogenies and host specialization in the alder ectomycorrhizal fungi *Alnicola*, *Alpova* and *Lactarius* (Basidiomycota) in Europe. BMC Evol Biol 11: 40.
26. Wijmstra TA (1969) Palynology of the first 30 metres of a 120 m deep section in Northern Greece. Acta Botanica Neerlandica 18: 511–527.
27. West RG (1980) Pleistocene forest history in East Anglia. New Phytol 85: 571–622.
28. Andersen ST (1965) Interglacialer og interstadialer i Danmarks Kvartaer. Meddelelser fra Dansk Geologisk Forening 15: 496–506.
29. Follieri M, Giardini M, Magri D, Sadori L (1998) Palynostratigraphy of the last glacial period in the volcanic region of Central Italy. Quatern Int 47–48: 3–20.
30. de Beaulieu J-L, Reille M (1984) A long Upper Pleistocene pollen record from Les Echets, near Lyon, France. Boreas 13: 111–132.
31. de Beaulieu J-L, Reille M (1992) The last climatic cycle at La Grande Pile (Vosges, France) a new pollen profile. Quat Sci Rev 11: 431–438.
32. Reille M, de Beaulieu J-L, Svobodova H, Andrieu-Ponel V, Goeury C (2000) Pollen analytical biostratigraphy of the last five climatic cycles from a long continental sequence from the Velay region (Massif Central, France). J Quaternary Sci 15: 665–685.
33. Müller UC (2000) A Late-Pleistocene pollen sequence from the Jammertal, south-western Germany with particular reference to location and altitude as factors determining Eemian forest composition. Veg Hist Archaeobot 9: 125–131.

34. King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. Mol Ecol 7: 1151–1161.
35. Douda J, Čejková A, Douda K, Kochánková J (2009) Development of alder carr after the abandonment of wet grasslands during the last 70 years. Ann For Sci 66: 712.
36. Douda J (2010) The role of landscape configuration in plant composition of floodplain forests across different physiographic areas. Journal of Vegetation Science 21: 1110–1124.
37. Tallantire PA (1974) The palaeohistory of the grey alder (*Alnus incana* (L.) Moench.) and black alder (*A. glutinosa* (L.) Gaertn.) in Fennoscandia. New Phytol 73: 529–546.
38. McVean DN (1953) Biological flora of the British Isles: *Alnus glutinosa* (L.) Gaertn. J Ecol 41: 447–466.
39. Kullman L (1992) The ecological status of grey alder (*Alnus incana* (L.) Moench) in the upper subalpine birch forest of the central Scandes. New Phytol 120: 445–451.
40. Moher D, Liberati A, Tetzlaff J, Altman DG (2010) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 8: 336–341.
41. Fyfe RM, de Beaulieu J-L, Binney H, Bradshaw RHW, Brewer S, et al. (2009) The European Pollen Database: past efforts and current activities. Veg Hist Archaeobot 18: 417–424.
42. Kuneš P, Abraham V, Kovařík O, Kopecký M, Contributors P (2009) Czech Quaternary Palynological Database – PALYCZ: review and basic statistics of the data. Preslia 81: 209–238.
43. Giesecke T, Davis B, Brewer S, Finsinger W, Wolters S, et al. (2013) Towards mapping the late Quaternary vegetation change of Europe. Veg Hist Archaeobot. doi: 10.1007/s00334-012-0390-y.
44. Blaauw M (2010) Methods and code for "classical" age-modelling of radiocarbon sequences. Quat Geochronol 5: 512–518.
45. R Development Core Team (2012) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
46. Punt W, Blackmore S, Hoen PP, Stafford PJ (2003) The Northwest European Pollen Flora, Volume 8. Elsevier. 194 p.
47. Leopold EB, Birkebak J, Reinink-Smith L, Jayachandar AP, Narváez P, et al. (2012) Pollen morphology of the three subgenera of *Alnus*. Palynology 36: 131–151.
48. Douda J (2008) Formalized classification of the vegetation of alder carr and floodplain forests in the Czech Republic. Preslia 80: 199–224.
49. Lisitsyna OV, Giesecke T, Hicks S (2011) Exploring pollen percentage threshold values as an indication for the regional presence of major European trees. Rev Palaeobot Palynol 166: 311–324.
50. Tallantire PA (1992) The alder [*Alnus glutinosa* (L.) Gaertn.] problem in the British Isles: a third approach to its palaeohistory. New Phytol 122: 717–731.
51. Montanari C (1996) Recent pollen deposition in alder woods and in other riverine plant comunities. Allionia 34: 309–323.
52. Ralska-Jasiewiczowa M, Latałowa M, Wasylikowa K, Tobolski K, Madeyska E, et al. (2004) Late Glacial and Holocene history of vegetation in Poland based on isopollen maps. Cracow, Poland: W. Szafer Institute of Botany, Polish Academy of Sciences. 444 p.
53. González-Sampériz P, Valero-Garcés BL, Carrión JS, Peña-Monné JL, García-Ruiz JM, et al. (2005) Glacial and Lateglacial vegetation in northeastern Spain: New data and a review. Quatern Int 140–141: 4–20.
54. Watts W, Allen JRM, Huntley B (1996) Vegetation history and palaeoclimate of the last glacial period of Lago Grande di Monticchio, southern Italy. Quat Sci Rev 15: 133–153.
55. Lucchi M (2008) Vegetation dynamics during the Last Interglacial–Glacial cycle in the Arno coastal plain (Tuscany, western Italy): location of a new tree refuge. Quat Sci Rev 27: 2456–2466.
56. Pini R, Ravazzi C, Donegana M (2009) Pollen stratigraphy, vegetation and climate history of the last 215ka in the Azzano Decimo core (plain of Friuli, north-eastern Italy). Quat Sci Rev 28: 1268–1290.
57. Pini R, Ravazzi C, Reimer PJ (2010) The vegetation and climate history of the last glacial cycle in a new pollen record from Lake Fimon (southern Alpine foreland, N-Italy). Quat Sci Rev 29: 3115–3137.
58. Barbier D (1999) Histoire de la vegetation du nord-mayennais de la fin du Weichselien à l'aube du XXIème siècle: mise en évidence d'un Tardiglaciaire armoricain; interactions homme-milieu. Univ. de Nantes (Thesis).
59. Aleshinskaya ZV, Gunova VS (1976) History of Lake Nero as reflection on the surrounding landscape dynamics. In: Kalinin GP, Klige RK, editors. Problemy Paleohydrologii. Moscow: Nauka. 214–222.
60. Zarrina YP, Spiridonova YA, Arslanov KA, Kolesnikova TD, Simonova GF (1973) New profile of Middle Valday deposits near the village Shenskoe (Mologo-Sheksna basin). In: Subakov VA, editor. Pleistocene Chronology and Climatic Stratigraphy. Leningrad: Nauka Press, Lenigrad Division. 160–167.
61. Voznyachuk LN, Valchik MA (1978) Morphology, structure and history of the valley of the Neman in Pleistocene and Holocene. Minsk: Nauka i Tekhnika.
62. Gaigalas AI, Dvaretskas VV, Banis YY, Davaynis GA, Kibilda ZA, et al. (1981) Radiocarbon age of river terraces in the southern Baltic. Isotopic and geochemical methods in biology, geology and archeology. Tartu. 28–32.
63. Kelly A, Charman DJ, Newnham RM (2010) A Last Glacial Maximum pollen record from Bodmin Moor showing a possible cryptic northern refugium in southwest England. J Quaternary Sci 25: 296–308.
64. Paus A, Svendsen JI, Matiouchkov A (2003) Late Weichselian (Valdaian) and Holocene vegetation and environmental history of the northern Timan Ridge, European Arctic Russia. Quat Sci Rev 22: 2285–2302.
65. Tonkov S, Possnert G, Bozilova E (2006) The lateglacial vegetation and radiocarbon dating of Lake Trilistnika, Rila Mountains (Bulgaria). Veg Hist Archaeobot 16: 15–22.
66. Finsinger W, Tinner W, van der Knaap WO, Ammann B (2006) The expansion of hazel (*Corylus avellana* L.) in the southern Alps: a key for understanding its early Holocene history in Europe? Quat Sci Rev 25: 612–631.
67. Magyari EK, Chapman JC, Gaydarska B, Marinova E, Deli T, et al. (2008) The "oriental" component of the Balkan flora: evidence of presence on the Thracian Plain during the Weichselian late-glacial. J Biogeogr 35: 865–883.
68. Reille M, Gamisans J, de Beaulieu J-L, Andrieu V (1997) The late-glacial at Lac de Creno (Corsica, France): a key site in the western Mediterranean basin. New Phytol 135: 547–559.
69. Lowe JJ, Watson C (1993) Lateglacial and early Holocene pollen stratigraphy of the northern Apennines, Italy. Quat Sci Rev 12: 727–738.
70. Kelly MG, Huntley B (1991) An 11 000-year record of vegetation and environment from Lago di Martignano, Latium, Italy. J Quaternary Sci 6: 209–224.
71. Björkman L, Feurdean A, Wohlfarth B (2003) Late-Glacial and Holocene forest dynamics at Steregoiu in the Gutaiului Mountains, Northwest Romania. Rev Palaeobot Palynol 124: 79–111.
72. Björck S, Möller P (1987) Late Weichselian environmental history in southeastern Sweden during the deglaciation of the Scandinavian ice sheet. Quaternary Res 28: 1–37.
73. Saarse L, Niinemets E, Amon L, Heinsalu A, Veski S, et al. (2009) Development of the late glacial Baltic basin and the succession of vegetation cover as revealed at Palaeolake Haljala, northern Estonia. Estonian Journal of Earth Sciences 58: 317–333.
74. Amon L, Veski S, Heinsalu A, Saarse L (2012) Timing of Lateglacial vegetation dynamics and respective palaeoenvironmental conditions in southern Estonia: evidence from the sediment record of Lake Nakri. J Quaternary Sci 27: 169–180.
75. Wohlfarth B, Filimonova L, Bennike O, Björkman L, Brunnberg L, et al. (2002) Late-Glacial and Early Holocene Environmental and Climatic Change at Lake Tambichozero, Southeastern Russian Karelia. Quaternary Res 58: 261–272.
76. Wohlfarth B, Tarasov P, Bennike O, Lacourse T, Subetto D, et al. (2006) Late Glacial and Holocene Palaeoenvironmental Changes in the Rostov-Yaroslavl' Area, West Central Russia. J Paleolimnol 35: 543–569.
77. Latałowa M, Borówka RK (2006) The Allerød/Younger Dryas transition in Wolin Island, northwest Poland, as reflected by pollen, macrofossils, and chemical content of an organic layer separating two aeolian series. Veg Hist Archaeobot 15: 321–331.
78. Stančikaitė M, Šinkūnas P, Šeirienė V, Kisielienė D (2008) Patterns and chronology of the Lateglacial environmental development at Pamerkiai and Kašučiai, Lithuania. Quat Sci Rev 27: 127–147.
79. Bos JAA, Huisman DJ, Kiden P, Hoek WZ, van Geel B (2005) Early Holocene environmental change in the Kreekrak area (Zeeland, SW-Netherlands): A multi-proxy analysis. Palaeogeogr Palaeoclimatol Palaeoecol 227: 259–289.
80. Waller MP, Marlow AD (1994) Flandrian vegetational history of south-eastern England. Stratigraphy of the Brede valley and pollen data from Brede Bridge. New Phytol 126: 369–392.
81. Farcas S, de Beaulieu J-L, Reille M, Coldea G, Diaconeasa B, et al. (1999) First 14C datings of Late Glacial and Holocene pollen sequences from Romanian Carpathes. Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie 322: 799–807.
82. Bozilova ED, Tonkov S, Pavlova D (1986) Pollen and plant macrofossil analyses of the Lake Sucho Ezero in the south Rila mountains. Annual of Sofia University, Faculty of Biology 80: 48–57.
83. Bozilova ED, Tonkov SB (2000) Pollen from Lake Sedmo Rilsko reveals southeast European postglacial vegetation in the highest mountain area of the Balkans. New Phytol 148: 315–325.
84. Tonkov S, Panovska H, Possnert G, Bozilova E (2002) The Holocene vegetation history of Northern Pirin Mountain, southwestern Bulgaria: pollen analysis and radiocarbon dating of a core from Lake Ribno Banderishko. Holocene 12: 201–210.
85. López-Merino L, Silva Sánchez N, Kaal J, López-Sáez JA, Martínez Cortizas A (2012) Post-disturbance vegetation dynamics during the Late Pleistocene and the Holocene: An example from NW Iberia. Glob Planet Change 92–93: 58–70.
86. Seppä H (1996) Post-glacial dynamics of vegetation and tree-lines in the far north of Fennoscandia. Fennia 174: 1–96.
87. Carcaillet C, Hörnberg G, Zackrisson O (2012) Woody vegetation, fuel and fire track the melting of the Scandinavian ice-sheet before 9500 cal yr BP. Quaternary Res 78: 540–548.
88. Scourse J (2010) Comment: A Last Glacial Maximum pollen record from Bodmin Moor showing a possible cryptic northern refugium in southwest England. (Kelly et al., 2010). J Quaternary Sci 25: 826–827.
89. Lepais O, Muller SD, Ben Saad-Limam S, Benslama M, Rhazi L, et al. (2013) High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. PLoS ONE 8: e75029. doi: 10.1371/journal.pone.0075029.

90. Ben Tiba B, Reille M (1982) Recherches pollenanalytiques dans les montagnes de Kroumirie (Tunisie septentrionale): premiers résultats. Ecol Mediterr 8: 75–86.
91. Stambouli-Essassi S, Roche E, Bouzid S (2007) Evolution of vegetation and cimatic changes in North-Western Tunisia during the last 40 millennia. Internat J Trop Geol Geogr Ecol 31: 171–214.
92. Chambers FM, Price S-M (1985) Palaeoecology of *Alnus* (alder): early post-glacial rise in a valley Mire, north-west Wales. New Phytol 101: 333–344.
93. Bush MB, Hall AR (1987) Flandrian *Alnus*: expansion or immigration? J Biogeogr 14: 479–481.
94. Chambers FM, Elliott L (1989) Spread and expansion of *Alnus* Mill in the British Isles: timing, agencies and possible vectors. J Biogeogr 16: 541–550.
95. Bennett KD, Birks HJB (1990) Postglacial history of alder (*Alnus glutinosa* (L.) Gaertn.) in the British Isles. J Quaternary Sci 5: 123–133.
96. Godwin H (1975) History of the British flora. 2nd ed. Cambridge: Cambridge University Press. 580 p.
97. Magny M, Vannière B, Calo C, Millet L, Leroux A, et al. (2011) Holocene hydrological changes in south-western Mediterranean as recorded by lake-level fluctuations at Lago Preola, a coastal lake in southern Sicily, Italy. Quat Sci Rev 30: 2459–2475.
98. Wick L, Tinner W (1997) Vegetation changes and timberline fluctuations in the Central Alps as indicators of Holocene climatic oscillations. Arct Antarct Alp Res 29: 445–458.
99. Magyari EK, Jakab G, Bálint M, Kern Z, Buczkó K, et al. (2012) Rapid vegetation response to Lateglacial and early Holocene climatic fluctuation in the South Carpathian Mountains (Romania). Quat Sci Rev 35: 116–130.
100. Brown A (1988) The palaeoecology of *Alnus* (alder) and the postglacial history of floodplain vegetation. Pollen percentage and influx data from the West Midlands, United Kingdom. New Phytol 110: 425–436.
101. Sugita S (1994) Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. Journal of Ecology 82: 881–897.
102. Gaillard M-J, Sugita S, Bunting MJ, Middleton R, Broström A, et al. (2008) The use of modelling and simulation approach in reconstructing past landscapes from fossil pollen data: a review and results from the POLLANDCAL network. Veg Hist Archaeobot 17: 419–443.
103. Eisenhut G (1961) Untersuchungen über die Morphologie und Ökologie der Pollenkörner heimischer und fremdländischer Waldbäume (translated into English by Jackson TS and Jaumann P, 1989). Hamburg: Paul Parey. 68 p.
104. van der Knaap W, van Leeuwen JF, Froyd CA, Willis KJ (2012) Detecting the provenance of Galapagos non-native pollen: The role of humans and air currents as transport mechanisms. Holocene 22: 1373–1383.
105. May L, Lacourse T (2012) Morphological differentiation of *Alnus* (alder) pollen from western North America. Rev Palaeobot Palynol 180: 15–24.
106. Yaltrik Y (1982) Betulaceae. In: Davis PH, editor. Flora of Turkey, Vol. 7. Edinburgh: Edinburgh University Press. 688–694.
107. Zare H, Amini T (2012) A review of the genus *Alnus* Gaertn. in Iran, new records and new species. Iran J Bot 18: 10–21.
108. Palmé AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. Heredity 91: 465–474.
109. Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, et al. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300: 1563–1565.
110. Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, Lascoux M (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. J Biogeogr 34: 1601–1610.
111. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, et al. (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92: 142–166.
112. Lepais O, Bacles CFE (2011) De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. J Hered 102: 627–632.
113. Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. Trends Ecol Evol 25: 410–418.

Tree Genetics & Genomes (2014) 10:865–873
DOI 10.1007/s11295-014-0727-z

ORIGINAL PAPER

Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*

Alena Drašnarová · Karol Krak · Petr Vít · Jana Doudová · Jan Douda · Věroslava Hadincová · Petr Zákravský · Bohumil Mandák

Received: 25 November 2013 / Revised: 14 February 2014 / Accepted: 18 March 2014 / Published online: 1 April 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract We investigated 39 previously developed *Betula*, *Alnus*, and *Corylus* simple sequence repeat (SSR) markers for their utility in the cross-generic amplification of two European alder species, i.e., *Alnus glutinosa* and *A. incana*. Of these markers, ten loci had successful amplification within *Alnus* species. Finally, we designed two multiplexes composed of eight and nine loci for *A. glutinosa* and *A. incana*, respectively. Multiplexes were tested on 100 samples from five different populations of each species across Europe. The majority of loci had a relatively high genetic diversity, were in Hardy–Weinberg equilibrium, and showed low error rates and low occurrence of null alleles. By comparing sequences of source species and both *Alnus* species, we concluded that repeat motifs of five of these ten loci differed from those described for the source species. These differences represent mainly the modifications of the original motifs and affected compound or interrupted repeats as well as pure ones. The repeat motifs of three loci of the two alder species also differed. These mutations could lead to erroneous estimates of allele homology, because alleles with identical lengths will not have the same number of repeat units. Hence, before using microsatellite markers in studies comparing two or more species, they should be carefully examined and sequenced to ensure that allele homology is really stable and not affected by various inserts that change the sequence.

Keywords Alder · Betulaceae · Microsatellites · DNA sequencing

Communicated by P. Ingvarsson

Electronic supplementary material The online version of this article (doi:10.1007/s11295-014-0727-z) contains supplementary material, which is available to authorized users.

A. Drašnarová (✉) · K. Krak · P. Vít · J. Doudová · J. Douda · V. Hadincová · P. Zákravský · B. Mandák
Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, 252 43 Průhonice, Czech Republic
e-mail: drasnarova@fzp.czu.cz

A. Drašnarová · J. Douda · B. Mandák
Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 21 Praha 6 – Suchdol, Czech Republic

Introduction

Alnus glutinosa and *A. incana* (family Betulaceae) are common species of European riparian and water-logged habitats (McVean 1953; Tallantire 1974). *A. glutinosa* ranges across lowlands and middle altitudes of the whole of Europe, except the extreme north. It extends as far as West Siberia and the mountains of West Turkey (eastwards different subspecies have been recorded, Yatrik 1982) and North Africa (McVean 1953). The distribution range of *A. incana* is divided into a northern part covering the entire Fennoscandia, which continues eastwards across European Russia to the western edge of its Asian part and a southern range with a patchy mountain occurrence in the Alps, the Hercynian Mountains, the Carpathians, and the Dynaric Alps. The two ranges overlap in a zone from southern France to western Russia and in southern Fennoscandia.

Compared to important timber trees such as *Picea abies* and the *Quercus* species, distribution ranges of these alders have not been much extended by human activities (Claessens et al. 2010); thus, there are plenty of non-planted populations, especially on waterlogged sites. These populations are ideal for studies of population genetic structure or phylogeographic patterns of both species as they represent stands with low human impact.

Such information could answer important questions regarding population genetic structure and past or present

connectivity of river systems and could explain recent or historical migrations of a given species. There is, however, a lack of molecular markers currently available for both *Alnus* species. Thanks to King and Ferris (1998), we have information about the postglacial history of *A. glutinosa* based on chloroplast DNA (cpDNA) diversity supporting previous fossil pollen data (Huntley and Birks 1983). Further, Cox et al. (2011) used a genome scan approach based on amplified fragment length polymorphism (AFLP) markers to detect a signature of adaptation to natural selection in populations of *A. glutinosa*.

Recently, 12 microsatellite markers (SSRs) were developed for *A. glutinosa* and co-amplified in a single multiplex PCR by Lepais and Bacles (2011) (ten of these also amplified successfully in *A. incana*). Moreover, 15 SSRs markers previously developed for birch (Kulju et al. 2004) were tested on *A. glutinosa* and *A. incana*, and eight of them amplified clearly scorable products (Zhuk et al. 2008). Zhuk et al. (2008) pointed out that these markers should be used with caution for population genetic studies due to the likely presence of null alleles. Further, SSR markers had been developed for the Betulaceae: *Betula pendula* (Kulju et al. 2004; Gürcan and Mehlenbacher 2010), *Betula maximowicziana* (Ogyu et al. 2003; Tsuda et al. 2009a, b), *Betula platyphylla* (Wu et al. 2002), *Betula pubescens* (Truong et al. 2005), *Alnus maritima* (Lance et al. 2009), *A. glutinosa* (Gürcan and Mehlenbacher 2010), and *Corylus avellana* (Gürcan et al. 2010). However, their amplification on both *A. glutinosa* and *A. incana* has not been tested.

Although several microsatellite markers have been published, more are still required. The limited number of loci available might reduce our ability to detect genetic diversity and heterozygosity (Selkoe and Toonen 2006) through the whole range of alder species. Modern statistical approaches based on Bayesian computations usually require multiple loci (>20; Wilson and Rannala 2003; Pearse and Crandall 2004) because the use of many loci can compensate for heterogeneity in mutational properties among them (Ellegren 2004).

In the present work, our aim was to increase the number of microsatellite markers for *A. glutinosa* and *A. incana* by cross-amplification of microsatellite markers previously developed for other members of Betulaceae. In order to obtain a more precise picture of population genetic diversity and the structure of alder populations in future studies, we selected markers with clear amplification profiles and sufficient polymorphism, paying particular attention to their usability and quality control. Furthermore, we sequenced these microsatellite loci in both species in order to determine and compare their repeat motifs. Specifically, we have asked whether mutations in both the SSR region and the flanking region might be responsible for variation in allele size among species.

Material and methods

Plant material

Samples of both *Alnus* species were collected from different locations across Europe in 2011 (see Table 1). Leaves from 235 trees were stored in silica gel. Individual samples were collected along transect with a pairwise distance of at least 50 m between trees.

DNA isolation

DNA from 35 individuals, for testing the microsatellite cross-amplification and optimization of multiplexes, was isolated as described in Štorchová et al. (2000) with a few modifications. In the first extraction step, 1.3 ml of extraction buffer and mercaptoethanol (1,000:1) was used and 100 μl of ethylenediaminetetraacetic acid (EDTA, 0.5 M, pH=8), and one half of a microspoon of polyvinylpolypyrrolidone (PVPP) was added to prevent DNA breakdown and to remove secondary metabolites. During the lysis step, 0.2 mg/ml of RNase (Fermentas, St. Leon, Germany) was added. For the DNA isolation of further ten populations, disruption of plant material was carried out using a TissueLyser II (Qiagen) with 2 cycles of 1-min disruption at 20 Hz, and DNA was extracted using a DNeasy 96 Plant Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The quality and yield of isolated DNA was checked on a 1 % agarose gel, and then, the DNA concentration was evaluated with a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). Highly concentrated samples were then diluted by 20–25 ng/μl ddH_2O.

Multiplex optimization

Choice of SSR markers and simplex reactions

First, a literature search of available markers from various species in the Betulaceae was made. Thirty-nine appropriate markers developed for the following species were found: *B. pendula* (12; Kulju et al. 2004; Gürcan and Mehlenbacher 2010), *B. maximowicziana* (7; Ogyu et al. 2003; Tsuda et al. 2009a, b), *B. platyphylla* (5; Wu et al. 2002), *B. pubescens* (1; Truong et al. 2005), *A. maritima* (4; Lance et al. 2009), *A. glutinosa* (2; Gürcan and Mehlenbacher 2010), and *C. avellana* (8; Gürcan et al. 2010) (additional information is provided in Online Resource 1).

Markers were initially tested in simplex reactions on four alder DNA samples (two from *A. glutinosa* and two from *A. incana*; Table 1 (a), samples 1–4) to check alder DNA amplification. PCRs were carried out in a final volume of 10 μl using 1× Qiagen Multiplex PCR Master Mix (Qiagen), 0.3 μM forward and reverse primers, and 1 μl

Table 1 Geographic locations of alder populations used for different steps of multiplex PCR optimization, i.e., (a) cross-amplification, (b) multiplexing, and (c) verification of final multiplexes

Site name	Country	Latitude	Longitude	Species	*N*
(a) Cross-amplification					
1. Stvolínky	CZE	50.6327	14.421961	*Alnus glutinosa*	1
2. Zemplínské jestrabie	SVK	48.48953	21.789322	*Alnus glutinosa*	1
3. Divoká Šárka	CZE	50.10785	14.338128	*Alnus incana*	1
4. Tyrolia	ITA	46.67132	10.520433	*Alnus incana*	1
5. Tübingen	GER	48.5577	9.045913	*Alnus glutinosa*	1
(b) Multiplexing					
6. Černiš	CZE	49.00914	14.436994	*Alnus glutinosa*	5
7. Tagliamento	ITA	46.13553	12.949137	*Alnus glutinosa*	5
8. Busche	SPA	46.03578	11.987871	*Alnus glutinosa*	5
9. Fagaraš	ROM	45.69358	24.572972	*Alnus incana*	5
10. Roccheta	ITA	46.26612	11.06297	*Alnus incana*	5
11. Bobbio	ITA	44.60601	9.296819	*Alnus incana*	5
(c) Multiplexes verification					
12. Wöllstein	GER	48.90805	9.955931	*Alnus glutinosa*	20
13. Fernitz	AUS	46.97412	15.487067	*Alnus glutinosa*	20
14. Praid	ROM	46.56867	25.167343	*Alnus glutinosa*, *Alnus incana*	20, 20
15. Chomutov	CZE	48.90834	13.827566	*Alnus glutinosa*, *Alnus incana*	20, 20
16. Busche	ITA	46.03578	11.987871	*Alnus glutinosa*, *Alnus incana*	20, 20
17. Bruckberg	GER	48.49655	11.995551	*Alnus incana*	20
18. Tauglbach	AUS	47.65932	13.153648	*Alnus incana*	20

AUS Austria, *CZE* Czech Republic, *GER* Germany, *ITA* Italy, *ROM* Romania, *SVK* Slovakia, *SPA* Spain, *N* number of individuals used in each optimization step

template DNA (20–25 ng/μl). Reactions were run on a Master Cycler Pro (Eppendorf, Hamburg, Germany) or Verity (Applied Biosystems, Foster City, CA, USA) thermal cyclers using the following conditions: 15 min of denaturation at 95 °C, followed by 40 cycles at 94 °C for 30 s, with three different annealing temperatures (56, 58, and 60 °C) for 30 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min. Amplification products were checked on 2 % agarose gel.

Markers which yielded clear bands were labeled by one of the fluorescent dyes NED, FAM, VIC, or PET (Applied Biosystems) and tested again by PCR on two samples (one *A. glutinosa* and one *A. incana*; Table 1 (a), samples 4 and 5) with the same conditions as before. One microliter of PCR product (ten times diluted) was mixed with 0.1 μl of GeneScan-500 LIZ internal size standard (Applied Biosystems) and 12 μl of Hi-Di formamide (Applied Biosystems) and electrophoresed using an ABI PRISM 3130 Genetic Analyser (Applied Biosystems).

Multiplex PCR optimization

Multiplex PCR optimization was done in two steps: multiplex optimization and multiplex verification on a large set of individuals and localities. Selected markers were tested in multiplexes on 30 samples (15 from *A. glutinosa* and 15 from *A. incana*). To obtain a better estimation of individual locus variability, we tested samples from three geographically distinct populations of each species (Table 1 (b)). PCRs were performed in a total volume of 5 μl using 1× Qiagen Multiplex PCR Master Mix (Qiagen), 0.15 μM primers, and 1 μl of template DNA (20–25 ng/μl) using the same thermal cycling condition as described above, but annealing temperature 58 °C which evinced the best amplification success during simplex reactions. Fragment analysis was performed as described above. Low quality profiles, in which markers showed little or no amplification, excessive stuttering, or lacked polymorphism, were excluded, and two multiplexes were built, one for each alder species. In subsequent PCRs, amplification patterns were further checked, and primer concentrations were optimized.

To verify the usability of our multiplexes in *A. incana* and *A. glutinosa*, both multiplexes were further tested on samples from five different populations (20 individuals per each population) of each species across Europe (Table 1 (c)), and the primer concentrations were further optimized (see [C] in Tables 2 and 3).

Genotype scoring and analyses

Allele size was determined using GeneMarker version 2.4.0 (SoftGenetics, State College, PA, USA). For each marker, automatic allele calling was defined by bins with manual correction afterwards. An individual was declared null

Table 2 Characteristics of final multiplex for *Alnus glutinosa* based on 100 analysed samples

No	Dye	[C][a] (μl)	Size[b] (bp)	ER/allele	N_A	H_O	H_E	F_{IS}	$_B$
A2	FAM	0.05	142–144	0	2	0.010	0.010	0.0001	0.000
A7	PET	0.15	172–176	0	3	0.120	0.133	0.100	0.006
A10	NED	0.25	111–133	0.020	7	0.700	0.699	−0.001	−0.007
A22	VIC	0.5	158–176	0.063	9	0.780	0.817	0.045	0.007
A26	PET	0.15	342–378	0.021	16	0.840	0.871	0.035	−0.009
A35	NED	0.1	220–248	0	15	0.880	0.856	−0.028	−0.025
A37	PET	0.8	245–273	0	12	0.760	0.802	0.052	0.010
A38	PET	0.1	108–143	0	10	0.400	0.416	0.037	0.002
Mean	–	–	–	0.013	9.25	0.561	0.575	0.025	−0.002

ER error rate, N_A total number of alleles, H_O observed heterozygosity, H_E expected heterozygosity, F_{IS} inbreeding coefficient (all loci are in Hardy–Weinberg equilibrium), $_B$ null allele frequency averaged over all populations (Brookfield method)

[a] Final concentration of each primer

[b] Size range of each allele

(nonamplifying at a locus) and treated as missing data after at least two amplification failures. To detect discrete size variants, allele binning was based on raw size using Autobin (http://www4.bordeaux-aquitaine.inra.fr/biogeco/Ressources/Logiciels/Autobin; see Guichoux et al. 2011a). Evidence of stuttering, allele dropout, and presence of null allele for each locus were tested by Micro-Checker version 2.2.3 (Van Oosterhout et al. 2004). Null allele frequencies were calculated using the Brookfield 1 equation (Brookfield 1996). The error rate estimate was achieved by counting the mismatches (Pompanon et al. 2005) from a subset of 25 duplicated samples for each species (five samples from each population, 25 % of the complete dataset) and was expressed as an error rate per allele for each locus (number of incorrect alleles divided by total number of alleles).

Summary data for SSR loci including total allele number (N_A), observed heterozygosity (H_O), and expected heterozygosity (H_E) and Weir and Cockerham's (1984) parameter $f(F_{IS})$, a measure of inbreeding within populations, were calculated using FSTAT (Goudet 1995). Deviation from the Hardy–Weinberg equilibrium was determined on the basis of 10,000 permutations in FSTAT as well as linkage disequilibrium that was determined for each locus pair across all populations.

Sequence analysis of SSR loci in *A. glutinosa* and *A. incana*

Sequence analyses were performed on successfully cross-amplified loci, in order to identify the repeat motifs in the two alder species and compare the motifs with the original sequences from GenBank (if available). For this purpose, one

Table 3 Characteristics of final multiplex for *Alnus incana* based on 100 analysed samples

No	Dye	[C][a] (μl)	Size[b] (bp)	ER/allele	N_A	H_O	H_E	F_{IS}	$_B$
A2	FAM	0.1	142–144	0.020	2	0.040	0.197	0.798*	0.092
A6	FAM	0.5	194–202	0	7	0.580	0.550	−0.055	−0.039
A10	NED	0.25	117–126	0	6	0.647	0.702	0.080	0.000
A18	VIC	0.8	223–269	0.023	21	0.798	0.876	0.090*	0.009
A22	VIC	0.5	156–178	0.080	8	0.414	0.473	0.124*	0.029
A26	PET	0.15	345–374	0	16	0.798	0.822	0.030	−0.019
A35	NED	0.1	220–238	0	9	0.808	0.821	0.013	−0.015
A37	PET	0.25	245–257	0	7	0.234	0.284	0.180*	0.023
A38	PET	0.1	109–126	0	7	0.560	0.619	0.096	0.013
Mean	–	–	–	0.014	9.22	0.542	0.594	0.087	0.010

ER error rate, N_A total number of alleles, H_O observed heterozygosity, H_E expected heterozygosity, F_{IS} inbreeding coefficient (loci deviated from Hardy–Weinberg equilibrium at $P<0.05$), $_B$ null allele frequency averaged over all populations (Brookfield method)

*$P<0.05$

[a] Final concentration of each primer

[b] Size range of each allele

individual of each species, homozygous for the particular locus, was selected. The loci were amplified as described above, with unlabeled primers. The PCR products were purified using the QIAquick purification kit (Qiagen) and sequenced using the PCR primers in GATC Biotech (Konstanz, Germany).

The resulting sequences were proofread in Chromas Lite 2.01 (Technelysium Pty. Ltd., Brisbane, Queensland, Australia), and each locus was aligned manually in BioEdit 7.4.0.1 (Hall 1999) with the previously published sequences (Wu et al. 2002; Ogyu et al. 2003; Kulju et al. 2004; Lance et al. 2009; Tsuda et al. 2009a, b; Gürcan et al. 2010). For two loci (AF310863 and Alma11), the repeat motifs could not be read unambiguously, due to the close proximity of the PCR primers to the microsatellite region. Therefore, these loci were cloned using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions but were downscaled to half reactions. For each sample, three to four white colonies were harvested, transferred to 20 μl deionized water, and denatured for 10 min at 95 °C. The inserts were re-amplified using the M13 primers (supplied with the cloning kit). The PCRs were performed in 25 μl reactions containing 1× PCR buffer with KCl (Fermentas), 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.2 μM of each primer, 0.5 U of Taq polymerase (Fermentas), and 1 μl of the denatured colonies as a template. The cycling conditions were as follows: 5 min at 94 °C followed by 35 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1.5 min and final extension at 72 °C for 10 min. The PCR products were purified and sequenced from both ends using the M13 primers.

Results and discussion

Cross-amplification and multiplex PCR optimization

Starting with the 39 loci previously developed for the Betulaceae, 16 were excluded (Online Resource 1, part C) because they failed to amplify or showed poor amplification patterns after cross-amplification. This led to a total of 23 markers tested in multiplexes using labeled primers. Thirteen of these were eliminated (five were poorly amplified: A4, A14, A16, A27, and A36; six showed inconsistent electrophoretic patterns: A3, A8, A17, A20, A24, and A34; and two had insufficient polymorphism levels: A19 and A21; Online Resource 1, part B).

Finally, we designed two multiplexes composed of eight and nine microsatellite loci for *A. glutinosa* (Tables 2 and 4; Online Resource 1, part A) and *A. incana* (Tables 3 and 4), respectively. All loci integrated into multiplexes showed well interpretable and reproducible polymorphic patterns (Fig. 1). We further validated both multiplexes on 200 individuals from five localities for each species representing the distribution

Table 4 Microsatellite loci included in the final multiplexes for *Alnus glutinosa* and *A. incana*

No	Original species	Locus	GenBank (original)	Source	Repeat			Insertions in flanking region	GenBank	
					Original organism	*A. glutinosa*	*A. incana*		*A. glutinosa*	*A. incana*
A2	BETPLA	Bp04	AB084474	Wu et al. (2002)	$(GT)_{12}\ldots(GA)_5$	$(GT)_8GAAAGC(GA)_2A\ A(GA)_5$	$(GT)_8GAAAGC\ (GA)_7$		KF724864	KF724864
A6	BETMAX	Bmax624	AB094100	Ogyu et al. (2003)	$(TC)_{14}$	n/a	$(TC)_3TT(TC)_9$		n/a	KF724866
A7	BETMAX	Bmax097	AB094104	Ogyu et al. (2003)	$(CT)_{13}$	$(CT)_4CC(CT)_2$	n/a		KF724867	n/a
A10	BETPEN	L5.5	AF310863	Kulju et al. (2004)	$C_{12}CTCC(CT)_7TT(CT)_5$	$C_3AAC(CT)_9CGTT(CT)_2$	$C_3TCCC(CT)_8CG\ TT(CT)_2$	1 bp (*A. incana*)	KF724868	KF724869
A18	CORAVE	B634	FJ986496	Gürcan et al. (2010)	$(AG)_{15}$	n/a	$(CA)_4\ldots(AG)_{23}$		n/a	KF724870
A22	CORAVE	B720	FJ986523	Gürcan et al. (2010)	$(AG)_{14}$	$(AG)_{13}$	$(AG)_8$		KF724872	KF724871
A26	ALNMAR	Alma11	n/a	Lance et al. (2009)	$(CT)_{11}$	$(CT)_2AT(CT)_9\ldots(CT)_2\ (GT)_6$	$(CT)_{15}\ldots(CT)_2(GT)_6$	7 bp (*A. incana*)	KF724873	KF724874
A35	BETMAX	CD276907	CD276907	Tsuda et al. (2009a)	$(TC)_8$	$(TC)_{12}$	$(TC)_{10}$		KF724876	KF724875
A37	BETMAX	CD277113	CD277113	Tsuda et al. (2009b)	$(TC)_9$	$(TC)_{14}$	$(TC)_{10}$		KF724878	KF724877
A38	BETMAX	CD278280	CD278280	Tsuda et al. (2009a)	$(CAA)_5$	$(CAA)_7$	$(CAA)_5$		KF724880	KF724879

BETPLA Betula platyphylla, *BETMAX B. maximowicziana*, *BETPEN B. pendula*, *CORAVE Corylus avellana*, *ALNMAR Alnus maritima*, *n/a* not available

Springer

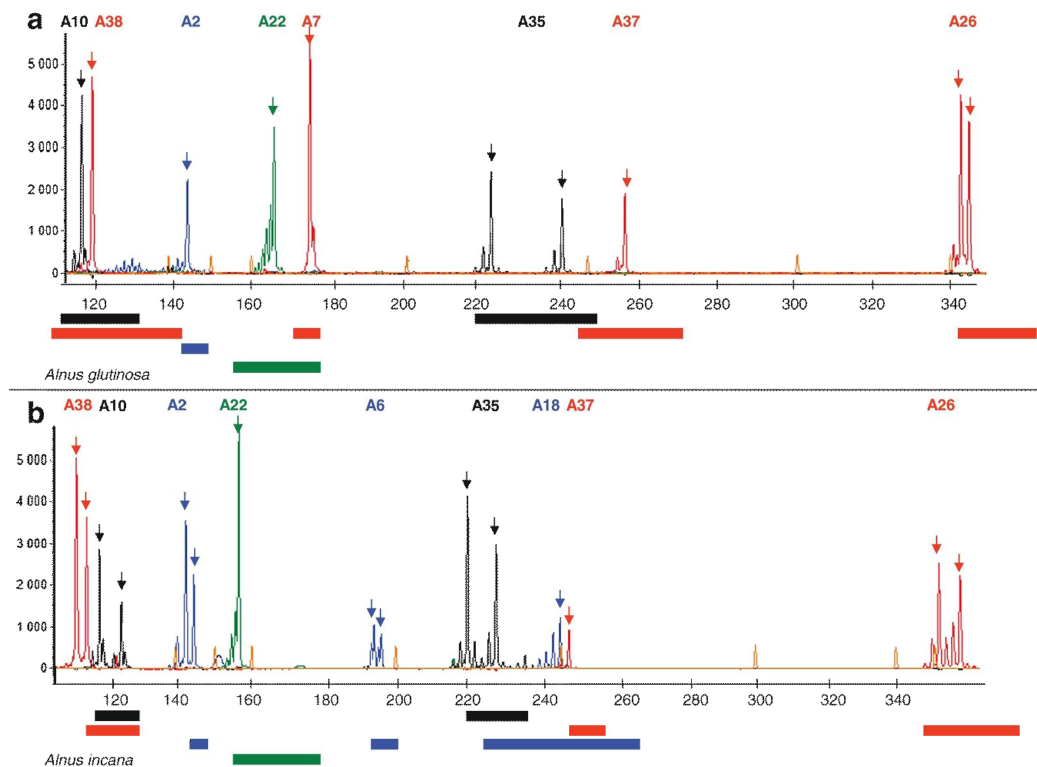


Fig. 1 Examples of an individual electropherograms and allele range sizes based on 100 individuals from across the range for *Alnus glutinosa* (**a**) and *A. incana* (**b**). *Each arrow* represents an actual allele

range (Table 1 (c)). Some microsatellite markers could he used in multiplex for only one species due to bad amplification or insufficient polymorphism in the second species. However, none of them was species specific, i.e., exclusively amplified in one species and failed to amplify in the other one. Specifically, loci A6 and A18 amplified inconsistently in *A. glutinosa*, whereas A7 amplified inconsistently in *A. incana* (Tables 2 and 3). PCR conditions were thoroughly tested in several hundred samples and, thus, are suitable for future large-scale population genetic studies of both *Alnus* species.

Genotype scoring and analyses

Most markers were characterized by a clear succession of dinucleotide repeats except A38 (trinucleotide). Off-ladder microvariants (i.e., variants differing from the expected periodicity of 2 or 3 bp) were observed for one locus of *A. glutinosa* (A26) and tbree loci of *A. incana* (A6, A10, and A26). Since these intermediate-sized variants were clearly separated from the neighhoring size classes, they did not cause problems for binning. Moreover, clearly defined off-ladder microvariants can improve the precision of subsequent analysis (Guichoux et al. 2011b). Presence of off-ladder microvariants can be caused by interrupted repeats of loci A10 and A26 (Table 4) which can mutate quickly and cause single-base pair differences.

Mean genotyping error rate combined overall loci, and all samples were 1.3 % per allele for *A. glutinosa* and 1.4 % per allele for *A. incana* (see error rate for each loci in Tables 2 and 3), illustrating the robustness of markers. The most problematic locus, A22, shows extra bands in some cases; for *A. glutinosa*, these bands are clearly separated from allele calling and can be simply omitted during scoring. Other errors were caused by amplification failure of one allele (A10, A22), occurrence of A+allele (A2), and different allele callings in hotb runs (A18, A26). When working at the population level, reasonable error rates (<2 % of genotypes mistyped) are unlikely to seriously bias results; on the other hand, serious problems can be found in parentage analysis (Bonin et al. 2004; Hoffman and Amos 2005).

Overall, tbe majority of loci showed relatively high genetic diversity (Tables 2 and 3). We identified 74 alleles at eight microsatellite loci in *A. glutinosa* and 83 alleles at nine microsatellite loci in *A. incana*, witb an average of 9.25 and 9.22 alleles per locus, respectively. It was shown that microsatellite mutation rates increase witb the number of repeats of tbe motif, indicating that long loci have higher allelic richness than short ones (Ellegren 2000; Petit et al. 2005). In our case, loci originally with less than ten repeats did not show any

repeat interruptions due to point mutation or compound repeat types in both alder species (i.e., A35, A37, A38; Table 4). However, even loci A37 and A38 showed a lower number of repeats, they expressed high allelic richness, mainly in *A. glutinosa* (Tables 2 and 3). In contrast, loci with few allele variants (i.e., A2, A7) were also included. They can be useful when allele distribution helps to distinguish populations and determine homozygote or heterozygote states. For locus, A2 alleles 142 and 144 were prevalent in *A. incana* and *A. glutinosa* populations, respectively. Locus A7 used only for *A. glutinosa* had two rare alleles. Allele 172 occurred in three out of five populations, and allele 174 occurred in four out of five populations with no clear geographical pattern.

The eight *A. glutinosa* microsatellites had no loci that differed significantly from Hardy–Weinberg expectations (Table 2). The same was not true for *A. incana* where four out of nine microsatellite loci significantly violated Hardy–Weinberg expectations (Table 3). Tests for linkage disequilibrium between pairs of loci revealed no significant results (data not shown) for both species. The mean H_O for *A. glutinosa* and *A. incana* was 0.561 and 0.542, respectively, and the mean H_E was 0.575 and 0.594, respectively. Low and nonsignificant values were obtained for inbreeding coefficient $f(F_{IS})$ for *A. glutinosa*. It varied from −0.0285 (locus A35) to 0.052 (locus A37) with a mean of 0.025. In *A. incana*, it varied from −0.055 (locus A6) to 0.796 (locus A2) with a mean 0.087.

The diagnostic results using Micro-Checker (Van Oosterhout et al. 2004) found no evidence of large allele dropout for any of the loci of either alder species. However, the analysis pointed to one out of 40 population/marker

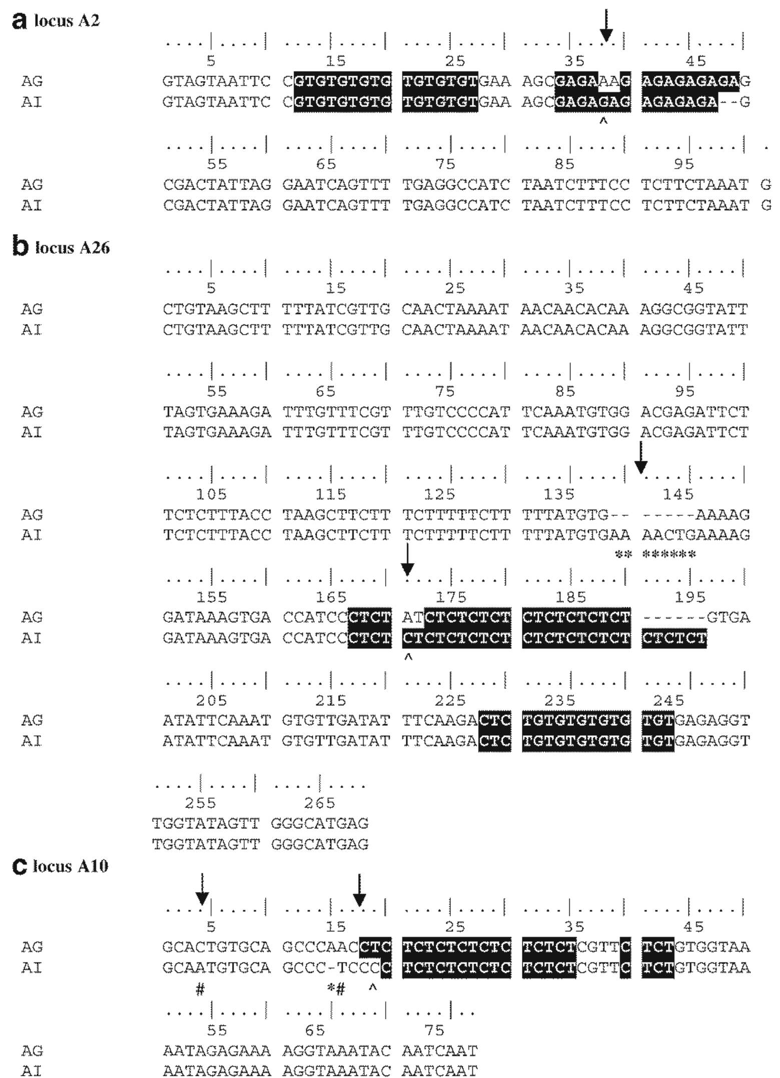


Fig. 2 Sequence variation in microsatellite loci among *Alnus glutinosa* and *A. incana*. **a** Locus A2, **b** locus A26, **c** locus A10. *AG* indicates sequences of *A. glutinosa* and *AI* sequences of *A. incana*; repeats are marked with *black*; variations in sequences are marked with *arrow* and distinguished by *symbols*: *circumflex accent* (substitutions within the repeat region), *asterisk* (indels within the flanking regions), and *number sign* (substitutions within the flanking regions)

combinations for *A. glutinosa* and three out of 45 population/marker combinations for *A. incana*, suggesting the presence of null alleles, involving one (A10) out of the eight loci for *A. glutinosa* and two (A2 and A22) out of the nine loci for *A. incana*. The average null allele frequency across all loci and all populations calculated using the Brookfield method was very low: −0.002 for *A. glutinosa* and 0.010 for *A. incana*, respectively (see results for each locus in Tables 2 and 3). According to the simulation study of Chapuis and Estoup (2007), ignoring the presence of null alleles up to 5 % on average across loci will only slightly bias classical estimates of population differentiation. Hence, our results can be directly used to explore population genetic diversity and structure. On the other hand, undetected null alleles can lead to overestimation of both F_{ST} and genetic distance in strongly differentiated populations (Chapuis and Estoup 2007), and also, their effect is most serious in fine-scale population studies and parentage analysis (Pompanon et al. 2005; Oddou-Muratorio et al. 2009). In these cases, we suggest omitting problematic loci from such analyses (Oddou-Muratorio et al. 2009) or use the correction options implemented in several programs, e.g., PAPA or Micro-Checker (Duchesne et al. 2002; Van Oosterhout et al. 2004).

Sequence variation in repeat motifs

All selected loci were sequenced in one individual of each *Alnus* species in order to ensure that the primers amplify microsatellites and characterize their repeat motif. Sequencing confirmed the presence of microsatellite repeats for all markers in both species. The repeat motifs of the markers A6, A7, A10, A18, and A26 found in both alder species differed from those described for the source species (see Table 4). However, these differences represent mainly the modifications of the original motifs and affected compound or interrupted (A10) as well as pure (A6, A7, A18, A26) repeats. Four loci with pure repeats in the source species (A6, A7, A18, and A26) showed repeat interruptions due to point mutation or compound repeat types in *A. incana*, *A. glutinosa*, or both. Loci which initially had compound repeats differed even more. A similar tendency of repeat motif modification among congeneric species or related genera has been previously reported, e.g., in Fabaceae (Peakall et al. 1998). Repeat motives of three loci (A2, A10, and A26) differed between the two alder species (Table 4). In the locus, A2 substitution of G for A in one of the GA repeat units in *A. glutinosa* resulted in repeat interruption when compared to *A. incana* (Fig. 2a). Similarly in A26, a C-to-A substitution in the CT stretch in *A. glutinosa* differentiates its repeat motif from that observed in *A. incana*. Moreover, *A. incana* contains a 7-bp-long insertion in the sequence flanking this microsatellite repeat (Fig. 2b). In the case of the A10 locus, sequences with different nucleotide compositions and length difference of 1 bp were flanking the CT motif (Fig. 2c). Homologous microsatellite loci are known to differ between closely related species, not only by repeat number but also by substitutions and indels in the sequence regions flanking them (Doyle et al. 1998; Peakall et al. 1998). These mutations could lead to erroneous estimates of allele homology, because alleles with identical lengths will not have the same number of repeat units (Peakall et al. 1998). Therefore, knowledge of the entire sequence amplified by the particular primer pair in all studied taxa is essential for studies aiming to use microsatellite loci at the interspecific level.

Conclusions

Cross-amplification of various species from the Betulaceae led to the development of two multiplexes for *A. glutinosa* and *A. incana*. All markers were precisely controlled in order to choose only markers with high levels of polymorphism, low error rates, and low occurrence of null alleles. In combination with previously developed markers for *Alnus*, our two multiplexes gained by cross-amplification of markers from Betulaceae should provide detailed information for future genetic analyses and enable use of modern statistical methods. Moreover, we have demonstrated that homologous SSR loci differ between *A. incana* and *A. glutinosa* by substitutions in repeat regions as well as by length mutations in regions flanking the repeats. These findings underline the necessity of sequence information for correct determination of homologous alleles in studies focused at the interspecific level.

Acknowledgments We would like to thank Martina Hadincová and Eva Ničová for their help in the lab and John P. Bailey for language editing of the manuscript. We also would like to thank Phil Gibson for his help with microsatellite markers originally developed for *Alnus maritima*. This study was supported by grant no. P504/11/0402 from the Grant Agency of the Czech Republic and as a long-term research development project no. RVO 67985939.

Data Archiving Statement The sequence data generated in this study have been deposited at GenBank under the accession numbers KF724864–KF724880.

References

Bonin A, Bellemain E, Bronken Eidesen P et al (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273. doi:10.1111/j.1365-294X.2004.02346.x

Brookfield J (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455. doi:10.1046/j.1365-294X.1996.00098.x

Chapuis M, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621–631. doi:10.1093/molbev/msl191

Claessens H, Oosterbaan A, Savill P, Rondeux J (2010) A review of the characteristics of black alder (*Alnus glutinosa* (L.) Gaertn.) and their implications for silvicultural practices. Forestry 83:163–175. doi:10.1093/forestry/cpp038

Cox K, Vanden Broeck A, Van Calster H, Mergeay J (2011) Temperature-related natural selection in a wind-pollinated tree across regional and continental scales. Mol Ecol 20:2724–2738. doi:10.1111/j.1365-294X.2011.05137.x

Doyle J, Morgante M, Scott V, Wayne P (1998) Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). Mol Biol Evol 15:215–218

Duchesne P, Godhout M-H, Bernatchez L (2002) PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental. Mol Ecol Notes 2:191–193. doi:10.1046/j.1471-8286.2002.00164.x

Ellegren H (2000) Microsatellite mutations in the germline. Trends Genet 16:551–558. doi:10.1093/mutage/ger028

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435–445. doi:10.1038/nrg1348

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Guichoux E, Lagache L, Wagner S et al (2011a) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611. doi:10.1111/j.1755-0998.2011.03014.x

Guichoux E, Lagache L, Wagner S et al (2011b) Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus* spp.). Mol Ecol Resour 11:578–585. doi:10.1111/j.1755-0998.2011.02983.x

Gürcan K, Mehlenbacher S (2010) Transferability of microsatellite markers in the Betulaceae. J Am Soc Hortic Sci 135:159–173

Gürcan K, Mehlenbacher SA, Botta R, Boccacci P (2010) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. Tree Genet Genomes 6:513–531. doi:10.1007/s11295-010-0269-y

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Mol Ecol 14:599–612. doi:10.1111/j.1365-294X.2004.02419.x

Huntley B, Birks H (1983) An atlas of past and present pollen maps for Europe: 0–13000 years ago. Cambridge University Press, Cambridge

King A, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. Mol Ecol 7:1151–1161. doi:10.1046/j.1365-294x.1998.00432.x

Kulju KKM, Pekkinen M, Varvio S (2004) Twenty-three microsatellite primer pairs for *Betula pendula* (Betulaceae). Mol Ecol Notes 4:471–473. doi:10.1111/j.1471-8286.2004.00704.x

Lance SL, Jones KL, Hagen C et al (2009) Development and characterization of nineteen polymorphic microsatellite loci from seaside alder, *Alnus maritima*. Conserv Genet 10:1907–1910. doi:10.1007/s10592-009-9851-y

Lepais O, Bacles CFE (2011) De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. J Hered 102:627–632. doi:10.1093/jhered/esr062

McVean D (1953) Account of *Alnus glutinosa* (L.) Gaertn. for the Biological Flora of the British Isles. J Ecol 41:447–466

Oddou-Muratorio S, Vendramin GG, Buiteveld J, Fady B (2009) Population estimators or progeny tests: what is the best method to assess null allele frequencies at SSR loci? Conserv Genet 10:1343–1347. doi:10.1007/s10592-008-9648-4

Ogyu K, Tsuda Y, Sugaya T et al (2003) Identification and characterization of microsatellite loci in *Betula maximowicziana* Regel. Mol Ecol Notes 3:268–269. doi:10.1046/j.1471-8286.2003.00419.x

Peakall R, Gilmore S, Keys W (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability. Mol Biol Evol 15:1275–1287

Pearse DE, Crandall KA (2004) Beyond FST: analysis of population genetic data for conservation. Conserv Genet 5:585–602. doi:10.1007/s10592-003-1863-4

Petit RJ, Deguilloux M, Chat J et al (2005) Standardizing for microsatellite length in comparisons of genetic diversity. Mol Ecol 14:885–890. doi:10.1111/j.1365-294X.2005.02446.x

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847–859. doi:10.1038/nrg1707

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615–629. doi:10.1111/j.1461-0248.2006.00889.x

Štorchová H, Hrdličková R, Chrtek J et al (2000) An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. Taxon 49:79–84. doi:10.2307/1223934

Tallantire (1974) The palaeohistory of the grey alder (*Alnus incana* (L.) Moench.) and black alder (*A. glutinosa* (L.) Gaertn.) in Fennoscandia. New Phytol 73:529–546. doi:10.1111/j.1469-8137.1974.tb02131.x

Truong C, Palme A, Felber F, Naciri-Graven Y (2005) Isolation and characterization of microsatellite markers in the tetraploid birch, *Betula pubescens* ssp. *tortuosa*. Mol Ecol Notes 5:96–98. doi:10.1111/j.1471-8286.2004.00848.x

Tsuda Y, Ueno S, Ide Y, Tsumura Y (2009a) Development of 14 EST-SSRs for *Betula maximowicziana* and their applicability to related species. Conserv Genet 10:661–664. doi:10.1007/s10592-008-9608-z

Tsuda Y, Ueno S, Ranta J et al (2009b) Development of 11 EST-SSRs for Japanese white birch, *Betula platyphylla* var. *japonica* and their transferability to related species. Conserv Genet 10:1385–1388. doi:10.1007/s10592-008-9701-3

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. doi:10.1111/j.1471-8286.2004.00684.x

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370. doi:10.2307/2408641

Wilson G, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191

Wu B, Lian C, Hogetsu T (2002) Development of microsatellite markers in white birch (*Betula platyphylla* var. *japonica*). Mol Ecol 3:413–415. doi:10.1046/j.1471-8278

Yatrik Y (1982) Betulaceae. In: Davis P (ed) Flora Turkey, vol 7. Edinburh University Press, Edinburgh, pp 688–694

Zhuk A, Veinberga I, Daugavietis M, Rungis D (2008) Cross-species amplification of *Betula pendula* Roth. simple sequence repeat markers in *Alnus* species. Balt For 14:116–121

Flow cytometry, microsatellites and niche models reveal the origins and geographic structure of *Alnus glutinosa* populations in Europe

Bohumil Mandák[1, 2, *], Petr Vít[1, 2], Karol Krak[1, 2], Pavel Trávníček[1, 3], Alena Havrdová[1, 2], Věroslava Hadincová[1], Petr Zákravský[1], Vlasta Jarolímová[1], Cecile Fanny Emilie Bacles[4] and Jan Douda[1, 2]

[1] Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43, Průhonice, Czech Republic; [2] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha 6 – Suchdol, CZ-165 21, Czech Republic; [3] Faculty of Agriculture, University of South Bohemia, Studentská 13, CZ-370 05, České Budějovice, Czech Republic; [4] University of Pau and Pays Adour, UFR Sciences et Techniques, Departement de Biologie, F-64100, Pau, France. * *For correspondence. E-mail bohumil.mandak@ibot.cas.cz*

ABSTRACT

- **Background and Aims** Polyploidy in plants is an intensively studied topic. In many groups, two or more cytotypes represent separate biological entities with distinct distribution, history and ecology. This paper examines the distribution and origins of cytotypes of *Alnus glutinosa* in Europe, Northern Africa and Western Asia.
- **Methods** Our combined approach involved flow cytometry and microsatellite analysis of twelve loci in 2,200 plants from 209 populations together with species distribution modelling by MIROC and CCSM climatic models to analyse (i) ploidy and genetic variation; (ii) the origin of tetraploid *A. glutinosa,* considering *A. incana* as a putative parent; and (iii) past distribution of the species.
- **Key Results** We for the first time report the occurrence of tetraploid populations of *A. glutinosa* in Europe. The distribution of tetraploids is far from random, forming two geographically well delimited clusters located in the Iberian Peninsula and in the Dinaric Alps. Based on microsatellite analysis, both tetraploid clusters are probably of autopolyploid origin with no indication that *A.*

incana was involved in their evolutionary history. A projection of the MIROC distribution model into the Last Glacial Maximum (LGM) showed that (i) populations occurring in the Iberian Peninsula and North Africa are likely to have been interconnected during the LGM, and that (ii) populations occurring in the Dinaric Alps did not exist throughout the last glacial periods and were retreated south into lowland areas of the Balkan Peninsula.

- **Conclusions** Newly discovered tetraploid populations are situated in the putative main glacial refugia, and neither of them was probably involved in the colonization of Central and Northern Europe after glacial withdrawal. This could mean that neither the Iberian Peninsula nor the western part of the Balkan Peninsula served as effective refugial areas for northward postglacial expansion of *A. glutinosa*.

Key words: *Alnus*, autopolyploidy, cytotype distribution, ecological niche models, flow cytometry, glacial refugia, microsatellites.

INTRODUCTION

Polyploidy, the possession of more than two sets of chromosomes, has been an important factor in eukaryote evolution (Otto, 2007), with as many as 80% of all taxa estimated to have had a polyploid origin (Stebbins, 1950, 1971; Grant, 1981; Masterson, 1994; Otto and Whitton, 2000). Recently it has been shown that almost all vascular plants have undergone at least one polyploidization event in their evolutionary history, one exception being *Amborella* (Amborella Genome Project, 2013). Beyond all doubt, genome doubling has been an important process in plant evolution, quickly producing novel cytotypes by two major processes. While hybridization between different cross-sterile taxa, which can escape from sterility by chromosome doubling, produces allopolyploid individuals (Lowe and Abbott, 1996; Kochert *et al.*, 1996; Cook *et al.*, 1998; Segraves *et al.*, 1999; Kolář *et al.*, 2009; Mandák *et al.*, 2012), hybridization between fully cross-fertile progenitors – accompanied by doubling of structurally similar, homologous genomes – produces autopolyploid individuals (Thompson and Lumaret, 1992; Parisod *et al.*, 2010). Allopolyploidy has been recognized as the most common process in polyploid formation for a long time (Stebbins, 1950, 1971; Wendel, 2000), but only recently have different authors shown that autopolyploidy has probably been overlooked and that the number of autopolyploids is underestimated (Mahy *et al.*, 2000; Soltis *et al.*, 2007; Parisod *et al.*, 2010).

Both allo- and autopolyploids potentially harbour more genetic variation than their diploid progenitors because they combine more than two gene copies, but differ in their mode of inheritance (disomic in allopolyploids vs polysomic in autopolyploids) (Catalán *et al.*, 2006; Parisod *et al.*, 2010). Moreover, recent studies have demonstrated that polyploid genomes can be highly dynamic and undergo rapid structural and functional alternations (Doyle *et al.*, 2008; Leitch and Leitch, 2008).

It has long been known that the frequency of polyploidy increases with latitude in the Northern Hemisphere (Hagerup, 1931; Löve and Löve, 1957; Johnson and Packer, 1965; Stebbins, 1950). Stebbins (1984, 1985) suggested a correlation between the frequency of polyploidy and the degree of glaciation, rather than purely between polyploidy and latitude. This led Stebbins to describe polyploidy as a process stabilizing new gene combinations derived from hybridization between

genetically distinct parental gene pools adapted to different ecological conditions ("secondary contact hypothesis"). The view of Stebbins (1984) was recently revisited by Brochmann *et al.* (2004), who used Arctic plants as a model system. They did not find any association between polyploidy and the degree of glaciation for the Arctic flora as a whole, but for Arctic specialist taxa with restricted distribution, the frequency of diploids was higher in largely unglaciated areas during the last Ice Age than in heavily glaciated areas. Even differences in current distribution among cytotypes might be significantly affected by history, the pattern of polyploidy increasing with latitude might not be absolutely clear in many taxa (Mandáková and Münzbergová, 2006; Kolář *et al.*, 2009, 2012; Duchoslav *et al.*, 2010). Therefore, the position of past cytotype refugia relative to sites that became available for colonization by a single cytotype is probably also important and may significantly complicate the pattern of increasing ploidy with latitude. Kolář *et al.* (2009, 2012) revealed a unique evolutionary pattern in *Knautia arvensis* agg. They determined a wide variety of processes and mechanisms which likely took part in the rapid evolution of this complex, including isolation in Holocene refugia, repeated colonization by distinct lineages, hybridization and recurrent polyploidization. All these processes generate diffuse patterns of cytotype distribution not related to the general trend of increasing ploidy levels from the south to the north.

Putative migration routes after glaciers retreated have mostly been described based on studies of postglacial tree migrations (Bennett *et al.*, 1991; Hewitt, 1996, 2000; Taberlet *et al.*, 1998; Cruzan and Templeton, 2000; Tzedakis *et al.*, 2013). These studies have demonstrated that the current distributions of different tree species in Europe are a result of multiple migration scenarios (Petit *et al.*, 2002; Magri *et al.*, 2006; Tollefsrud *et al.*, 2008; Parducci *et al.*, 2012, Wachowiak *et al.*, 2013). King and Ferris (1998) suggested, based on an analysis of chloroplast DNA, that *Alnus glutinosa* (L.) Gaertn., a foundation tree species of alluvial forests, may have expanded from a Balkan refugium to Northern and Western Europe after the Ice Age. Even though the species was present in Italian and Iberian refugia, these populations do not seem to have participated in postglacial recolonization of northern Europe including the far north. However, the sampling of King and Ferris (1998) completely missed the Western Russian Plains, where at least one important refugium is supposed to be located. Furthermore, Douda *et al.* (2014) showed, on the

basis of radiocarbon dated pollen and macrofossil sites, that *Alnus* trees were likely to have withstood the LGM in Western Europe, the northern foothills of the Alps, the Romanian Carpathians and a large area of northeastern Europe. It follows that, after withdrawal of glaciers, *Alnus* rapidly colonized southern Sweden and gradually expanded northward, most likely predominantly from a refugium located in today's Belarus and western Russia. The increase in *Alnus* occurrence in more southerly ice-free areas of Europe seems to reflect local expansions originating from regional refugia.

Alnus glutinosa and *A. incana* (family *Betulaceae*) have been reported to be diploid (2n = 2x = 28) by many authors (Fedorov, 1969; Goldblatt, 1981, 1984, 1985, 1988; Goldblatt and Johnson, 1990, 1991, 1994, 1996, 1998, 2000, 2003; Ivanova *et al.*, 2006). However, there is one tetraploid record (2n = 4x = 56) for *A. glutinosa* reported in Fedorov (1969) that is not based on chromosome counting of European samples, but American material collected by Woodworth (1929, 1931). Recently, Lepais *et al.* (2013) described putative tetraploid populations from North Africa (Morocco) based on nuclear microsatellite genotyping.

Both *Alnus glutinosa* and *A. incana* are wind-pollinated, self-incompatible trees of riparian and water-logged habitats (McVean, 1953; Tallantire, 1974, Douda *et al.*, 2012). The distribution of the two alder species significantly differ (Jalas and Suominen, 1976). *Alnus glutinosa* grows in lowlands and midlands throughout Europe, except the extreme north, extending as far as Siberia and the mountains of Turkey and North Africa (McVean, 1953). It is common in southern Fennoscandia, but northwards it is associated with a coastal oceanic climate (Tallantire, 1974). The range of *A. incana* is divided into a northern and a southern part, similarly as that of Norway spruce (*Picea abies*) (Jalas and Suominen, 1976). However, the species' distribution range during the LGM was probably significantly wider, covering also Western Europe and part of the Iberian Peninsula, as has been showed by Svenning *et al.* (2008) using species distribution modelling. Both species were therefore probably in contact elsewhere across Europe throughout the last Ice Age.

Some individuals of *A. glutinosa* collected in 2011 in the Iberian Peninsula and analysed using microsatellites turned out not to be diploids, contrary to reports of numerous authors (but see Fedorov, 1969; Lepaise *et al.*, 2013), but probably tetraploids. The occurrence of a previously nearly unknown ploidy level prompted

the study of cytogeographical patterns in Europe using a method other than microsatellites, which do not provide us with precise information concerning the ploidy level of all collected individuals. We therefore conducted extensive screening of *A. glutinosa* populations across its distribution range to determine ploidy levels distribution and also included *A. incana* (L.) Moench. to test whether tetraploids are of an auto- or allopolyploid origin.

One of the hypotheses tested was whether the distribution of different ploidy levels fits the classical concept of increasing polyploidy frequency from southern to northern Europe, i.e. whether polyploids are more likely to persist under extreme climatic conditions. The following questions were asked: (1) What are the frequencies and distribution patterns of plants of different ploidy levels? (2) Are there mixed-cytotype populations? (3) Are particular cytotypes restricted to certain geographical areas? (4) What is the origin of the tetraploid cytotype? Specifically, have tetraploids evolved by autopolyploidy, or are they a result of hybridization with *A. incana* followed by a polyploidization event? (5) What is the Late Pleistocene history of the tetrapoloid cytotype?

MATERIALS AND METHODS

Study area and sampling procedure

The research area covers most of the European distribution range of *A. glutinosa*, although populations from Siberia were not included and the area of the mountains of Turkey was represented by only a few samples. Samples of *A. glutinosa* were collected across this distribution range in summer months between 2011 and 2013 exclusively from natural unmanaged forest stands. Several sampling strategies were applied to include as many samples as possible in the analysis. (1) **Detailed sampling** (94 localities) – a stratified random sampling procedure was used to sample populations and individuals within populations. Populations were at least 100 km apart. Whenever possible, 20 individuals per population were collected. Within each population, samples were collected along a linear transect from individuals at least 50 m apart; i.e. each population represents an area of at least one kilometre (Fig. 1, Appendix 1). Areas with the detected presence of tetraploids were visited several

times to ascertain ploidy homogeneity of populations and better delimit the boundaries between diploid and tetraploid populations. Some populations are therefore listed under “detailed sampling” although they comprised less than 20 collected individuals and were less than 100 km apart. (2) **Coarse sampling** (70 localities) – at least three individuals per population were collected in some part of the range by our collaborators. In these cases, individuals were at least 500 m apart (Fig. 1, Appendix 1). (3) **Samples from seed collections** (45 localities) – to extend the number of examined populations, we also included populations from whole Europe maintained by the International Alder Seed Bank at the Research Institute for Nature and Forest in Belgium (Fig. 1, Appendix 1). Plants were cultivated in the experimental garden of the Institute of Botany, The Czech Academy of Sciences, Průhonice, Czech Republic (49º59'30"N, 14º34'00"E, *ca* 320 m above sea level). Seeds were germinated in 5 × 5 cm bedding cells with homogeneous garden compost and later moved from the bedding cells to 19 × 19 × 19 cm (6.9 L) pots filled with a common garden substrate.

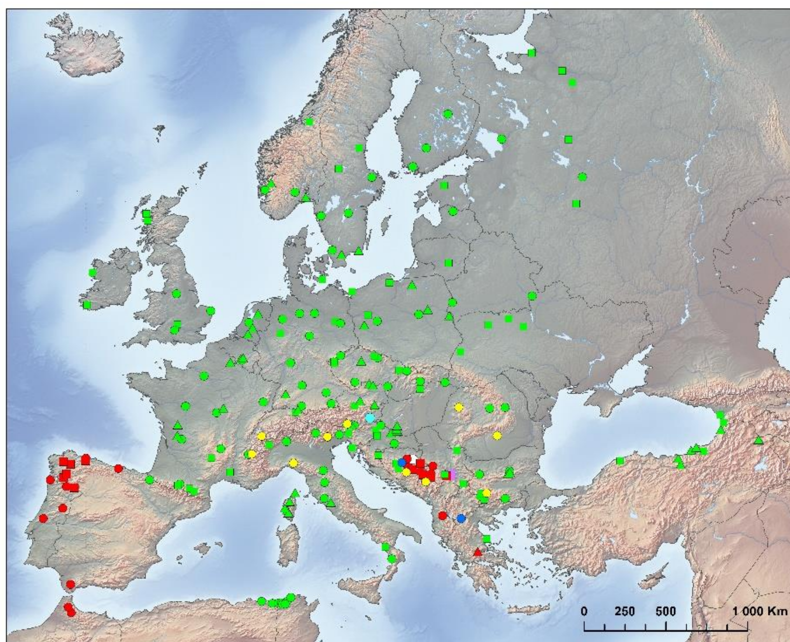


Fig. 1 Geographic locations of populations of *Alnus glutinosa* and *A. incana* across the entire study area. *A. incana* – yellow circles, *A. glutinosa* – others symbols and colours. Different sampling strategies are represented by different symbols: circles – detailed sampling, squares – coarse sampling, triangles – samples from seeds collections (see Materials and Methods for details). Cytotypes within *A. glutinosa* populations are represented by different colours: green – diploid populations, red – tetraploid populations, violet – mixed populations of diploids and tetraploids, light blue – mixed populations of diploids and triploids, white – mixed populations of triploids and tetraploids, dark blue – mixed populations of diploids, triploids and tetraploids.

For detailed sampling, all samples were stored in silica gel, and whenever possible, fresh leaf were transported to the laboratory and immediately analysed. This, however, was not the case for all samples because some expeditions were too long to keep samples fresh enough for flow cytometry analysis or DNA extraction. While all coarse samples were stored in silica gel and analysed later, all samples from collections planted in the experimental garden were analysed fresh.

To test the hypothesis that tetraploid *A. glutinosa* has an allopolyploid origin with *A. incana* serving as one parental species, we also collected samples of *A. incana* in the same way as in the detailed sampling of *A. glutinosa*. Only populations growing in south Europe, i.e. close to tetraploid populations, were included in the microsatellite analysis (see Fig. 1, Appendix 1).

Estimation of DNA ploidy level and genome size

DNA ploidy levels (Suda *et al.*, 2006) and absolute genome sizes (C-values; Greilhuber *et al.*, 2005) of *Alnus* species were estimated using flow cytometry. For flow cytometry analyses, 209 populations and 2200 individuals of *A. glutinosa* were collected in total. Moreover, 33 plants from 13 populations were subjected to analysis of absolute genome size, i.e. seventeen plants estimated by flow cytometry as diploids from populations 1, 85, 86, 91, 98 and 100, and sixteen plants estimated as tetraploids from populations 12, 34, 48, 63, 66, 102 and 106 (see Appendix 1 for exact locations). Both fresh leaves and leaves stored in silica gel were used for analyses. Young, intact leaf tissue of the analysed plants and an appropriate amount of leaf tissue of the internal reference standard [*Bellis perennis*; 2C-value set to 3.38 pg following Schönswetter *et al.* (2007)] were chopped together using a sharp razor blade in a plastic Petri-dish containing 0.5 ml of ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20) (Otto, 1990; Doležel *et al.*, 2007). The crude suspension was filtered through a 0.42 μm nylon mesh to remove tissue debris and then incubated for at least 30 min at room temperature. Isolated nuclei were stained with 1 ml of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12\ H_2O$) supplemented with the AT-selective fluorochrome 4’,6-diamidino-2-phenylindole (DAPI) and ß-mercaptoethanol at final concentrations of 4 μg/ml and 2 μg/ml, respectively. Immediately after staining, the relative fluorescence intensity of at least 3,000 particles was recorded on a CyFlow Space flow cytometer (Partec

GmbH, Münster, Germany) equipped with a diode UV chip as an excitation light source.

A different staining procedure was used for absolute genome size estimation. The suspension of isolated nuclei was stained with a solution containing 1 ml of Otto II buffer (0.4M $Na_2HPO_4 \cdot 12\ H_2O$), β-mercaptoethanol (final concentration of 2 μl/ml), propidium iodide (PI) and RNase IIA (both at final concentrations of 50 μg/ml). Samples were stained for 5 min at room temperature before being run through the flow cytometer CyFlow SL (Partec GmbH, Münster, Germany). Isolated stained nuclei were excited with a laser beam of 532 nm (solid-state laser Samba, Cobolt AB, Solna, Sweden), and the fluorescence intensity of 5,000 particles was recorded. Resulting histograms were evaluated using the application FloMax (Partec GmbH, Münster, Germany); DNA ploidy levels and absolute genome sizes were determined on the basis of the sample/standard ratio. Each plant was analysed separately. Our previous pilot study confirmed the lack of variation in the sample/standard ratio between fresh and silica-dried samples analysed in the same way. The reliability of FCM measurements (i.e. between-plant differences) was repeatedly confirmed in simultaneous runs of *Alnus* accessions yielding distinct fluorescence intensities (i.e. resulting in furcate double peaks in FCM histograms (Greilhuber, 2005).

Chromosome counts

To confirm the reliability of the ploidy estimates, FCM results were supplemented by conventional chromosome counts. Two diploids (locality No. 1) and two tetraploids (locality No. 48) were analysed. Chromosome counts were obtained from somatic mitotic cells in root-tips of pot cultivated plants. The root tips were pre-treated in a saturated water solution of p-dichlorbenzene for approximately 2 hours, then fixed in a 3:1 mixture of 96% ethanol and acetic acid, macerated in a 1:1 mixture of ethanol and hydrochloric acid for 15 s, washed in water, and stained with lacto-propionic orcein. The number of chromosomes was determined under an NU Zeiss microsope with an Olympus E 510 camera attached.

Microsatellite analysis

DNA extraction

Fresh leaves were collected and stored in silica gel. DNA from *A. glutinosa* and *A. incana* samples was isolated using the DNeasy 96 Plant Kit (Qiagen, Germany). The quality and yield of isolated DNA was checked on 1% agarose gels, and then precisely measured for DNA concentration and purity using NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). All samples were then diluted to a 20–25 ng/µl concentration prior to the PCR (for more details see also Drašnarová *et al.*, 2014).

PCR and fragment analyses

For microsatellite analysis, 31 populations and 619 individuals of *A. glutinosa,* and 10 populations and 194 individuals of *A. incana* were used (see Fig. 1 for the geographical distribution of populations). We analysed genetic variation at 12 nuclear microsatellite loci in 619 samples of *A. glutinosa* and 194 samples of *A. incana* from 31 and 10 populations, respectively (Appendix 1). These loci have been cross-amplified from closely related species by Drašnarová *et al.* (2014) (multiplex PCR 1: A2, A22, A35, A37, A38) or developed specifically for both *Alnus* species by Lepais and Bacles (2011) (multiplex PCR 2: Ag1, Ag5, Ag9, Ag10, Ag13, Ag20, Ag30). Concerning multiplex PCR 1, we used only five microsatellite loci out of ten published in Drašnarová *et al.* (2014) due to an amplification failure in one of the species studied (A6, A7, A18) and mutations in the sequence region flanking microsatellite loci (A2, A10, A26). Such mutations can lead to erroneous estimates of allele homology because alleles with identical lengths will not have the same number of repeat units (Drašnarová *et al.*, 2014). In the case of multiplex PCR 2, only seven microsatellite loci out of 12 published by Lepais and Bacles (2011) were used. The reasons for rejection of some loci were the same as in the case of multiplex PCR 1, i.e. amplification failure in one of the species studied (Ag14, Ag23, Ag25, Ag35) and different fragment length in homologous loci between species (Ag9).

DNA was amplified using the QIAGEN Multiplex PCR kit (QIAGEN, Germany) in a total reaction volume of 5 µl of PCR mix plus 5 µl of mineral oil to keep the PCR mix from evaporating. The mix contained 20–25 ng, i.e. 1 µl, of DNA, 0.1–0.5 µM of each primer and 2.5 µl of Master Mix (QIAGEN). PCR amplifications were conducted in a Mastercycler (Eppendorf, Germany) under the following conditions for multiplex PCR1: 15 min of denaturation at 95 °C, followed by 40 cycles at 94 °C for 30 s, 30 s at 58 °C, 60 s at 72 °C and a final extension of 10 min at 72 °C; and for multiplex PCR2: 5 min denaturation at 95°C, 30 cycles at 95°C for 30 s, 58°C for 3 min, 72 °C for 30 s and extension of 30 min at 60 °C. PCR products were electrophoresed in an ABI PRISM 3130 sequencer (Applied Biosystems, USA). One microlitre of PCR product was mixed with 0.2 µl of GeneScan-500 LIZ (Applied Biosystems) and 12 µl of Hi–Di formamide (Applied Biosystems). Allele sizes were determined using GeneMarker version 2.4.0 (SoftGenetics, USA). Microsatellite locus was treated as missing data after two or more amplification failures.

Data analysis

Genetic diversity

To examine the genetic diversity of diploid and tetraploid populations of *A. glutinosa* and diploid *A. incana*, we computed Nei's (1978) gene diversity (H_e) corrected for sample size and the average number of alleles (A) in each population using SPAGeDi version 1.2 (Hardy and Vekemans, 2002), a programme that computes statistics and permutation tests of relatedness and differentiation among populations for organisms of any ploidy level (Hardy and Vekemans, 2002). We used analysis of variance to compare gene diversity (H_e) and the average number of alleles (A) between *A. glutinosa* and *A. incana,* and among cytotypes within *A. glutinosa* (StatSoft, Inc., 2013).

Population structure and the origin of tetraploid populations

To resolve the origin of tetraploid populations situated in two distinct geographic locations (Fig. 1), we used two approaches: (1) To examine genetic similarities and

relationships among individuals of the two species and among different cytotypes within *A. glutinosa,* we performed a PCoA in R (R core team, 2014) using the POLYSAT package (Clark and Jasieniuk, 2011). A pairwise distance matrix among all samples was calculated using Bruvo distances (Bruvo *et al.*, 2004) as implemented in POLYSAT 1.3-0. This measure of genetic distance has been developed specifically for polyploids and takes distances between microsatellite alleles into account without knowledge of the allele copy number (Clark and Jasieniuk, 2011). (2) Bayesian model-based clustering of microsatellite data was employed using the procedure implemented in STRUCTURE 2.3.3 (Pritchard *et al.*, 2000), accounting for different ploidy levels and allele copy ambiguity in the co-dominant dataset (Falush *et al.*, 2007). This analysis was run to infer homogeneous genetic clusters and detect genetic admixture. POLYSAT 1.3-0 was used to generate the input data file for STRUCTURE computations. Ten replicates for each K = 1–10 (the user-defined number of clusters) with the burn-in length of 100,000 generations and the data collection of additional 1,000,000 generations were run, using the admixture model and correlated allele frequencies. We analysed three datasets in order to determine the origin of polyploid individuals of *A. glutinosa*. (1) Analysis of the whole dataset containing di-, tri- and tetraploid *A. glutinosa* individuals and *A. incana* to infer the allopolyploid origin of tetraploid *A. glutinosa*. (2) Analysis of di-, tri- and tetraploid *A. glutinosa* individuals mainly to infer the relationship between diploids and two geographically separate clusters of tetraploids. (3) Analysis of a subset of di-, tri- and tetraploid *A. glutinosa* individuals from the Balkans, i.e. populations 24, 29, 30, 32, 33, 34, 35 and 36, to infer the origin of triploids in cytotypically mixed populations. Population 13 from Austria was also added to this dataset due to the presence of one triploid individual. The STRUCTURE output data were parsed using the program Structure-sum running in the R runtime (Ehrich *et al.*, 2007), mainly to determine the optimal K value following the method of Nordborg *et al.* (2005). Alignment of cluster assignments across replicate analyses was then conducted in CLUMPP 1.1.2 (Jakobsson and Rosenberg, 2007), and subsequently visualized using DISTRUCT 1.1 (Rosenberg, 2004).

Palaeo-distribution modelling

We used the maximum entropy machine-learning approach implemented in Maxent 3.3.3k (Elith *et al.*, 2006; Phillips *et al.*, 2006) to infer the potential present-day distribution of tetraploid populations and their potential range during the Last Glacial Maximum (LGM, ~21 ka) and Last Interglacial periods (LIG, ~115 ka). The aim was to explore whether the potential distribution of tetraploids during the LGM corresponds to their recent range or whether they have experienced population withdrawal or expansion during the LGM in some areas. For distributional modelling, 33 coordinates of sites where the presence of tetraploids was detected using flow cytometry analysis were used. Climatic layers for 19 climatic variables were obtained from the WorldClim database (Hijmans *et al.*, 2005; http://www.worldclim.org). The LIG climatic model simulation with 30 s (~1 km) resolution followed Otto-Bliesner *et al.* (2006). For LGM, two climatic models with 2.5 min (~4 km) resolution were used: MIROC (Model for Interdisciplinary Research on Climate; Hasumi and Emori, 2004) and CCSM (Community Climate System Model; Collins *et al.*, 2006). First, we clipped all climatic layers to span from 35.90˚N to 75.71° N and from 12.07° W to 48.05° E, including all populations of *A. glutinosa* with determined ploidy level. Then we selected the nine most biologically relevant and relatively uncorrelated ($r < 0.7$) bioclimatic variables, i.e. annual mean temperature (BIO1), temperature seasonality (BIO4), minimum temperature of the coldest month (BIO6), mean temperature of the wettest quarter (BIO8), mean temperature of the warmest quarter (BIO10), mean temperature of the coldest quarter (BIO11), annual precipitation (BIO12), precipitation seasonality (BIO15) and precipitation of the warmest quarter (BIO18).
In MAXENT, 5-fold cross-validation and default settings with the “auto features” option, 10^{-5} convergence threshold, 1,000 maximum iteration, regularization parameter $\beta = 1$ and 100,000 background points were selected. In each of 5 replicates, 80% of sites were used as the training dataset, and 20% of the remaining sites formed the testing dataset. The quality of the final model was evaluated using the AUC statistic (Peterson *et al.*, 2008). The resulting model projections were the medians from 5 replicates. To delimit the potential area with suitable conditions for the occurrence of tetraploids at present, the LGM and the LIG, we used the “equal

training sensitivity and specificity" logistic threshold (Liu *et al.*, 2005). The potential area of long-term stable occurrence of tetraploid populations resulted from an overlap of suitable areas predicted for the present, LGM and LIG.

RESULTS

Cytotype composition

Chromosome numbers were obtained for two putatively diploid and tetraploid individuals, confirming the existence of diploid (2n = 2x = 28) and tetraploid (2n = 4x = 56) cytotypes of *Alnus glutinosa* (Fig. 2). A total of 2,200 *A. glutinosa* samples from 209 populations were analysed to assess their ploidy level, and a subset was used to estimate absolute nuclear DNA content (genome size). The three sampling strategies – detailed sampling, coarse sampling and sampling of seed collections – yielded 94, 70 and 45 localities, respectively. Three different DNA ploidy levels were recorded. Of the 209 populations sampled, diploids occurred in 85.6%, triploid in 1.9% and tetraploids in 15.8%. Populations consisted of one, two

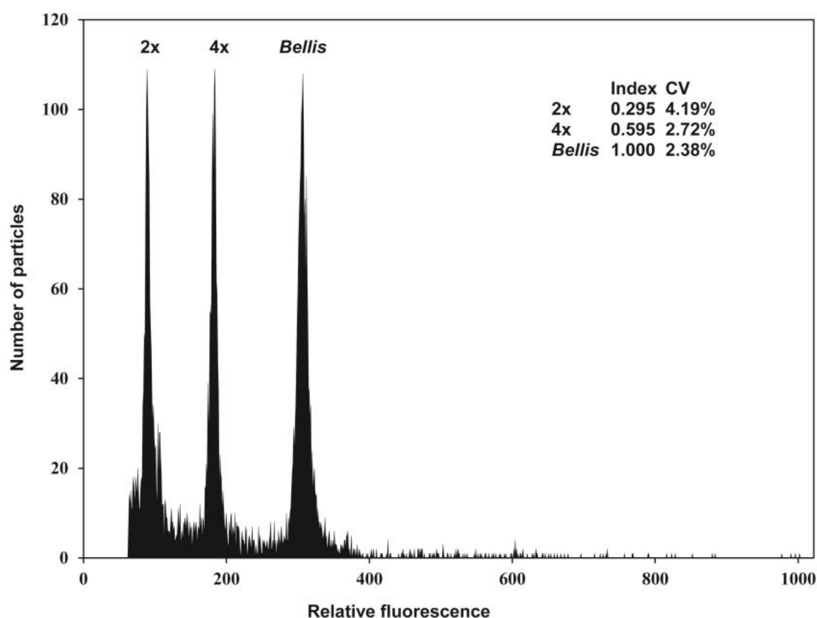


Fig. 2 Flow cytometric histogram of DAPI-stained nuclei of both ploidy levels (2× and 4×) of *Alnus glutinosa* analysed simultaneously with the internal standard *Bellis perennis*. Sample to standard ratio of peaks is represented by index.

or three ploidy levels (Fig. 1, Appendix 1). Most of the populations sampled (97.6%) consisted exclusively of one ploidy level. Populations comprising two or three ploidy levels were extremely rare (1.4% and 1.0%, respectively). Among the populations consisting of a single ploidy level, 83.7% consisted of diploids, and 13.9% consisted of tetraploids.

The distribution of ploidy levels in Europe, North Africa and Western Asia (Georgia, Turkey) departed from a random pattern significantly (Fig. 1). Diploid populations prevailed in south-eastern, eastern, western, central and northern Europe, and in northern Africa (Algeria, Tunisia) and Western Asia (Fig. 1). Tetraploid populations occupied two geographically distinct areas. One was situated in the Iberian Peninsula (Portugal, Spain) and northern Africa (Morocco), the other in the Dinaric Alps and adjacent areas of Greece, Albania, Montenegro, Serbia, and Bosnia and Herzegovina (Fig. 1). We found two mixed populations consisting of all three ploidy levels in the contact zone between the area dominated by tetraploids in the Dinaric Alps and diploids (Fig. 1). No mixed populations were found in the Iberian Peninsula or northern Africa (Fig. 1).

Mean 2C-values±SD for diploids, Iberian tetraploids and Dinaric tetraploids of *A. glutinosa* cytotypes were estimated to be 1.010±0.010 pg, 2.091±0.016 pg and 2.070±0.007 pg, respectively. Monoploid genome sizes (1Cx-values±SD) were 0.505±0.005 pg, 0.522±0.004 pg and 0.518±0.002 pg for diploids, Iberian tetraploids and Dinaric tetraploids, respectively (significantly different according to one-way ANOVA, $F_{2,\ 33} = 61.4$, $P < 10^{-6}$). Diploids differed significantly from tetraploid populations in both areas, but tetraploid populations from the Iberian Peninsula and the Dinaric Alps did not differ significantly at $P < 0.05$ according to a multiple-range Tukey's test.

Gene diversity

The overall mean number of alleles per locus (*A*) and gene diversity (*H*e) for *A. glutinosa* and *A. incana* (see Appendix 1 for individual values) differed significantly ($F_{1,\ 492} = 91.98$, $P < 10^{-2}$ and $F_{1,\ 492} = 54.69$, $P < 10^{-6}$, respectively); *A. glutinosa* reached higher values (Table 1). A comparison of the number of alleles per locus (*A*) and gene diversity (*H*e) within *A. glutinosa* for different regions characterized by the

occurrence of different ploidy levels also yielded a significant result ($F_{3, 372} = 23.29$, $P < 10^{-6}$; $F_{3, 372} = 9.11$, $P < 10^{-5}$). The highest genetic diversity was found in tetraploid populations of *A. glutinosa* from the Balkan Peninsula, and the lowest was found in diploid populations from northern Africa (Table 1).

Table 1 Summary of the average number of alleles (*A*) and gene diversities (*H*e) for *Alnus glutinosa* and *A. incana*, individual ploidy levels and regions

Species	Ploidy	Region	*A*	*H*e
A. incana	2x	Europe	3.77	0.437
A. glutinosa total	2x + 4x	Europe + North Africa	6.70	0.643
A. glutinosa	2x	Europe	6.09 a	0.611a
	2x	North Africa	4.67 b	0.522 a
	4x	Iberia	7.97 c	0.721 b
	4x	Balkan	8.50 c	0.730 b

Means within a column followed by different superscript letter in *A. glutinosa* were significantly different at P < 0.05 in multiple range Tukey's test.

Origin of polyploids and gene exchange between diploids and polyploids

PCoA clearly separated *A. glutinosa* and *A. incana* as two distinct species irrespective of the ploidy level within *A. glutinosa* (Fig. 3). The same was true for the STRUCTURE analyses of the whole dataset. Similarity coefficients indicated that two and three clusters best explained the genetic structuring of *Alnus* populations (Supplementary Data Fig. S1). Two clusters of individuals clearly separated the two

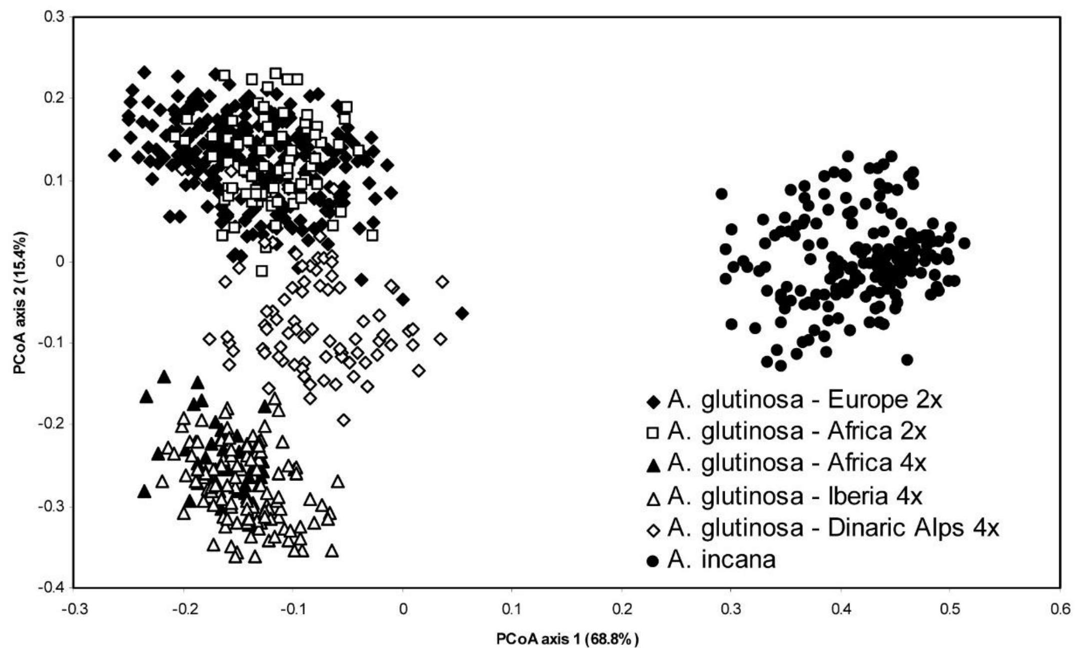


Fig. 3 PCoA using Bruvo distances performed in POLYSAT for *Alnus incana* and different cytotypes of *A. glutinosa* (see also Fig. 1). The percentage of variance explained by each axis is provided within the figure.

species irrespective of the ploidy level (Fig. 4A), and three clusters corresponded to (i) *A. incana*, (ii) diploid populations of *A. glutinosa* from Europe and North Africa and tetraploid populations from the Dinaric Alps, and (iii) tetraploid *A. glutinosa* populations from the Iberian Peninsula and North Africa (Fig. 4A). The hypothesis that *A. glutinosa* is of allopolyploid origin and that *A. incana* is one of the parental

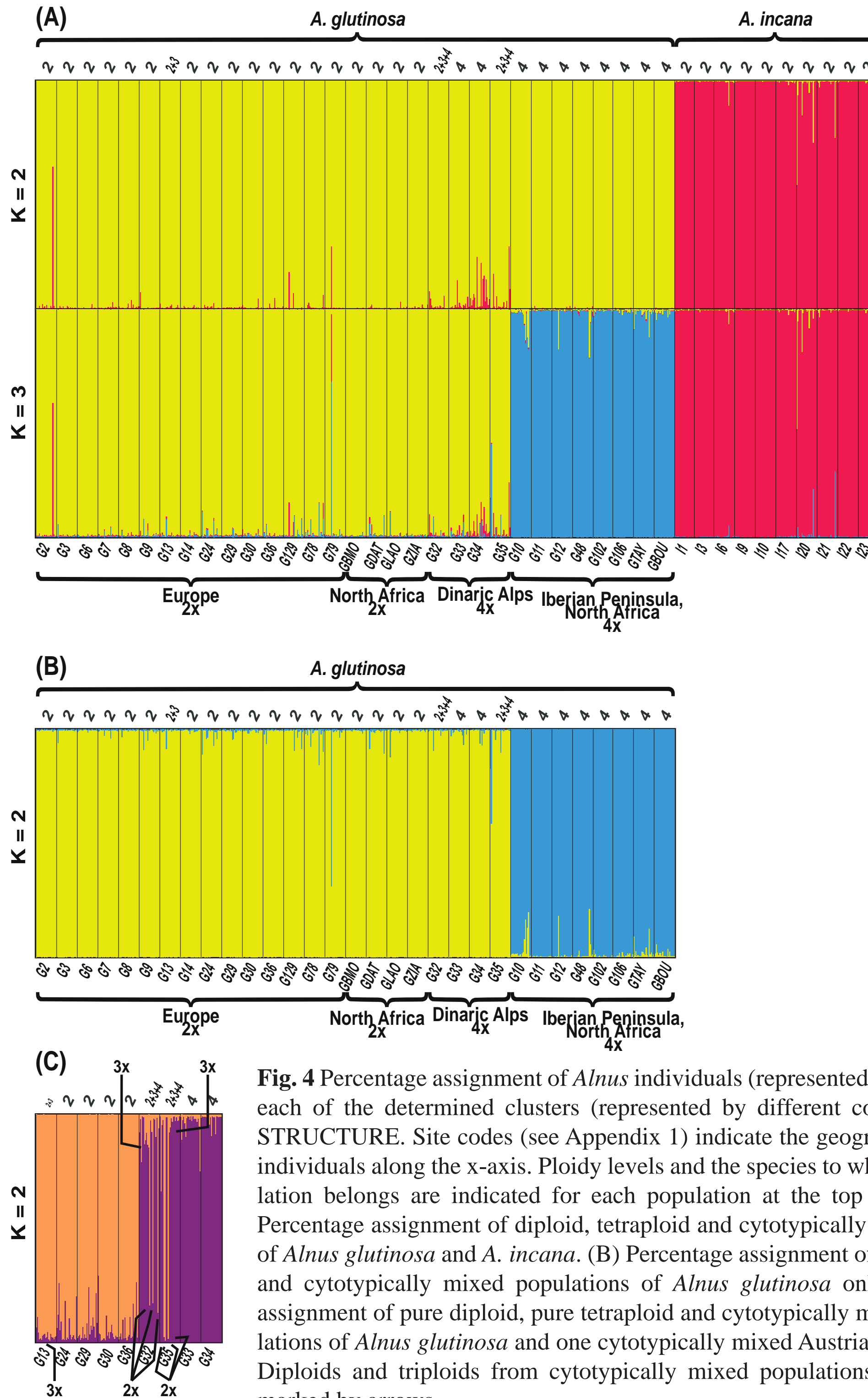


Fig. 4 Percentage assignment of *Alnus* individuals (represented by vertical bars) to each of the determined clusters (represented by different colours) inferred by STRUCTURE. Site codes (see Appendix 1) indicate the geographical location of individuals along the x-axis. Ploidy levels and the species to which the each population belongs are indicated for each population at the top of the figure. (A) Percentage assignment of diploid, tetraploid and cytotypically mixed populations of *Alnus glutinosa* and *A. incana*. (B) Percentage assignment of diploid, tetraploid and cytotypically mixed populations of *Alnus glutinosa* onlx. (C) Percentage assignment of pure diploid, pure tetraploid and cytotypically mixed Balkan populations of *Alnus glutinosa* and one cytotypically mixed Austrian populations (13). Diploids and triploids from cytotypically mixed populations (13,, 32, 35) are marked by arrows.

species is therefore not supported, suggesting that tetraploid cytotypes of *A. glutinosa* are probably of autopolyploid origin.

Separate analyses of all *A. glutinosa* populations yielded a similar pattern. The similarity coefficients indicated that two clusters best explained the genetic structuring of *A. glutinosa* populations (Supplementary Data Fig. S2). Two clusters of individuals clearly separated Iberian and North-African tetraploid populations from the rest of European diploid and Dinaric tetraploid populations (Fig. 4B).

The analysis of Balkan populations combining exclusively diploid and tetraploid populations together with mixed populations comprising di, tri- and tetraploids individuals and one Austrian population (13) comprising diploids and a triploid, showing that two clusters best explained the genetic structuring of these populations (Supplementary Data Fig. S3). The analysis clearly separated diploid and tetraploid populations as well as individuals of different ploidy levels from cytotypically mixed populations (Fig. 4C). All triploids in Balkan populations had a high proportion of the tetraploid genetic cluster (Fig. 4C), indicating that they have not arisen by fusion of reduced and unreduced gametes of diploids, but by hybridization of diploid and tetraploid individuals. In this case, tetraploids provided two-thirds of the genome. This was reflected in the STRUCTURE analysis, which assigned triploids to the tetraploid genetic cluster. A different origin was detected in the case of triploid individual from the Austrian population assigned to the diploid genetic cluster (Fig. 4C). This means that the triploid evolved by fusion of reduced and unreduced gametes of diploid individuals with no participation of tetraploids.

Palaeo-distribution modelling

A high value of the operating characteristic curve (AUC = 0.967) indicated a better than random prediction of tetraploid distribution based on nine BIOCLIM variables. A prediction of the recent range showed that the area with suitable conditions for tetraploid populations is larger than their current occurrence, also including large areas of temperate Europe (Figs. 5, 6). On the contrary, LGM models suggested that the distribution of tetraploids was limited mainly to the Mediterranean area (Figs. 5, 6).

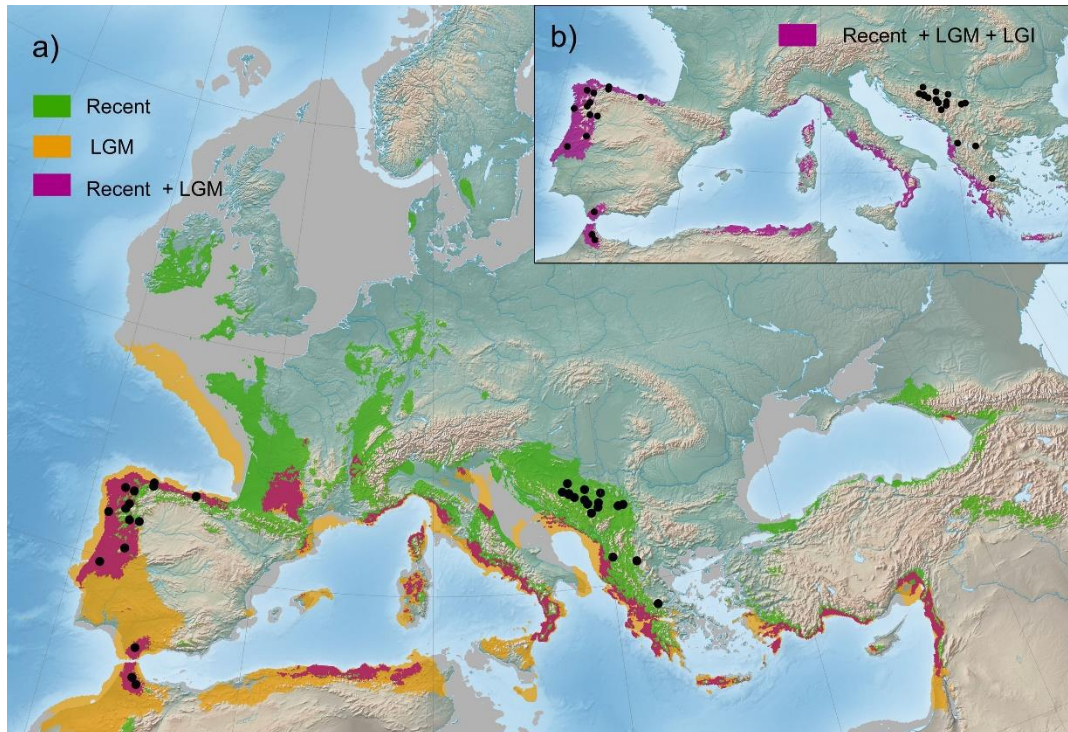


Fig. 5 Distributions of *Alnus glutinosa* tetraploids since a) the Last Glacial Maximum (LGM, *c*. 21 ka) and b) the Last Interglacial period (LIG, *c*. 115 ka) predicted based on the MIROC climatic model. The equal training sensitivity and specificity threshold of 0.143 was used for delimiting the area with suitable climatic conditions. "Recent + LGM" and "Recent + LGM + LGI" mark areas predicted as suitable in more than one time period. Grey in coastal zones indicates land areas above the see line during the LGM.

Balkan and Dinaric populations were retreated to the western coastal area of the Balkan Peninsula based on both the MIROC and the CCSM climatic models (Figs. 5, 6). Both LGM climatic models supported a continuous distribution of tetraploids in lowlands of southern Portugal and Spain connecting African and Iberian populations (Figs. 5, 6). The occurrence of tetraploids in northern Portugal, north-eastern Spain and southern France was predicted only by the MIROC climatic model (Fig. 5).

An overlap among recent, LGM and LIG predictions based on the MIROC model indicated suitable conditions for survival of tetraploid populations in the Balkan, Iberian and Apennine Peninsulas and northern Africa, at least since the Last Interglacial period (Fig. 5).

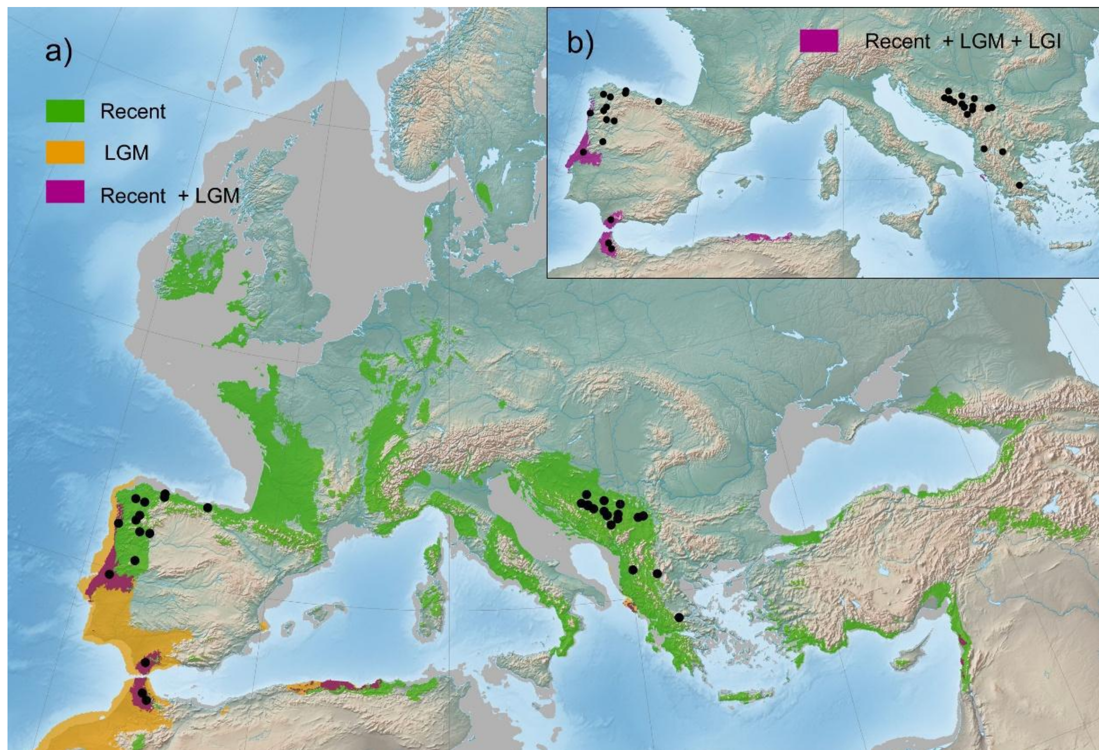


Fig. 6 Distributions of *Alnus glutinosa* tetraploids since a) the Last Glacial Maximum (LGM, *c*. 21 ka) and b) the Last Interglacial period (LIG, *c*. 115 ka) predicted based on the CCSM climatic model. The equal training sensitivity and specificity threshold of 0.143 was used for delimiting the area with suitable climatic conditions. “Recent + LGM” and “Recent + LGM + LGI” mark areas predicted as suitable in more than one time period. Grey in coastal zones indicates land areas above the see line during the LGM.

DISCUSSION

Distribution and origin of tetraploid cytotypes

Analysis of ploidy-level variation in populations of *A. glutinosa* all across Europe combined with microsatellite analysis and palaeodistribution modelling allowed us to infer the distribution, origins and survival of different cytotypes of *A. glutinosa*. *Alnus glutinosa* has been considered a diploid species by a large number of cytological studies from many European countries (e.g. Fedorov, 1969; Goldblatt, 1981, 1985, 1988; Goldblatt and Johnson, 1990, 1991, 1994, 2000, 2003; Ivanova *et al.*, 2006). The first indication that this might not be entirely true has been brought up by Woodworth (1929, 1931), who was later cited by Fedorov (1969). Woodworth for the first time discovered tetraploid individuals of *A. glutinosa*. He, however, did not work with European samples directly, but collected the material for his studies at the

Arnold Arboretum of Harvard University in Boston. *Alnus glutinosa*, being a species alien to North America, had to be introduced to the Arnold Arboreum from somewhere. Unfortunately, it is impossible to identify the location of the source population. Source trees that were growing at the Arnold Arboretum around the period when Robert H. Woodworth worked there are no longer alive, and there is no chance of determining the source population by genetic means (Michael S. Dosmann, Curator of Living Collections, The Arnold Arboretum of Harvard University, letter correspondence). The possibility that *A. glutinosa* has more than one cytotype in Europe has been completely overlooked, probably because there are so many diploid chromosome counts and because Woodworth's counts seem doubtful. A second study pointing out the existence of tetraploid *A. glutinosa* has recently been published by Lepais *et al.* (2013), who report tetraploid populations from Morocco based on microsatellite typing. Lepais *et al.* (2013) postulate that these relictual populations could have evolved by hybridization between differentiated *A. glutinosa* populations and subsequent polyploidization at the southern edge of the species' distribution range. In this study, we for the first time report the occurrence of tetraploid populations of *A. glutinosa* in Europe. The distribution of tetraploids is far from random, as it forms two geographically very well delimited populations. The first is situated in the Iberian Peninsula, extending to North Africa, where it has previously been reported by Lepais *et al.* (2013). The second one lies in the Dinaric Alps, extending to south-western Greece. Bayesian clustering analysis revealed a clear pattern of genetic structure spanning *A. incana* as well as both diploid and tetraploid populations of *A. glutinosa*. PCoA analysis yielded the same results. Both tetraploid populations are therefore probably of autopolyploid origin with no indication that *A. incana* has been involved in their evolutionary history.

Establishment and maintenance of polyploids

Two scenarios have been proposed to explain differences in patterns of cytotype distributions. First, the adaptive evolutionary scenario is based on the assumption that newly arising polyploids possess novel genetic combinations enabling them to thrive under a wider range of ecological conditions (Levin, 1983; Soltis and Soltis,

1993; Otto and Whitton, 2000; Soltis *et al.*, 2003). It has been repeatedly shown that polyploid populations are able to successfully colonize different niches than those inhabited by their progenitors (Flégrová and Krahulec, 1999; Ramsey, 2011; Mráz *et al.*, 2012; Collins *et al.*, 2013; Hahn *et al.*, 2013). Due to ecological sorting along abiotic or biotic gradients, polyploids can occupy different ecological niches, resulting in spatial segregation within diploid-polyploid complexes. The other scenario is the non-adaptive scenario (the so-called "minority cytotype exclusion model", Levin, 1975; Fowler and Levin, 1984; Ramsey and Schemske, 1998) is used to explain spatial segregation by frequency-dependent mating success that results from low fitness of hybrids formed from between-cytotype mating gradually leading to the elimination of the minority cytotype from the population. In this case, cytologically uniform populations occur in different locations, and the coexistence of multiple cytotypes is viewed as a transient phenomenon.

Although our knowledge concerning the distribution of *A. glutinosa* cytotypes in the field does not allow an accurate distinction between the adaptive and non-adaptive scenarios (Levin, 1975; Husband *et al.*, 2013), many studies of other groups provide information about contact zones between or among cytotypes. There are two types of contact zones: narrow contact zones with only a low number of cytotypically mixed populations (Husband and Schemske, 1998; Baack, 2004) and contact zones extending over large areas (Burton and Husband, 1999; Halverson *et al.*, 2008; Duchoslav *et al.*, 2010, Fialová *et al.*, 2014). Even in areas where the distribution of two cytotypes is diffuse, mixed-cytotype populations tend to be quite rare (Kolář *et al.*, 2009; Trávníček *et al.*, 2010; Castro *et al.*, 2012). Contact zones may result from secondary contact between previously allopatric chromosomal races (secondary contact zones; Petit *et al.*, 2002) or the expansion of newly formed polyploids from within diploid populations. It is, however, difficult to apply the concept of secondary contact zones to *A. glutinosa* cytotypes. In our view, the distribution of its cytotypes has been formed by past climatic changes during glacial and interglacial times.

In the case of *A. glutinosa*, the distribution areas of diploids and tetraploids overlap only to a small degree in the Dinaric Alps. Diploid and tetraploid populations are almost parapatric, suggesting differences in ecological tolerance. Whereas pure tetraploid populations occur almost exclusively at the bottoms of

deep valleys, diploids are distributed at lower altitudes around core tetraploid populations. The observed pattern might be explainable by variation in ecological tolerance, as polyploids have a wider spectrum of tolerance and are adapted to ecological conditions not suitable for diploids (Levin, 1983).

The small number of triploids found suggests restricted gene exchange and increasing reproductive isolation. However, the occurrence of triploids was detected in both (i) populations situated within the range of diploids, i.e. locality 13 in Austria, and (ii) on the border between diploid and tetraploid populations ranges, i.e. localities 32 and 35. In the case of Austrian population 13, the triploid individual originated by fusion of reduced and unreduced gametes, as has been shown by the STRUCTURE analysis. On the other hand, triploids in Balkan populations originated by hybridization between di- and tetraploid individuals. We have thus demonstrated restricted gene flow between diploids and tetraploids that could probably have occurred anywhere individuals of different ploidy levels are in contact.

Glacial refugia for Alnus glutinosa and ploidy level distribution

Two studies consider the postglacial migration of *A. glutinosa* in Europe (King and Ferris, 1998; Douda *at al.*, 2014). The first looked at variation in chloroplast DNA to reveal main migration routes of the species after the retreat of glaciers and located LGM refugia in the Carpathians. Douda *et al.* (2014), by contrast, showed, based on a review of radiocarbon-dated pollen and macrofossil sites for species of the subgenus *Alnus* from the last 50,000 years in Europe, that an expansion of *Alnus* is supported by the presence of alders during the LGM in southern peninsulas and northerly areas of western Europe, the foothills of the Alps, the Carpathians and northeastern Europe. However, postglacial recolonization routes have been interpreted without information on cytotype variation of *A. glutinosa* in Europe. Tetraploids were found in two separate populations located in two important glacial refugia, one in the Iberian Peninsula and the second covering the western part of the Balkan Peninsula. The results of this study corroborate previous studies indicating that northward postglacial expansion from the Iberian Peninsula is unlikely and also rules the western part of the Balkan Peninsula as a putative source

refugium. Therefore, diploid populations located in southern European peninsulas that might be taken into consideration for northward expansion are located only along the border between the Iberian Peninsula and Europe, i.e. in the Pyrenees, in the eastern part of the Balkan Peninsula, in the Apennine Peninsula and in more northern refugial areas suggested by Douda *et al.* (2014). To sum up, because some putative glacial refugia harbour the tetraploid cytotype, these areas could not have served as effective refugia for *A. glutinosa* diploids growing in the rest of Europe.

Ecological niche models and tetraploid glacial refugia

Areas predicted as currently suitable for tetraploids of *A. glutinosa* by ecological niche model projections onto current climate layers cover a larger area than the actual distribution of the species. Iberian tetraploid populations might occur in western France, southern Ireland and western Great Britain, and also in the foothills of the Alps and certain areas of Germany. The same is true for tetraploids of the Dinaric Alps, whose predicted area covers the whole of former Yugoslavia, extending to south-eastern Greece. The wider distribution predicted by the model compared to the actual situation might not be simply explainable by sparse sampling in the predicted areas because the whole Europe was sampled quite densely and evenly. One reason for this mismatch might reside in the fact that species distribution models do not consider information on species history, dispersal ability, water availability, soil types, competing vegetation or genetic diversity, and assume a species-climate equilibrium, namely that species occupy all environmentally suitable areas (Nogués-Bravo, 2009; Angert *et al.*, 2011; Abeli *et al.*, 2013).

The projection of the distribution model into the LGM situated the distribution range of the tetraploid cytotype exclusively in the Mediterranean area. It has been shown that populations found in the Iberian Peninsula and North Africa are likely to had been interconnected during the LGM, which is also in agreement with the genetic structure – tetraploid populations from the Iberian Peninsula form the same genetic cluster as populations from North Africa. Moreover, some Iberian populations were probably present at localities of continuous occurrence during the last glacial period. This should support the view that formerly recognized refugia

did not constitute single refugia throughout the Quaternary, but instead several independent or interconnected “refugia within refugia” (Gómez and Lunt, 2006; Nieto-Feliner, 2011; Fernández-Mazuecos and Vargas, 2013). Hence, as Gutiérrez-Larena *et al.* (2002) or Martín-Bravo *et al.* (2010) argued in the case of the Quaternary history of Mediterranean species, altitudinal migrations were more likely than large-scale range shifts. Long-term isolation within each refugium led to the evolution of genetically very distinct tetraploid populations in Europe and North Africa as well as diploid populations in North Africa.

However, a different picture might be drawn in the case of tetraploid populations occurring in the Dinaric Alps. Most of them probably did not exist throughout the last glacial period and found refuge south in lowland and coastal areas of the Balkan Peninsula. By combining data from ecological niche models and Bayesian analyses of microsatellite data, it is therefore possible to postulate two contrasting scenarios for each area. While some populations in the Iberian Peninsula were probably stable over a very long period of time, tetraploid populations in the Dinaric Alps withdrew during glacial times and expanded to new suitable localities in interglacial times. These populations now survive at the bottom of deep river valleys with specific micro-climatic conditions, for example, in the canyons of the rivers Tara and Mrtvica Rivers in Montenegro.

CONCLUSIONS

Our cytological and molecular data combined with species distribution modelling brings new insights into the cytological variation, cytotype origin and late Quaternary history of *A. glutinosa*. Newly discovered tetraploid populations in the Iberian Peninsula and the Dinaric Alps are likely to have persisted there for a long time. Both tetraploid clusters are situated in the putative main glacial refugia, and neither of them was probably involved in the colonization of Central and Northern Europe after glacial withdrawal. This means that neither the Iberian Peninsula nor the western part of the Balkan Peninsula served as effective refugial areas for northward postglacial expansion because the ploidy levels are different. The diploid populations currently growing in most of Europe are therefore likely to have descended from populations located in western Europe and the eastern Balkan

Peninsula or the Apennine Peninsula. This result partly supports the outcomes of King and Ferris (1998), who placed the main refugium of *A. glutinosa* in the Carpathians, and Douda *et al.* (2014), who emphasized the importance of easterly located refugia (i.e. Belarus and western Russia) and also individual regional refugia serving as important centres of local spreading.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Fig. S1–3: Average similarity coefficients for each *K* with standard deviations.

ACKNOWLEDGEMENTS

We would like to thank Martina Hadincová and Eva Ničová for their help in the lab and Fred Rooks for helping us out with the language of the manuscript. Michael S. Dosmann kindly informed us about the history of *Alnus* accessions growing at the Arnold Arboreum of Harvard University. We also wish to thank all people who collected samples of *Alnus* for us, especially Filip Kolář, Vít Dvořák, Milan Štech and Boudewijn Michiels. Serge Müller deserves thanks for sharing samples from North Africa. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated. This study was supported by the Grant Agency of the Czech Republic (P504/11/0402) and as part of long-term research development project (RVO 67985939) and project Postdoc USB (CZ.1.07/2.3.00/30.0006).

LITERATURE CITED

Abeli T, Gentili R, Mondoni A, Orsenigo S, Rossi G. 2013. Effects of marginality on plant population performance. *Journal of Biogeography* **41:** 239–249.

Amborella Genome Project. 2013. The Amborella Genome and the Evolution of Flowering Plants. *Science* **342:** 1241089.

Angert AL, Crozier LG, Rissler LJ, Gilman SE, Tewksbury JJ, Chunco AJ. 2011. Do species' traits predict recent shifts at expanding range edges? *Ecology Letters* **14:** 677–689.

Baack EJ. 2004. Cytotype segregation on regional and microgeographic scales in snow buttercups (*Ranunculus adoneus*: Ranunculaceae). *American Journal of Botany* **91:** 1783–1788.

Bennett KD, Tzedakis PC, Willis KJ. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* **18:** 103–115.

Brochmann C, Brysting AK, Alsos IG, *et al.* 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* **82:** 521–536.

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13:** 2101–2106.

Burton TL, Husband BC. 1999. Population cytotype structure in the polyploidy *Galax urceolata* (Diapensiaceae). *Heredity* **82:** 381–390.

Castro S, Loureiro J, Procházka T, Münzbergová Z. 2012. Cytotype distribution at a diploid–hexaploid contact zone in *Aster amellus* (Asteraceae). *Annals of Botany* **110:** 1047–1055.

Catalán P, Segarra-Moragues JG, Palop-Esteban M, Moreno C, Gonzáles-Candelas F. 2006. A Bayesian approach for discriminating among alternative inheritance hypotheses in plant polyploids: The allotetraploid origin of genus *Borderea* (Dioscoreaceae). *Genetics* **172:** 1939–1953.

Clark LV, Jasieniuk M. 2011. POLYSAT: and R package for polyploid microsatellite analysis. *Molecular Ecology Resources* **11:** 526–566.

Collins AR, Thalmann D, Müller-Schärer H. 2013. Cytotypes of *Centaurea stoebe* found to differ in root growth using growth pouches. *Weed Research* **53:** 159–163.

Collins WD, Bitz CM, Blackmon ML, *et al.* 2006. The Community Climate System Model Version 3 (CCSM3). *Journal of Climate* **19:** 2122–2143.

Cook LM, Soltis PS, Brunsfeld SJ, Soltis DE. 1998. Multiple independent formations of *Tragopogon* tetraploids (Asteraceae): evidence from RAPD markers. *Molecular Ecology* **7:** 1293–1302.

Cruzan MB, Templeton AR. 2000. Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution* **15:** 491–496.

Doležel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2:** 2233–2244.

Douda J, Doudová J, Drašnarová A, *et al.* 2014. Migration patterns of subgenus *Alnus* in Europe since the Last Glacial Maximum. A systematic review. *PLoS ONE* **9:** e88709. doi:10.1371/journal.pone.0088709.

Douda J, Doudová-Kochánková J, Boublík K, Drašnarová A. 2012. Plant species coexistence at local scale in temperate swamp forest: test of habitat heterogeneity hypothesis. *Oecologia* **169:** 523–534.

Doyle JJ, Flagel LE, Paterson AH, *et al.* 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* **42:** 443–461.

Drašnarová A, Krak K, Vít P, *et al.* 2014. Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genetics and Genomes* **10:** 865–873.

Duchoslav M, Šafářová L, Krahulec F. 2010. Complex distribution patterns, ecology and coexistence of ploidy levels of *Allium oleraceum* (Alliaceae) in the Czech Republic. *Annals of Botany* **105:** 719–735.

Ehrich D, Gaudeul M, Assefa A, *et al.* 2007. Genetic consequences of Pleistocene range shifts: Contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* **16:** 2542–2559.

Elith J, Graham CH, Anderson RP, *et al.* 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29:** 129–151.

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7:** 574–578.

Fedorov AA (ed). 1969. Chromosome Numbers of Flowering Plants. Nauka, Leningrad.

Fernández-Mazuecos M, Vargas P. 2013. Congruence between distribution modelling and phylogeographical analyses reveals Quaternary survival of a toadflax species (*Linaria elegans*) in oceanic climate areas of a mountain ring range. *New Phytologist 198:* 1274–1289.

Fialová M, Jandová M, Ohryzek J, Duchoslav M. 2014. Biology of the polyploid geophyte *Allium oleraceum* (Amaryllidaceae): Variation in size, sexual and asexual reproduction and germination within and between tetra-, penta- and hexaploid cytotype. *Flora* **209:** 312–324.

Flégrová M, Krahulec F. 1999. *Anthoxanthum odoratum* and *A. alpinum* : Life history parameters at two different altitudes. *Folia Geobotanica* **34:** 19–31.

Fowler NL, Levin HS. 1984. Ecological constrains on the establishment of a novel polyploidy in competition with its diploid progenitor. *American Naturalist* **124:** 703–711.

Goldblatt P (ed). 1981. Index to plant chromosome numbers 1975–1978. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 5.

Goldblatt P (ed). 1984. Index to plant chromosome numbers 1979–1981. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 8.

Goldblatt P (ed). 1985. Index to plant chromosome numbers 1982–1983. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 13.

Goldblatt P (ed). 1988. Index to plant chromosome numbers 1984–1985. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 23.

Goldblatt P, Johnson DE (eds). 1990. Index to plant chromosome numbers 1986–1987. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 30.

Goldblatt P, Johnson DE (eds). 1991. Index to plant chromosome numbers 1988–1989. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 40.

Goldblatt P, Johnson DE (eds). 1994. Index to plant chromosome numbers 1990–1991. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 51.

Goldblatt P, Johnson DE (eds). 1996. Index to plant chromosome numbers 1992–1993. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 58.

Goldblatt P, Johnson DE (eds). 1998. Index to plant chromosome numbers 1994–1995. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 69.

Goldblatt P, Johnson DE (eds). 2000. Index to plant chromosome numbers 1996–1997. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 81.

Goldblatt P, Johnson DE (eds). 2003. Index to plant chromosome numbers 1998–2000. *Monographs in Systematic Botany from the Missouri Botanical Garden*, 94.

Gómez A, Lunt DH. 2006. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N, eds. *Phylogeography of southern European refugia*. Dordrecht, the Netherlands: Springer, 155–188.

Grant V. 1981. Plant speciation. New York, NY, USA: Columbia University Press.

Greilhuber J, Doležel J, Lysák MA, Bennett MD. 2005. The origin, evolution and proposed stabilization of the terms "genome size" and "C-value" to describe nuclear DNA contents. **Annals of Botany 95:** 255–260.

Greilhuber J. 2005. Intraspecific variation in genome size in Angiosperms: identifying its existence. **Annals of Botany 95:** 91–98.

Gutiérrez-Larena B, Fuertes-Aguilar J, Nieto-Feliner G. 2002. Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. *Molecular Ecology* **11:** 1965–1974.

Hagerup O. 1931. Über Polyploidie in Beziehung zu Klima, Ökologie und Phylogenie. *Hereditas* **16:** 19–40.

Hahn MA, Müller-Schärer H. 2013. Cytotype differences modulate eco-geographical differentiation in the widespread plant *Centaurea stoebe*. *Ecology* **94:** 1005–1014.

Halverson K, Heard SB, Nason JD, Stireman JO. 2008. Origins, distribution, and local co-occurrence of polyploidy cytotypes in *Solidago altissima* (Asteraceae). *American Journal of Botany* **95:** 50–58.

Hardy OJ, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2:** 618–620.

Hasumi H, Emori S. 2004. K1-coupled GCM (MIROC) description. K1 Technical Report No. 1. Tokyo: Center for Climate System Research.

Hewitt G.M. 1996. Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58:** 247–276.

Hewitt G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International *Journal of Climatology* **25:** 1965–1978.

Husband BC, Baldwin SJ, Suda J. 2013. The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. In: Greilhuber J, Doležel J, Wendel JF, eds. *Plant genome diversity*. Vienna: Springer, 2: 255–276.

Husband BC, Schemske DW. 1998. Cytotype distribution at a diploid-tetraploid contact zone in *Chamerion* (*Epilobium*) *angustifolium* (Onagraceae). *American Journal of Botany* **85:** 1688–1694.

Ivanova D, Dimitrova D, Vladimirov V. 2006. Chromosome numbers of selected woody species from the Bulgarian flora. *Phytologia Balcanica* **12:** 79–84.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801–1806.

Jalas J, Suominen J (eds). 1976. Atlas Florae Europaeae. Distribution of Vascular Plants in Europe. 3. Salicaceae to Balanophoraceae. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

Johnson AW, Packer JG. 1965. Polyploidy and environment in arctic Alaska. *Science* **148:** 237–239.

King RA, Ferris C. 1998. Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn., *Molecular Ecology* **7:** 1151–1161.

Kochert G, Stalker HT, Gimenes M, Galgaro L, Lopes CR, Moore K. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae), *American Journal of Botany* **83:** 1282–1291.

Kolář F, Fér T, Štech M, *et al.* 2012. Bringing Together Evolution on Serpentine and Polyploidy: Spatiotemporal History of the Diploid-Tetraploid Complex of *Knautia arvensis* (Dipsacaceae). *PlosONE* **7:** e39988.

Kolář F, Štech M, Trávníček P, *et al.* 2009. Towards resolving the *Knautia arvensis* agg. (Dipsacaceae) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales. *Annals of Botany* **103:** 963–974.

Leitch AR, Leitch IJ. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* **320:** 481–483.

Lepais O, Bacles CFE. 2011. De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. *Journal of Heredity* **102:** 627–632.

Lepais O, Muller SD, Ben Saad-Limam S, *et al.* 2013. High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PLoS ONE* **8:** e75029. doi:10.1371/journal.pone.0075029.

Levin DA. 1975. Minority cytotype exclusion in local plant populations. *Taxon* **24:** 35–43.

Levin DA. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* **122:** 1–25.

Liu C, Pam M, Dawson TP, Pearson RG. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28:** 385–393.

Löve Á, Löve D. 1957. Arctic polyploidy. *Proceedings of the Genetics Society of Canada* **2:** 23–27.

Lowe AJ, Abbott RJ. 1996. Origins of the new allopolyploid species S*enecio cambrensis* (Asteraceae) and its relationship to the Canary Islands endemic *Senecio teneriffae*, *American Journal of Botany* **83:** 1365–1372.

Magri D, Vendramin GG, Comps B, *et al.* 2006. A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist* **171:** 199–221.

Mahy G, Bruederle LP, Connors B, van Hofwegen M, Vorsa N. 2000. Allozyme evidence for genetic autopolyploidy and high genetic diversity in tetraploid

cranberry, *Vaccinium oxycoccos* (Ericaceae). *American Journal of Botany* **87:** 1882–1889.

Mandák B, Trávníček P, Paštová L, Kořínková D. 2012. Is hybridization involved in the evolution of the *Chenopodium album* aggregate? An analysis based on chromosome counts and genome size estimation. *Flora* **207:** 530–540.

Mandáková T, Münzbergová Z. 2006. Distribution and ecology of cytotypes of the *Aster amellus* aggregates in the Czech Republic. *Annals of Botany* **98:** 845–856.

Manzaneda AJ, Rey PJ, Bastida JM, Weiss-Lehman C, Raskin E, Mitchell-Olds T. 2011. Environmental aridity is associated with cytotype segregation and polyploidy occurrence in *Brachypodium distachyon* (Poaceae). *New Phytologist* **193:** 797–805.

Martín-Bravo S, Valcárcel V, Vargas P, Luceño M. 2010. Geographical speciation related to Pleistocene range shifts in the western Mediterranean mountains (*Reseda* sect. *Glaucoreseda*, Resedaceae). *Taxon* **59:** 466–482.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in the majority of angiosperms. *Science* **264:** 421–423.

McVean DN. 1953. Account of *Alnus glutinosa* (L.) Gaertn. for the Biological Flora of the British Isles. *Journal of Ecology* **41:** 447–466.

Mráz P, Španiel S, Keller A, Bowmann G, Farkas A, Šingliarová B, Rohr RP, Broennimann O, Müller-Schärer H. 2012. Anthropogenic disturbance as a driver of microspatial and microhabitat segregation of cytotypes of *Centaurea stoebe* and cytotype interactions in secondary contact zones. *Annals of Botany* **110:** 615–627.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89:** 583–590.

Nieto-Feliner G. 2011. Southern European glacial refugia: a tale of tales. *Taxon* **60:** 365–372.

Nogués-Bravo D. 2009. Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography* **18:** 521–531.

Nordborg M, Hu TT, Ishino Y, *et al.* 2005. The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol* **3:** e196. 10.1371/journal.pbio.0030196.

Otto F. 1990. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman H, Darzynkiewicz AZ, eds. *Methods in cell biology*. New York: Academic Press, 33: 105–110.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review Genetics* **34:** 401–437.

Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell* **131:** 452–462.

Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. 2006. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311:** 1751–1753.

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

Parisod C, Holderegger R, Brochmann C. 2010. Evolutionary consequences of autopolyploidy. *New Phytologist* **186:** 5–17.

Peterson AT, Papes M, Soberón J. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modelling. *Ecological Modelling* **213:** 63–72.

Petit RJ, Brewer S, Bordacs S, *et al.* 2002. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management* **156:** 49–74.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945 959.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploidy formation in flowering plants. *Annual Review of Ecology and Systematics* **29:** 467–501.

Ramsey J. 2011. Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences, USA* **108:** 7096–7101.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137–138.

Schönswetter P, Suda J, Popp M, Weiss-Schneeweiss H, Brochmann C. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP Wngerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and Evolution* **42:** 92–103.

Segraves KA, Thompson JN, Soltis PS, Soltis DE. 1999. Multiple origins of polyploidy and the geographic structure of *Heuchera grossulariifolia*. *Molecular Ecology* **8:** 253–262.

Soltis DE, Soltis PS, Schemske DW, *et al.* 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* **56:** 13–30.

Soltis DE, Soltis PS, Tate JA. 2003. Advances in the study of polyploidy since Plant speciation. *New Phytologist* **161:** 173–191.

Soltis DE, Soltis PS. 1993. Molecular data and the dynamics nature of polyploidy. *Critical Reviews in Plant Sciences* **12:** 348–352.

StatSoft, Inc. 2013. STATISTICA (data analysis software system), version 12. www.statsoft.com.

Stebbins GL. 1950. Variation and evolution in plants. New York, USA: Columbia University Press.

Stebbins GL. 1971. Chromosomal evolution in higher plants. London, UK: Edward Arnold.

Stebbins GL. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. *Botanica Helvetica* **94:** 1–13.

Stebbins GL. 1985. Polyploidy, hybridization, and the invasion of new habitats. *Annals of the Missouri Botanical Garden* **72:** 824–832.

Suda J, Krahulcová A, Trávníček P, Krahulec F. 2006. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxon* **55:** 447–450.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cossons J-F. 1998. Comparative phylogeography and postaglacial colonization routes in Europe. *Molecular Ecology* **7:** 453–464.

Tallantire PA. 1974. The palaeohistory of the grey alder (*Alnus incana* (L.) Moench) and black alder (*A. glutinosa* (L.) Gaertn.) in Fennoscandia. *New Phytologist* **73:** 529–546.

Thompson JL, Lumaret R. 1992. The evolutionary dynamics of polyploidy plants: origin, establishment and persistence. *Trends in Ecology and Evolution* **7:** 302–306.

Tollefsrud MM, Kissling R, Gugerli F, *et al.* 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.

Trávníček P, Eliášová A, Suda J. 2010. The distribution of cytotypes of *Vicia cracca* in Central Europe: the changes that have occurred over the last four decades. *Preslia* **82:** 149–163.

Tzedakis PC, Emerson BC, Hewitt G.M. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology and Evolution* **28:** 696–704.

Wachowiak W, Iason GR, Cavers S. 2013. Among population differentiation at nuclear genes in native Scots pine (*Pinus sylvestris* L.) in Scotland. *Flora* **208:** 79–86.

Wendel JF. 2000. Genome evolution in polyploid. *Plant Molecular Biology* **42:** 225–249.

Woodworth RH. 1929. Cytological studies in Betulaceae. II. *Corylus* and *Alnus*. *Botanical Gazette* **88:** 383–399.

Woodworth RH. 1931. Polyploidy in the Betulaceae. *Journal of the Arnold Arboretum* **12:** 206–217.

Appendix 1 Summary of geographic locations and genetic diversities of *Alnus glutinosa* and *A. incana* populations. Sp: species, G – *Alnus glutinosa*, I – *A. incana*; Pop: the number of population in our database – populations with the prefix L and KF are related to samples previously published elsewhere, i.e. L – Lepais *et al.* (2013), KF – King & Ferris (1998); Ploidy: ploidy level determined by flow cytometry, i.e. 2x = 28, 3x = 42, 4x = 56; Latitude, Longitude (GPS coordinates of sampling sites are in WGS84); N: number of individuals sampled from each population; SS: Sampling strategy, DS – detailed sampling, CS – coarse sampling, SC – samples from seed collections (see Materials and Methods for details); Country: country abbreviation; Locality: population location; M: population used for microsatellite analysis; *A*: average number of alleles; *He*: average gene diversity.

Sp	Pop	Ploidy	LAT	LON	N	SS	Country	Locality	M	*A*	*He*
I	1		46.049572	11.994668	20	DS	ITA	Busche	M	3.42	0.398
I	3		44.606500	9.298157	20	DS	ITA	Gorreto	M	3.58	0.423
I	6		46.722217	13.635433	20	DS	AUT	Paternion	M	4.33	0.469
I	9		45.041896	6.050286	20	DS	FRA	Le Bourg-D´Oisans	M	2.83	0.353
I	10		46.047280	6.763980	20	DS	CHE	Sixt-Fer- à-Cheval	M	2.83	0.361
I	17		46.950255	22.723963	20	DS	ROU	Negreni	M	4.75	0.499
I	20		44.993851	25.284784	20	DS	ROU	Tatarani	M	4.50	0.542
I	21		42.006754	23.640078	20	DS	BGR	Yakoruda	M	4.00	0.457
I	22		43.144686	19.297479	20	DS	MNE	Žabljak	M	3.92	0.447
I	23		44.184260	17.742016	16	DS	BIH	Vitez	M	3.50	0.417
G	1	2x	49.009135	14.436995	20	DS	CZE	České Budějovice			
G	2	2x	46.135526	12.949137	20	DS	ITA	San Daniele del Friuli	M	5.67	0.560
G	3	2x	46.035783	11.987871	20	DS	ITA	Busche	M	5.92	0.611
G	4	2x	46.266120	11.062970	20	DS	ITA	Denno			
G	5	2x	44.197562	11.602491	20	DS	ITA	Casola Valsenio			
G	6	2x	43.525650	11.685982	20	DS	ITA	Montalto	M	5.58	0.596
G	7	2x	44.606015	9.296819	20	DS	ITA	Gorreto	M	6.50	0.633
G	8	2x	43.029703	0.787468	20	DS	FRA	Aspet	M	5.25	0.573
G	9	2x	43.014991	−1.507978	20	DS	ESP	Eugi	M	5.33	0.595
G	10	4x	43.332940	−3.970437	20	DS	ESP	Vargas	M	7.92	0.733
G	11	4x	41.753617	−7.458716	20	DS	PRT	Chaves	M	7.33	0.705
G	12	4x	41.881106	−8.723562	36	DS	PRT	Vilares	M	8.08	0.724
G	13	2x, 3x	46.974117	15.487067	20	DS	AUT	Fernitz bei Graz	M	6.42	0.631
G	14	2x	46.523240	16.299652	20	DS	SVN	Gibina	M	6.00	0.609
G	15	2x	46.620793	14.179016	20	DS	AUT	Klagenfurt am Wörthersee			
G	16	2x	47.883150	12.385533	20	DS	DEU	Breitbrunn am Chimsee			

G	17	2x	48.496548	11.995551	20	DS	DEU	Bruckberg			
G	18	2x	45.794781	8.733642	20	DS	ITA	Cazzago Brabbia			
G	19	2x	44.998195	5.777764	20	DS	FRA	Laffrey			
G	20	2x	47.773332	9.921796	20	DS	DEU	Kisslegg			
G	21	2x	48.908048	9.955931	20	DS	DEU	Abtsgmünd			
G	22	2x	50.500603	13.329811	20	DS	CZE	Chomutov			
G	23	2x	50.237385	16.537688	20	DS	CZE	Orlické Záhoří			
G	24	2x	46.950346	22.724281	20	DS	ROU	Negreni	M	6.75	0.628
G	25	2x	46.568674	25.167343	20	DS	ROU	Praid			
G	26	2x	46.436926	26.370439	20	DS	ROU	Comăneşti			
G	27	2x	44.994135	25.284871	20	DS	ROU	Tatarani			
G	28	2x	42.931827	25.635955	20	DS	BGR	Vaglevtsi			
G	29	2x	41.882727	23.145251	20	DS	BGR	Simitli	M	6.42	0.609
G	30	2x	43.087282	23.380836	20	DS	BGR	Lakatnik	M	6.42	0.624
G	31	2x	41.514264	24.886083	20	DS	BGR	Rudozem			
G	32	2x, 3x, 4x	40.859622	21.487761	20	DS	GRC	Marina	M	9.00	0.721
G	33	4x	41.159971	20.192281	20	DS	ALB	Librazhd	M	9.17	0.741
G	34	4x	43.147406	19.296672	20	DS	MNE	Žabljak	M	7.92	0.727
G	35	2x, 3x, 4x	44.188385	17.737766	19	DS	BIH	Vitez	M	7.92	0.732
G	36	2x	45.444461	17.243923	20	DS	HRV	Pakrac	M	6.17	0.629
G	37	4x	44.056037	18.106082	4	CS	BIH	Čatići			
G	38	2x	51.760210	8.009907	1	CS	DEU	Beckum			
G	40	2x	52.413221	13.004794	3	CS	DEU	Potsdam			
G	42	2x	54.005942	14.701656	2	CS	POL	Międziwodzie			
G	43	2x	48.270556	19.821944	4	CS	SVK	Fiľakovo			
G	44	4x	41.783111	−6.888333	3	CS	PRT	Alimonde			
G	45	2x	51.746714	−1.248544	2	CS	GBR	Oxford			
G	46	2x	50.103330	12.792170	3	CS	CZE	Krásno			
G	47	2x	50.192153	12.766139	3	CS	CZE	Loket			
G	48	4x	36.520361	−5.619806	23	DS	ESP	Alcalá de los Gazules	M	8.25	0.727
G	49	2x	43.951389	4.477194	6	CS	FRA	Collias			
G	56	2x	41.841170	33.735940	5	CS	TUR	Ersizlerdere			
G	57	2x	40.682940	39.662490	3	CS	TUR	Maçka			
G	58	2x	39.506292	23.061651	1	CS	GRC	Pouri			
G	62	2x	47.416745	9.412158	1	CS	CHE	Sankt Galen			
G	63	4x	43.955735	19.982071	13	DS	SRB	Bjeloperica			

G	64	4x	43.550256	19.779239	15	DS	SRB	Negbina			
G	65	4x	43.552636	19.130500	14	DS	BIH	Čajniče			
G	66	4x	43.753039	18.992944	19	DS	BIH	Rogatica			
G	67	3x, 4x	44.098459	18.107900	18	DS	BIH	Kakanj			
G	68	4x	44.526268	18.094453	16	DS	BIH	Maglaj			
G	69	2x	52.283306	13.453375	20	DS	DEU	Rangsdorf			
G	70	2x	52.885725	11.124123	20	DS	DEU	Salzwedel			
G	71	2x	52.882186	9.743381	20	DS	DEU	Vierde			
G	72	2x	52.544087	8.179687	20	DS	DEU	Holdorf			
G	73	2x	52.285302	5.142885	20	DS	NLD	Naarden			
G	74	2x	51.402706	−1.436850	20	DS	GBR	Kintbury			
G	75	2x	50.184217	4.742358	20	DS	BEL	Soulme			
G	76	2x	49.120367	1.646080	20	DS	FRA	Fourges			
G	77	2x	47.338355	0.436995	20	DS	FRA	Langeais			
G	78	2x	45.490310	0.478290	20	DS	FRA	Les Graulges	M	6.92	0.659
G	79	2x	45.214136	3.602951	20	DS	FRA	Sainte-Marguerite	M	5.75	0.593
G	80	2x	47.177084	3.007033	20	DS	FRA	La Charité-sur-Loire			
G	81	2x	47.948524	6.755554	20	DS	FRA	Saulxures-sur-Moselotte			
G	82	2x	50.150774	8.972370	20	DS	DEU	Erlensee			
G	83	2x	51.641773	10.631134	20	DS	DEU	Zorge			
G	84	2x	56.330425	13.003779	20	DS	SWE	Tullstorp			
G	85	2x	58.253694	11.971803	20	DS	SWE	Ljungskile			
G	86	2x	59.565325	9.191127	20	DS	NOR	Notodden			
G	87	2x	59.635095	5.904407	20	DS	NOR	Etne			
G	88	2x	63.438946	10.830525	1	CS	NOR	Muruvik			
G	89	2x	62.901204	27.603156	20	DS	FIN	Kuopio			
G	90	2x	61.334245	24.268583	20	DS	FIN	Pälkäne			
G	91	2x	60.461435	22.388690	20	DS	FIN	Turku			
G	92	2x	60.188525	17.710743	20	DS	SWE	Örbyhus			
G	93	2x	61.853668	16.622170	6	CS	SWE	Friggesund			
G	94	2x	58.331634	14.824180	20	DS	SWE	Väderstad			
G	95	2x	54.743651	11.926331	3	CS	DNK	Nykøbing Falster			
G	96	2x	60.817195	14.104830	4	CS	SWE	Norra Kättbo			
G	97	2x	41.667906	23.387050	2	CS	BGR	Sandanski			
G	98	2x	48.621684	17.159782	20	DS	SVK	Šaštín-Stráže			
G	99	2x	48.704411	19.922088	20	DS	SVK	Tisovec			

G	100	2x	48.434422	21.955853	20	DS	SVK	Krá ľovský Chlmec			
G	101	2x	49.356967	18.904203	11	DS	SVK	Klubina			
G	102	4x	39.723900	−8.507950	19	DS	PRT	Caxarias	M	8.08	0.717
G	103	4x	40.537750	−7.337490	11	CS	PRT	Vila Soeiro			
G	104	4x	42.185720	−7.818060	10	CS	ESP	Allariz			
G	105	4x	43.004230	−7.574270	10	CS	ESP	Lugo			
G	106	4x	43.529750	−6.534630	20	DS	ESP	Luarca	M	8.08	0.737
G	107	4x	43.384030	−6.530950	5	CS	ESP	Navelgas			
G	108	4x	43.057490	−8.121810	3	CS	ESP	Foro			
G	109	2x	42.031111	23.274250	1	CS	BGR	Blagovegrad			
G	110	2x	46.265492	15.867676	3	CS	HRV	Macelj			
G	111	4x	42.420550	−7.673300	3	CS	ESP	Lornís			
G	112	2x	57.426147	−6.186376	1	CS	GBR	Portree			
G	113	2x	57.065272	−5.898843	3	CS	GBR	Armadale			
G	114	2x	42.658250	11.635333	17	DS	ITA	Sovana			
G	115	2x	44.621917	22.023700	3	CS	ROU	Cozla			
G	116	2x	60.760013	32.812489	20	DS	RUS	Olonets			
G	117	2x	42.854000	1.554500	3	CS	FRA	Surba			
G	118	2x	42.718833	1.910000	4	CS	FRA	Ax-les-Thermes			
G	119	2x	42.940000	0.647833	5	CS	FRA	Chaum			
G	120	2x	44.651667	2.934333	4	CS	FRA	Aubrac			
G	121	2x	45.588667	7.374333	4	CS	ITA	Cogne			
G	122	4x	43.905850	18.411617	2	CS	BIH	Sarajevo			
G	123	4x	43.565050	19.151217	2	CS	BIH	Staronići			
G	124	2x	44.347450	17.226783	4	CS	BIH	Jajce			
G	125	2x	41.690639	41.829861	5	CS	GEO	Chakvistskali			
G	126	2x	42.145524	41.851716	3	CS	GEO	Poti			
G	128	2x	39.797806	15.890444	7	CS	ITA	Orsomarso			
G	129	2x	39.106806	16.253944	20	DS	ITA	Martirano	M	6.25	0.610
G	130	2x	42.732607	22.062353	2	CS	SRB	Vladičin Han			
G	131	2x, 4x	43.336517	21.205204	3	CS	SRB	Čučale			
G	132	4x	43.306288	20.899779	3	CS	SRB	Brzeće			
G	133	2x	43.399503	20.790895	1	CS	SRB	Jošanička Banja			
G	134	4x	43.365756	19.724729	2	CS	SRB	Prijepolje			
G	135	4x	44.188049	19.071714	3	CS	BIH	Milići			
G	136	2x	44.042583	17.374110	3	CS	BIH	Bugojno			

G	137	2x	44.881809	15.898816	3	CS	BIH	Bihać			
G	138	2x	45.943771	15.942472	3	CS	HRV	Pila			
G	139	2x	52.263695	16.800589	20	DS	POL	Mosina			
G	140	2x	52.357538	20.626312	20	DS	POL	Leszno			
G	141	2x	52.702965	23.831147	20	DS	POL	Białowieża			
G	142	2x	57.621585	25.677251	20	DS	LVA	Strenči			
G	143	2x	56.009539	37.857896	5	CS	RUS	Pushkino			
G	144	2x	57.238451	39.473761	16	DS	RUS	Rostov			
G	145	2x	54.876544	24.202453	3	CS	LTU	Rumšiškės			
G	146	2x	59.117782	25.372891	2	CS	EST	Ardu			
G	147	2x	49.441820	18.263610	3	CS	CZE	Prostřední Bečva			
G	148	2x	45.386890	13.925666	4	CS	HRV	Sovinjak			
G	149	2x	52.644944	16.034833	2	CS	POL	Sieraków			
G	150	2x	54.306444	18.308472	3	CS	POL	Żukowo			
G	152	2x	47.689385	14.293535	4	CS	AUT	Rossleithen			
G	153	2x	51.999694	31.015861	20	DS	UKR	Novi Yarylovychi			
G	154	2x	50.494778	29.535556	3	CS	UKR	Komarivka			
G	155	2x	51.144528	28.449917	3	CS	UKR	Lypnyky			
G	156	2x	51.086556	26.578861	3	CS	UKR	Malyns′k			
G	157	2x	49.936389	23.703500	7	CS	UKR	Lelekhivka			
G	158	2x	62.194511	42.776253	3	CS	RUS	Shenkursk			
G	159	2x	63.028826	42.326628	1	CS	RUS	Bereznik			
G	160	2x	64.610619	39.818228	1	CS	RUS	Severodvinsk			
G	161	2x	59.432248	39.711738	3	CS	RUS	Kubenskoye			
G	162	2x	52.730932	1.571506	20	DS	GBR	Potter Heigham			
G	163	2x	53.401544	−1.782336	20	DS	GBR	Sheffield			
G	164	2x	50.367655	16.267707	3	CS	CZE	Nový Hrádek			
G	165	2x	46.198180	13.692720	3	CS	SVN	Gabrje			
G	166	2x	51.754650	−9.568960	3	CS	IRL	Glengarriff			
G	167	2x	53.559620	−9.875070	2	CS	IRL	Letterfrack			
G	L_BMO	2x	36.843333	7.980833	20	DS	DZA	Ben Mehidi	M	4.25	0.514
G	L_DAT	2x	36.931389	7.247500	20	DS	DZA	Guerbes	M	5.25	0.527
G	L_DER	2x	36.867222	8.388889	20	DS	DZA	El Kala			
G	L_LAO	2x	36.882500	8.580833	20	DS	DZA	El Kala	M	4.50	0.540
G	L_DHA	2x	36.800833	8.659722	20	DS	TUN	Ain Draham			
G	L_TIT	2x	36.964167	8.962778	20	DS	TUN	Ouchtata			

G	L_ZIA	2x	37.189167	9.213333	20	DS	TUN	Cap Serrat	M	4.67	0.508
G	L_ZLE	2x	36.823889	8.838056	20	DS	TUN	Ain Draham			
G	L_BOU	4x	35.013611	−5.190833	20	DS	MAR	Bab Taza	M	7.67	0.722
G	L_TAY	4x	35.272778	−5.448333	20	DS	MAR	Tayenza	M	8.33	0.703
G	KF1	2x	46.681498	13.968251	1	SC	AUT	Bodensdorf			
G	KF2	2x	48.867808	15.623271	1	SC	AUT	Drosendorf an der Thaya			
G	KF3	2x	47.676915	15.936001	3	SC	AUT	Gloggnitz			
G	KF4	2x	46.804967	13.567049	1	SC	AUT	Millstatt			
G	KF5	2x	48.758053	15.946232	2	SC	AUT	Retz			
G	KF6	2x	47.957743	14.773464	1	SC	AUT	Waidhofen an der Ybbs			
G	KF7	2x	50.251579	4.659573	3	SC	BEL	Florennes			
G	KF8	2x	50.315208	4.839036	3	SC	BEL	Anheé			
G	KF9	2x	42.827547	25.484146	3	SC	BGR	Plachkovtsi			
G	KF11	2x	42.876279	25.642505	4	SC	BGR	Voneshta Voda			
G	KF12	2x	42.484251	9.044495	1	SC	FRA	Asco			
G	KF13	2x	42.015759	8.885536	3	SC	FRA	Carazzi			
G	KF14	2x	41.869914	8.978614	3	SC	FRA	Cardo-Torgia			
G	KF17	2x	42.113824	8.702188	3	SC	FRA	Sagone			
G	KF18	2x	42.984558	9.453735	3	SC	FRA	Macinaggio			
G	KF19	2x	46.039742	17.052929	3	SC	HRV	Đurđevac			
G	KF24	2x	50.142781	15.112743	3	SC	CZE	Poděbrady			
G	KF34	2x	46.267254	0.009378	3	SC	FRA	Lezay			
G	KF36	2x	50.018108	3.785419	1	SC	FRA	Le Nouvion-en-Thiérache			
G	KF39	2x	45.670489	0.107054	3	SC	FRA	Angoulême			
G	KF40	2x	47.472018	3.528215	3	SC	FRA	Clamecy			
G	KF44	2x	48.428847	8.593329	3	SC	DEU	Dettingen			
G	KF47	2x	47.566881	9.637479	3	SC	DEU	Wasserburg			
G	KF48	4x	38.915613	22.272720	1	SC	GRC	Leianokladi			
G	KF49	2x	46.119946	17.598059	3	SC	HUN	Homokszentgyörgy			
G	KF50	2x	46.235339	17.535228	2	SC	HUN	Mike			
G	KF53	2x	46.415750	17.597299	3	SC	HUN	Somogysárd			
G	KF57	2x	42.424217	12.108994	1	SC	ITA	Viterbo			
G	KF64	2x	52.801181	5.915308	2	SC	NLD	Kuinre			
G	KF65	2x	60.051616	6.538543	2	SC	NOR	Eidesåsen			
G	KF68	2x	59.301238	10.367812	3	SC	NOR	Jarlsberg			
G	KF72	2x	52.016609	23.326928	3	SC	POL	Biała Podlaska			

G	KF75	2x	54.026528	20.410194	2	SC	POL	Wichrowo			
G	KF76	2x	52.551567	21.507783	3	SC	POL	Wyszków			
G	KF84	2x	56.230184	15.629082	3	SC	SWE	Lyckeby			
G	KF86	2x	56.083467	13.906670	3	SC	SWE	Vanneberga			
G	KF90	2x	39.998279	43.560745	2	SC	TUR	Göktaş Köyü			
G	KF91	2x	41.396940	41.423950	3	SC	TUR	Hopa			
G	KF96	2x	40.501269	37.776489	2	SC	TUR	Mesudiye			
G	KF97	2x	41.012030	38.849030	3	SC	TUR	Tirebolu			
G	KF98	2x	40.874194	37.762756	3	SC	TUR	Ulubey			
G	KF99	2x	41.045958	39.280186	1	SC	TUR	Vakfikebir			
G	KF115	2x	52.106505	15.611572	3	SC	POL	Sulechów			
G	KF117	2x	52.007295	5.500183	3	SC	NLD	Veenendaal			
G	KF118	2x	51.652495	5.122242	3	SC	NLD	Tilburg			

Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation

Bohumil Mandák[1, 2], Alena Havrdová[1, 2], Karol Krak[1, 2], Věroslava Hadincová[1], Petr Vít[1, 2], Petr Zákravský[1] and Jan Douda[1, 2]

[1] Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, CZ-252 43, Průhonice, Czech Republic; [2] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha 6 – Suchdol, CZ-165 21, Czech Republic; bohumil.mandak@ibot.cas.cz

Summary

- We reconstructed the historical pattern of postglacial biogeographic range expansion of the boreal tree species *Alnus incana* in Europe.
- To assess population genetic structure and diversity, we performed a combined analysis of nuclear microsatellite loci and chloroplast DNA sequences (65 populations, 1004 individuals).
- Analysis of haplotype and microsatellite diversity revealed that (i) southeastern refugial populations did not spread north and cannot be considered as important source populations for postglacial recolonization of Europe; (ii) eastern populations did not establish Fennoscandian populations; (iii) Northern Europe, i.e. Fennoscandia and Eastern Europe, has no unique genetic cluster, but represents a mix with a predominant cluster typical for Central Europe; and that (iv) colonization of Fennoscandia and Eastern Europe took place from Central Europe and founding Central-European populations most likely in the Alps.
- Our findings highlight the importance of an effective refugium in Central Europe located outside classical southern refugia confirming the existence of northern refugia for boreal trees in Europe. The postglacial range expansion of *A. incana* did not follow the model established for *Picea abies.* Fennoscandian populations are not derived from Eastern-European ones, but from Central-European ones that originated from populations in the Alps.

Key words: *Alnus*, approximate Bayesian computation, cpDNA, microsatellite, phylogeography, population structure, postglacial migration.

Introduction

The distribution of most taxa is markedly influenced by historical migrations (Bennett *et al.*, 1991; Taberlet *et al.*, 1998; Cruzan & Templeton, 2000; Gavin *et al.*, 2014), especially by direct influence of glaciations during the last 2.5 million years and colonization after the last Ice Age (i.e. about 12,000 BP) (Taberlet, 1998). While a majority of temperate tree species survived the Late Pleniglacial (LPG), 24 to 15 kyr BP (Tzedakis *et al.*, 2013), presumably in southerly located refugia (i.e. the Iberian, Italian and Balkan peninsulas) according to the "southern refugia paradigm" (e.g. Hewitt, 2000; but see Magri *et al.*, 2006; de Lafontaine *et al.*, 2013, 2014a, b), boreal tree species were probably distributed more widely (Tzedakis *et al.*, 2013). This is supported by the presence of genetically very rich populations found in the area north of the Alps (e.g. *Salix caprea* and *Betula pendula* – Palmé *et al.*, 2003, 2004; Maliouchenko *et al.*, 2007), which point to the existence of numerous populations fragmented into small refugia due to climatically unfavourable conditions of the open taiga and hemiboreal forests dominated by *Larix*, *Pinus*, *Picea* and *Betula* which likely occurred in the northern Carpathians, Belarus and the northwestern Russian Plains (Jankovská & Pokorný, 2008; Kuneš *et al.*, 2008; Binney *et al.*, 2009).

The most comprehensive studies focused on postglacial migration of boreal tree species have been done on *Picea abies* (i.e. Giesecke & Bennett, 2004; Latałowa & Van der Knaap, 2006; Tollesfrud *et al.*, 2008, 2009). Tollesfrud *et al.* (2008) showed, based on a combined analysis of variation in the mitochondrial *nad*1 gene containing two minisatellite regions and fossil pollen data, that during the last glaciation *Picea abies* survived in at least seven refugial areas from which it expanded during the Holocene. Because the distribution range of *Picea abies* is divided into a northern and a southern part, an important refugium for the northern part of Europe existed in the Russian Plains. The southern part of the species' range was colonized mainly from refugia located in the southeastern Alps, the southern Bohemian Massif and the Western Carpathians, but not from the Balkan Peninsula. Moreover, Parducci *et al.* (2012) documented the presence of ice-free refugia during most of the last Ice Age in Scandinavia; however, this conclusion has received considerable criticism (Birks *et al.*, 2012). Parducci *et al.* (2012) pointed out the possible occurrence of a specific DNA haplotype in northwestern Scandinavia, found also in sediments dating back 10,300

yr. In their view, populations of *Picea abies* located in northwestern Norway might have survived the LPG in microenvironmentally favourable pockets and colonized mainly the western part of Scandinavia after climate warming.

Palaeoecology provides extremely important information concerning the postglacial recolonization of Europe (e.g. Huntley & Birks, 1983), but some congener species are indistinguishable from each other in palaeoecological studies based on pollen analysis, so only macrofossil remains or detailed molecular studies can be used when reconstructing their postglacial colonization routes. Typical examples of this are species of the subgenus *Alnus*, i.e. *A. glutinosa* (a temperate tree species) and *A. incana* (a boreal tree species). In their classic study, Huntley & Birks (1983) assumed that the main source refugia for the *Alnus* expansion after the LPG lay in the eastern Alps, the Carpathians and the Ukrainian lowlands, and that other LPG refugia were located in Corsica, western France, northern Spain and northwestern Russia. Huntley & Birks (1983) supposed that the Holocene migration of *Alnus* likely began somewhere in Eastern Europe and continued by the northward expansion of *A. glutinosa* and *A. incana* to the Baltic region and Fennoscandia, and by the westward expansion of *A. glutinosa* along the southern shore of the North Sea as far as the British Isles (Huntley & Birks, 1983). Furthermore, Douda *et al.* (2014) showed, based on radiocarbon-dated pollen and macrofossil sites, that *Alnus* species were likely to have withstood the LPG in Western Europe, the northern foothills of the Alps, the Romanian Carpathians and a large area of Northeastern Europe. After the withdrawal of glaciers, *Alnus* rapidly colonized southern Sweden and gradually expanded northward, most likely predominantly from a refugium located in today's Belarus and western Russia. However, the picture drawn on the basis of fossil pollen records is probably a mixture of postglacial recolonization of two species with different ecological requirements, i.e. temperate *A. glutinosa* and boreal *A. incana*. Although a phylogeographical study of *A. glutinosa* has been published some time ago (King & Ferris, 1998), until now no follow-up study focusing on the postglacial range expansion of the palaeoecologically indistinguishable congener species *A. incana* has since been carried out.

The distribution range of *A. incana* is divided similarly to that of *Picea abies* (Tollefsrud *et al.*, 2008) into a northern and a southern part (Fig. 1). Here we assess variation in maternally inherited cpDNA and biparentally inherited microsatellites across the entire distribution range of *A. incana* in light of a previously published

review on the distribution of subgenus *Alnus* based on radiocarbon-dated pollen and macrofossil sites (Douda *et al.*, 2014). Taking into account the similarity of distribution ranges, we hypothesized that the postglacial migration pattern of *A. incana* should follow the pattern established for *Picea abies*. The following questions were asked: (1) How are *A. incana* populations genetically structured across the species' distribution range? (2) Is the pattern of postglacial range expansion congruent with that established for *Picea abies*? (3) Is there any evidence of cryptic refugia occurring alongside southern refugial areas? (4) Did admixture occur between recolonizing populations during the expansion of the postglacial range of this species?

Materials and Methods

Study species

Alnus incana (L.) Moench subsp. *incana* (grey alder) (family *Betulaceae*) is a diploid (2n = 2x = 28), wind-pollinated, self-incompatible, relatively short-lived woody species (ca. 20–50 years) of riparian and water-logged habitats (Tallantire, 1974; Douda *et al.*, 2009, 2010). It can reproduce after 6 to 15 years of age, the maximum lifespan being 60 to 100 years (Tallantire, 1974). In the north, *A. incana* covers the entire cold Fennoscandia and extends eastwards across European Russia to northern Asia (Fig. 1). This contrasts with its patchy occurrence in the southern part of the range, which is confined to the Alps, Hercynian Mountains, Carpathians and Dinaric Alps (Fig. 1).

Study area and sampling procedure

The research area covers most of the European distribution range of *A. incana*, although populations from the easternmost part of the distribution range were not included and the area of the Caucasus Mountains was represented by only a few samples (Fig. 1). Samples of *A. incana* were collected throughout Europe in summer

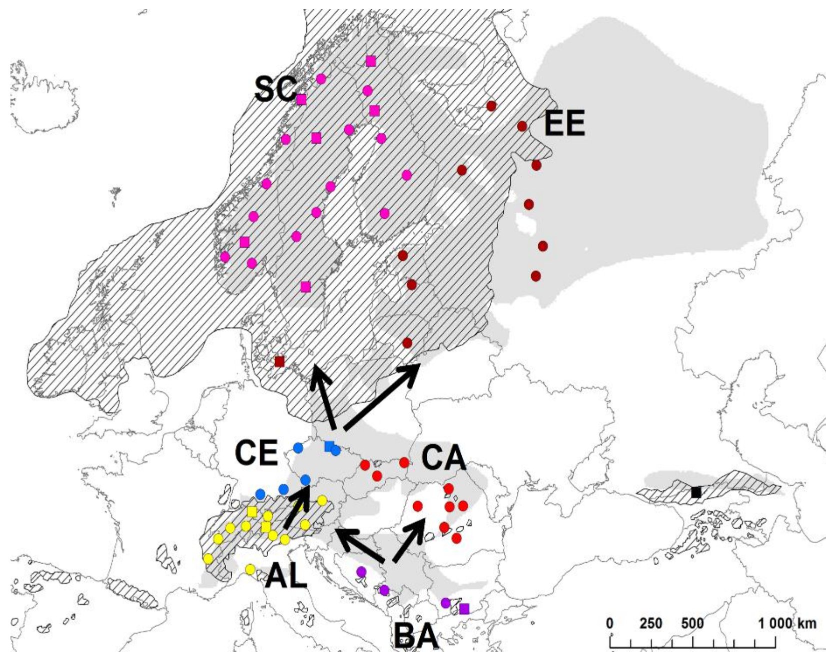


Fig. 1 Map of 65 localities of *Alnus incana*. The current distribution of *A. incana* is shadow, and the extent of glacier in LPG is hatched. Marks indicate populations sampled in this study – circles: detailed sampling, squares: coarse sampling (for more information concerning the sampling strategy, see Materials and Methods). European regions distinguished in this study are indicated by different colours – purple: Balkan region (BA), red: the Carpathians (CA), yellow: Alps (AL), blue: Central Europe (CE), pink: Fennoscandia (SC), brown: Baltic and Northeastern Europe (EE) and black: Caucasus. Arrows indicate the most probable colonization scenario for European *A. incana* populations determined by ABC analysis.

months from 2011 to 2013, exclusively from natural unmanaged forest stands. We applied several sampling strategies to include as many samples as possible in our analyses. (1) Detailed sampling – a stratified random sampling procedure was used to sample populations and individuals within populations for microsatellite and cpDNA analyses. Populations were at least 100 km apart. We collected 20 individuals per population if possible. Within each population, individuals were collected along linear transects at least 50 m apart, i.e. each population sample represents a one-kilometre-long or longer transect (Table 1, Fig. 1). All samples collected per population were analysed for variation in microsatellite loci. For cpDNA variation analyses, by contrast, usually the first, tenth and twentieth samples were included to ensure analysis of as distant samples as possible. All samples were stored in silica gel, and if possible fresh leaves were quickly transported to the laboratory and immediately analysed. (2) Coarse sampling – only three individuals per population were collected in some part of the range by our collaborators and were later used only for cpDNA analyses. Individuals always grew at least 500 m apart (Table 1, Fig. 1). In addition, three samples of *Alnus viridis* from the locality Zlatá Koruna, Czech Republic

Table 1 Summary of genetic diversity within 51 populations of *Alnus incana* based on eighteen microsatellite loci. The table is divided into two parts: Detailed sampling, i.e. samples analysed for microsatellite and cpDNA variation, and Coarse sampling, i.e. analysis of cpDNA variation only. Pop. = Population – the last three letters of each population name indicate the state in which population was found, i.e. AUT – Austria, BGR – Bulgaria, BIH – Bosnia and Herzegovina, CHE – Switzerland, CZE – Czech Republic, DEU – Deutschland, FIN – Finland, FRA – France, ITA – Italy, MNE – Monte Negro, NOR – Norway, ROU – Romania, RUS – Russia, SVK – Slovakia, SWE – Sweden. Group = geographically defined areas according Douda *et al.* (2014) to which individual populations belong, i.e., Balkan region (BA), Carpathians (CA), Alps (AL), Hercynian Mountains, i.e. Central Europe (CE), Scandinavia (SC) and Baltic and Northeastern Europe (EE); GPS coordinates of sampling sites are in WGS84; N – number of individuals sampled from each population; A – average number of alleles per locus; H_S – mean gene diversity over all loci; R_S – mean allelic richness; H_O – observed heterozygosity; H_E – expected heterozygosity; $f(F_{IS})$ – inbreeding coefficient according to Weir & Cockerham (1984); Haplo – cpDNA haplotype present in sample. Populations deviating from Hardy-Weinberg equilibrium at $P < 0.05$ are marked by asterisk. SD – standard deviation.

Pop.	Group	Locality	Latitude	Longitude	N	A	H_S	R_S	H_O	H_E	$f(F_{IS})$	Haplo
Detailed sampling												
1ITA	AL	Busche	46.04957	11.99467	19	3.9	0.447	3.623	0.427	0.446	0.043	1
2ITA	AL	Denno	46.26748	11.06451	19	3.9	0.449	3.637	0.430	0.449	0.044	2, 8
3ITA	AL	Gorreto	44.60650	9.29816	20	4.4	0.473	4.013	0.461	0.473	0.025	1
4AUT	AL	Lunz am See	47.85642	15.12061	20	3.5	0.453	3.253	0.472	0.453	–0.043	1
5AUT	AL	Hallein	47.65932	13.15365	20	4.1	0.454	3.609	0.422	0.453	0.069	1, 7
6AUT	AL	Paternion	46.72222	13.63543	20	4.9	0.510	4.425	0.428	0.508	0.161*	1, 13
7DEU	CE	Bruckberg	48.50245	11.99185	20	4.2	0.462	3.766	0.444	0.462	0.039	1
8AUT	AL	Mils bei Imst	47.20395	10.67570	19	3.8	0.469	3.536	0.456	0.469	0.027	1
9FRA	AL	Le Bourg-d'Oisans	45.04190	6.05029	20	3.1	0.390	2.917	0.383	0.390	0.018	1
10CHE	AL	Sixt-Fer-à-Cheval	46.04728	6.76398	20	3.3	0.415	3.054	0.400	0.415	0.036	1, 19
11CHE	AL	Frutigen	46.61223	7.68083	20	3.4	0.392	3.102	0.404	0.392	–0.031	1, 4
12CHE	AL	Sumvitg	46.72211	8.91563	20	3.6	0.420	3.319	0.433	0.421	–0.031	1, 2
13DEU	CE	Vöhringen	48.27184	10.06316	13	4.0	0.465	3.578	0.447	0.464	0.038	1
14CZE	CE	Lenora	48.91268	13.82311	20	4.5	0.456	3.983	0.425	0.455	0.068*	1
15CZE	CE	Chomutov	50.50060	13.32981	20	3.4	0.454	3.389	0.466	0.454	–0.026	1
16CZE	CE	Orlické Záhoří	50.23715	16.53814	18	3.8	0.452	3.595	0.468	0.452	–0.036	1, 5
17ROU	CA	Negreni	46.95026	22.72396	20	5.4	0.544	4.797	0.557	0.544	–0.024	13

18ROU	CA	Praid	46.57658	25.22393	20	5.7	0.562	4.946	0.515	0.561	0.085*	13
19ROU	CA	Preluci	46.46125	26.26796	20	4.9	0.524	4.422	0.522	0.524	0.004	11, 14
20ROU	CA	Tatarani	44.99385	25.28478	20	5.4	0.583	4.806	0.580	0.583	0.006	13
21BGR	BA	Yakoruda	42.00675	23.64008	20	4.7	0.512	4.257	0.516	0.512	–0.008	17
22MNE	BA	Žabljak	43.14469	19.29748	20	4.8	0.516	4.323	0.504	0.516	0.023	18
23BIH	BA	Vitez	44.18426	17.74202	16	4.1	0.461	3.914	0.444	0.460	0.035	1
24ROU	CA	Valea Putnei	47.44844	25.43936	20	4.3	0.511	3.974	0.493	0.510	0.035	1
25SVK	CA	Osadné	49.16437	22.20100	20	4.6	0.545	4.167	0.525	0.544	0.036	1
35ROU	CA	Porumbacu de Sus	45.66133	24.51872	20	4.5	0.541	4.130	0.527	0.541	0.027	13
38NOR	SC	Notodden	59.56533	9.19113	20	3.9	0.448	3.594	0.450	0.448	–0.004	1
39NOR	SC	Markhus	59.82816	6.24663	20	4.1	0.461	3.707	0.453	0.461	0.018	1, 19
40NOR	SC	Sel	61.84673	9.33083	20	3.9	0.470	3.642	0.481	0.470	–0.023	1, 19
41NOR	SC	Muruvik	63.43664	10.84057	20	4.0	0.455	3.736	0.464	0.456	–0.019	1, 19
42NOR	SC	Grane	65.57752	13.40024	20	4.1	0.472	3.771	0.474	0.472	–0.005	1
44SWE	SC	Boden	65.68637	21.64997	20	4.0	0.440	3.617	0.453	0.440	–0.029	1, 10
45SWE	SC	Abisco	68.35760	18.74583	20	3.4	0.438	3.116	0.428	0.437	0.023	1, 6
46FIN	SC	Kaukonen	67.37421	24.89247	20	3.8	0.442	3.539	0.469	0.443	–0.060	1
47FIN	SC	Oulu	64.94730	25.47806	20	4.2	0.449	3.751	0.409	0.448	0.089*	1
48FIN	SC	Kuopio	62.89678	27.61179	20	4.8	0.504	4.320	0.501	0.504	0.005	1
49FIN	SC	Pälkäne	61.33422	24.26574	20	4.6	0.492	4.150	0.484	0.492	0.016	1, 12
50SWE	SC	Friggesund	61.90295	16.52971	20	4.0	0.478	3.667	0.486	0.478	–0.018	1
51SWE	SC	Gävunda	60.81671	14.11120	20	4.3	0.476	3.872	0.462	0.475	0.029	1
57SVK	CA	Pohronská Polhora	48.74888	19.81227	20	4.3	0.524	3.985	0.519	0.524	0.010	1
58SVK	CA	Klubina	49.35969	18.90837	20	4.8	0.499	4.243	0.503	0.499	–0.008	1
64RUS	EE	Kivach	62.27591	33.99370	20	4.8	0.502	4.283	0.478	0.501	0.048	1
66LTU	EE	Rumšiškės	54.87927	24.20110	20	4.6	0.487	4.167	0.503	0.487	–0.033	1, 19
67LVA	EE	Strenči	57.61913	25.70187	20	4.6	0.482	4.069	0.494	0.482	–0.026	1
68EST	EE	Ardu	59.12385	25.35398	20	4.4	0.490	3.958	0.469	0.490	0.044	1

69RUS	EE	Pushkino	56.02788	37.80823	20	4.6	0.502	4.124	0.516	0.502	−0.028	20
70RUS	EE	Rostov	57.23391	39.49042	20	4.8	0.497	4.382	0.480	0.497	0.034	1, 20
75RUS	EE	Vologda	59.40823	39.70186	20	4.7	0.502	4.196	0.488	0.502	0.028	1, 20
76RUS	EE	Velsk	60.97661	42.06729	20	4.6	0.493	4.125	0.478	0.492	0.030	1, 20
77RUS	EE	Bereznik	63.03129	42.31052	20	5.0	0.498	4.346	0.492	0.498	0.012	20
78RUS	EE	Severodvinsk	64.61863	39.81693	20	4.9	0.478	4.269	0.482	0.478	−0.009	9, 20
Mean						4.3	0.479	3.905	0.472	0.479	0.007	
SD						0.6	0.040	0.451	0.040	0.040	0.073	
Coarse sampling												
26BGR		Rudozem	41.51600	24.88277	3							16
32ITA		Laudes	46.67132	10.52043	3							1
37CHE		St. Gallen	47.41675	9.41216	3							1, 7
43SWE		Sorsele	65.51446	17.32806	3							1
52DNK		Nykøbing Falster	54.73927	11.92178	2							1
54NOR		Hol	60.59220	8.34363	3							19
55FIN		Rovaniemi	66.34268	25.34149	3							1
56SWE		Ödeshög	58.33016	14.81612	3							1
59FIN		Njurkulahti	68.75009	26.17882	3							1
60NOR		Lakshol	67.45785	15.78301	3							1
62SWE		Docksta	63.08016	18.47626	3							1
65GEO		Ghebi	42.78626	43.43936	5							15
71CZE		Deštné	50.32504	16.32902	3							1, 3

(48°51'5.358"N, 14°20'16.309"E) were used as an outgroup in analyses of cpDNA variation.

For cpDNA variation analyses, a total of 65 populations and 171 individuals were collected. For microsatellite analyses, 51 populations and 1004 individuals were used (Table 1, Fig. 1).

Analysis of cpDNA

Total genomic DNA was extracted from silica-dried leaves as described in Štorchová *et al.* (2000). The *ndh*F-*rpl*32, *psb*J-*pet*A and *3'rps*16-*5'trn*K intergenic spacers were amplified using the primers published by Shaw *et al.* (2007). The PCRs were performed in 25 µl containing 1x Plain PP Mastermix (TopBio, Prague, Czech Republic) 0.2 mM of each primer and a few nanogrammes of genomic DNA. The cycling conditions were: 95°C for 3 min, followed by 40 cycles of 95°C for 30s, 49°C (*ndh*F-*rpl*32, *psb*J-*pet*A) or 52°C (*3'rps*16-*5'trn*K) for 30s and 72°C for 2 min. The reactions were completed by a final elongation step at 72°C for 15 min. The PCR products were checked on 1% agarose gel and sent to Macrogen (Amsterdam, Netherlands) for sequencing.

Microsatellite analysis

DNA was isolated using the DNeasy 96 Plant Kit (Qiagen, Germany). The quality and yield of isolated DNA was checked on 1% agarose gels and then by precisely measuring DNA concentration and purity using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). All samples were then diluted to 20–25 ng/µl prior to PCR (for more details, see also Drašnarová *et al.*, 2014).

We analysed genetic variation at 18 nuclear microsatellite loci in 1004 samples of *A. incana* from 51 populations (Table 1). These loci have been cross-amplified from closely related species by Drašnarová *et al.* (2014) (multiplex PCR 1: A2, A6, A10, A18, A22, A26, A35, A37, A38) or developed specifically for both *Alnus glutinosa* and *A. incana* by Lepais & Bacles (2011) (multiplex PCR 2: Ag1, Ag5, Ag9, Ag10, Ag13, Ag14, Ag20, Ag27, Ag30).

DNA was amplified using the QIAGEN Multiplex PCR kit (QIAGEN, Germany) in a total reaction volume of 5 µl of PCR mix plus 5 µl of mineral oil to avoid PCR mix evaporation, containing 20–25 ng of DNA, 0.1–0.5 µM of each primer and 2.5 µl of Master Mix (QIAGEN). PCR amplifications were conducted in a Mastercycler (Eppendorf, Germany) under the following conditions for multiplex PCR1: 15 min of denaturation at 95 °C, followed by 40 cycles at 94 °C for 30 s, 30 s at 58 °C, 60 s at 72 °C and a final extension of 10 min at 72 °C, and for multiplex PCR2: 5 min denaturation at 95°C, 30 cycles at 95°C for 30 s, 58°C for 3 min, 72 °C for 30 s and extension of 30 min at 60 °C. PCR products were electrophoresed in an ABI PRISM 3130 sequencer (Applied Biosystems, Foster City, CA, USA). One microlitre of PCR product was mixed with 0.2 µl of GeneScan-500 LIZ (Applied Biosystems) and 12 µl of Hi–Di formamide (Applied Biosystems). Allele sizes were determined using GeneMarker version 2.4.0 (SoftGenetics, USA). An individual was scored as null (nonamplifying at a locus) and treated as missing data after two or more amplification failures.

Data analysis

Analysis of cpDNA variation

All sequences were proofread using Chromas Lite 2.01 (Technelysium Pty. Ltd., Australia) and aligned manually in BioEdit 7.4.0.1 (Hall, 1999). Sequences of the three cpDNA regions were combined into a single dataset using FaBox (Villesen, 2007). Mononucleotide repeats were excluded due to a potentially high level of homoplasy (Ingvarsson *et al.*, 2003) prior to further processing of the data. Indels were coded by the simple gap coding method (Simmons & Ochoterena, 2000) as implemented in SeqState 1.4.1 (Müller, 2005). A chloroplast DNA haplotype network was constructed using a two-step procedure. First, a phylogenetic tree was constructed using MrBayes v 3.2.2 (Ronquist & Huelsenbeck, 2003). The data were divided into two parts (nucleotide sequence and the presence/absence matrix for coded indels) prior to the analysis. The GTR+I model was selected according to the AIC criterion using MrModeltest 2.3 (Nylander, 2004) for the sequence data as the appropriate model of nucleotide substitution. Default model settings were left for the

indel data. Two replicate analyses with four chains each were computed for 5 million generations, sampling every 1 000th generation. All statistical parameters indicated that convergence was reached. The first 1500 trees per run (30%) were discarded as burn-in, and the remaining 7002 trees were used to construct the consensus tree. Second, a cpDNA haplotype network was constructed based on this tree and the sequence alignment in HapView (Salzburger *et al.*, 2011). The original characters (0/1) in the presence/absence matrix representing indels were manually replaced by A and T, respectively, and the matrix was attached to the sequence alignment.

In order to identify cpDNA haplotype lineages, additional phylogenetic analyses were performed. Each haplotype was represented by one sequence, and sequences of *Alnus viridis* were included as an outgroup. Indel coding and Bayesian analyses were performed as described above, with the only difference that they were run for 2 million generations. A maximum parsimony analysis using PAUP* 4.0.b10 (Swofford, 2002) was performed using heuristic search with 100 replicates of random sequence addition and TBR branch swapping. Bootstrap analysis with 1000 replicates was performed to evaluate the support of the resulting clades.

Haplotypes belonging to Lineage 1 (see Fig. 2) possess a clear star-like phylogeny with one predominant haplotype (1) and numerous rare, closely related derivatives. This should be an indication of recent population expansion. We used the mismatch distribution approach using Arlequin (Schneider & Excoffier, 1999; Excoffier *et al.*, 2005) to detect population expansion events based on the frequency and nucleotide divergence of cpDNA haplotypes. The mode of mismatch distribution (τ) was used to date the expansion events following the equation of Rogers (1995): $t = \tau/2\mu$, where t is time and μ is the mutation rate for the whole sequence. Because the mutation rate has not been estimated for cpDNA in *Alnus*, we used the substitution rate estimated for synonymous substitutions in cpDNA of seed plants (i.e. 1.01×10^{-9} per site per year; Graur & Li, 1999).

Microsatellite genetic diversity

Summary data for SSR loci, including the average number of alleles per locus (A), mean gene diversity over all loci (H_S), mean allelic richness (R_S) [here allelic richness is a metric that uses a rarefaction index to take into account differences in

sample size (Goudet, 1995; El Mousadik & Petit, 1996)] and Weir & Cockerham's parameter $f(F_{IS})$ (1984), a measure of deviation from random mating within a population, were calculated using FSTAT (Goudet, 1995). Observed (H_O) and expected (H_E) heterozygosities were calculated using Arlequin (Excoffier *et al.*, 2005), and deviation from the Hardy-Weinberg equilibrium was determined based on 10,000 permutations in FSTAT.

Comparisons of genetic diversity parameters among regions defined following Douda *et al.* (2014) (i.e. the Balkan region (BA), the Carpathians (CA), the Alps (AL), Central Europe (CE), Fennoscandia (SC) and Baltic and Northeastern Europe (EE), see Table 1 for population distribution in individual groups and Fig. 1) were performed in FSTAT with 10,000 permutations.

Population structure based on microsatellite data

We used STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000) to estimate the number of genetic clusters (K) and to fractionally assign individuals sampled across the *A. incana* distribution range to the inferred groups. We applied the model which allows population admixture and correlated allele frequency (Pritchard *et al.*, 2000). Ten replicates for each K = 1–30 (the user-defined number of clusters) were set up to confirm the repeatability of the results. Each run comprised a burn-in period of 25,000 iterations followed by 100,000 Markov chain Monte Carlo (MCMC) steps. The STRUCTURE output data were parsed using the Structure-sum script in R (Ehrich *et al.*, 2007), mainly to determine the optimal K value following the method of Nordborg *et al.* (2005) and Evanno *et al.* (2005). Alignment of cluster assignments across replicate analyses was then conducted in CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) and subsequently visualized using DISTRUCT 1.1 (Rosenberg, 2004).

Testing alternative scenarios by ABC modelling

To obtain a more detailed inference of the postglacial history and migration patterns of European populations of *A. incana*, we compared several scenarios using the approximate Bayesian computation procedure (ABC; Beaumont *et al.*, 2002)

implemented in DIYABC 0.8.1 (Cornuet *et al.*, 2008). Based on Bayesian assignment of populations in STRUCTURE and different populations' Holocene history inferred from palaeoecological records (Douda *et al.*, 2014), we delimited populations in six large regions of *A. incana* (Table 1, Fig. 1), whose different origin was tested. Balkan populations represented an area that is considered to encompass southern LPG European refugia of trees (BA, 56 individuals). Baltic and Northeastern Europe (EE, 200 individuals) and Fennoscandia (SC, 260 individuals) were analysed separately because Fennoscandia was largely covered by the Scandinavian ice sheet during the LPG whereas northeastern European plains were not. Hence, Fennoscandian populations might be the result of postglacial expansion from northeastern European plains, a scenario resembling that published for *Picea abies* (Tollesfrud *et al.*, 2009). The Carpathians (CA, 180 individuals) and the Alps (AL, 217 individuals) represent areas of potential LPG refugia for many boreal trees. Central Europe to the north of the Alps (CE, 91 individuals) served as an ice-free corridor between the Alps and the Scandinavian ice sheet. The origins of BA, EE, SC, CA, CE, AL were investigated separately (e.g. Bryja *et al.*, 2010) by comparison of 4–6 scenarios for each regional population, i.e. 21 scenarios in total (Table 2, Fig. S1). For each scenario, priors and their minimum-maximum range were set up (Table 3). For simulations, we used the Generalized Stepwise Mutation model (GSM; Estoup *et al.*, 2002) and default values in DIY ABC for genetic parameters characterizing mutation rate (Cornuet *et al.*, 2008). The simulation of each scenario was performed using 1×10^6 iterations. Ten summary statistics were used to determine posterior probabilities of each scenario using direct estimation and logistic regression. These statistics included the mean number of alleles, mean allele size variance and the M index (Garza & Williamson, 2001; Excoffier *et al.*, 2005) for each population and population pair; and F_{ST} (Weir & Cockerham, 1984), classification index (Rannala & Moutain, 1997; Pascual *et al.*, 2007) and shared allele distance (Chakraborty & Jin, 1993) for population pairs. To evaluate confidence in scenario choice, we calculated false negative and false positive error rates from 500 pseudo-observed datasets for each scenario (Cornuet *et al.*, 2010). Demographic parameters of the most likely scenario were determined from 10,000 simulated datasets closest to the observed data (1% of the simulations).

Results

CpDNA haplotype diversity

Three non-coding cpDNA regions, *ndh*F-*rpl*32, *psb*J-*pet*A and *3'rps*16-*5'trn*K, were sequenced in 171 individuals from 65 populations of *A. incana*. The concatenated alignment of these three cpDNA regions consisted of 2782 bp. In this dataset, 30 variable sites were found, 23 of which were substitutions and 7 were indels.

Table 2 Posterior probability of 21 scenarios based on the logistic approach. Calculated confidence in scenarios described by both false positive and false negative error rates for logistic regression. Most likely scenarios are shaded out. Balkan region (BA), Carpathians (CA), Alps (AL), Central Europe (CE), Scandinavia (SC) and Baltic and Northeastern Europe (EE) (see Table 1 and Fig. 1).

Scenario	Origin	Logistic approach	False negative	False positive
1. Alps and Carpathians (BA)				
Scenario 1.1	AL, CA, BA refugium	0.0004 [0.0003, 0.0005]	0.096	0.089
Scenario 1.2	From BA (Holocene); BA refugium	0.9996 [0.9995, 0.9997]	0.072	0.083
Scenario 1.3	From BA (Pleistocene); BA refugium*	0.0000 [0.0000, 0.0000]	0.016	0.012
Scenario 1.4	From BA (Pleistocene); BA refugium**	0.0000 [0.0000, 0.0000]	0.092	0.118
Scenario 1.5	From BA (Pleistocene); BA	0.0000 [0.0000, 0.0000]	0.122	0.095
2. Central Europe (AL, CA)				
Scenario 2.1	AL, CA, CE refugium	0.0000 [0.0000, 0.1389]	0.135	0.104
Scenario 2.2	From CA; CA and AL refugium	0.0003 [0.0000, 0.1391]	0.116	0.100
Scenario 2.3	From AL; AL and CA refugium	0.1915 [0.0520, 0.3309]	0.104	0.109
Scenario 2.4	Admixture of AL and CA	0.0296 [0.0000, 0.1629]	0.156	0.141
Scenario 2.5	From AL; only AL refugium	0.7697 [0.7307, 0.8088]	0.138	0.155
Scenario 2.6	From CA; only CA refugium	0.0090 [0.0000, 0.2005]	0.151	0.188
3. Eastern Europe (CA, CE)				
Scenario 3.1	EE, CA, CE refugium	0.0000 [0.0000, 0.0000]	0.168	0.121
Scenario 3.2	From CE; CA and CE refugium	0.0000 [0.0000, 0.0000]	0.12	0.146
Scenario 3.3	From CA; CA and CE refugium	0.0000 [0.0000, 0.0000]	0.114	0.172
Scenario 3.4	Admixture of CE and CA	0.0000 [0.0000, 0.0000]	0.33	0.237
Scenario 3.5	From CA; only CA refugium	0.0182 [0.0154, 0.0209]	0.168	0.231
Scenario 3.6	From CE, only CE refugium	0.9818 [0.9791, 0.9846]	0.084	0.078
4. Fennoscandia (EE and CE)				
Scenario 4.1	From EE; EE and CE refugium	0.0010 [0.0007, 0.0012]	0.058	0.126
Scenario 4.2	From CE; EE and CE refugium	0.0007 [0.0005, 0.0009]	0.108	0.152
Scenario 4.3	Admixture of CE and EE	0.0010 [0.0007, 0.0012]	0.298	0.164
Scenario 4.4	From CE; only CE refugium	0.9974 [0.9969, 0.9979]	0.036	0.064

* Scenario 1.3, Colonization of AL and CA from BA during the Pleistocene (10,000–100,000 generations); **Scenario 1.4, Colonization of CA from BA refugium during the Pleistocene (10,000–100,000 generations), colonization of AL from CA during the Holocene; *** Scenario 1.5, Colonization of AL from BA refugium during the Pleistocene (10,000–100,000 generations), colonization of CA from AL during the Holocene

Twenty cpDNA haplotypes were identified (GenBank accession numbers KP244611–KP244673), 13 of which were unique to single accessions. Three

haplotype lineages (Fig. 2) and one ungrouped haplotype (18) were recognized based on the results of the phylogenetic analyses. Haplotypes and haplotype lineages showed distinct distribution ranges (Fig. 2), with haplotype 15 found exclusively in the Caucasus, and haplotypes 11, 13, 14 and 16–18 in Southeastern Europe. The most common haplotype 1 (found in 120 individuals), with the highest number of connections to other haplotypes usually separated by one mutation, was widely distributed across Southern, Central, Northern and partly also Eastern Europe, but was quite rare in Southeastern Europe. Haplotype 20 prevailed in Eastern Europe, but did not occur in Fennoscandia. Haplotype 19 was quite common in western Scandinavia (Norway), but was also found in Lithuania (Baltic area) and Switzerland (the Alps) (Fig. 2). Moreover, haplotype 13, common in the Carpathians, was also present in one population in the Alps, and haplotype 11, separated from Eastern-European haplotype 20 by one mutation, was present in the Carpathians (Fig. 2).

Table 3 Prior settings for all characteristics used to simulate 21 scenarios. Time ranges are expressed as the number of generations. Balkan region (BA), Carpathians (CA), Alps (AL), Central Europe (CE), Fennoscandia (SC) and Baltic and Northeastern Europe (EE) (see Table 1 and Fig. 1).

	Interpretation	Distribution	Minimum	Maximum
NBA, NEE, NSC, NCA, NAL, NCE	Population effective size	Uniform	1.00E+01	1.00E+05
t3	early Pleistocene divergence time of CE, CA, BA populations	Uniform	1.00E+04	1.00E+05
t3a	Time of colonization event of the Alps and the Carpathians in early Plesitocene	Uniform	1.00E+04	1.00E+05
t2	Split into refugial populations during LPG	Uniform	1.00E+03	1.00E+04
t1	Time of split or admixture of populations during Holocene expansion	Uniform	5.00E+02	1.00E+03
t1a	Time of colonization event of new regions during Holocene	Uniform	1.00E+01	1.00E+03
N_XF	Founding popultion size of regions	Uniform	1.00E+01	1.00E+04
r_x	Admixture rate of populations during Holocene expansion	Uniform		
μ	Mean mutation rate	Uniform	1.00E-04	1.00E-03
μi	Individual mutation rate	Gamma (shape parametr = 2)	1.00E-05	1.00E-02
μ_{SNI}	Mean single nucleotide insertion/deletion rate	Uniform	1.00E-08	1.00E-05
ind_{SNI}	Individual locus single nucleotide indel rate	Gamma (shape parametr = 2)	1.00E-09	1.00E-04
P_i	Individual locus probability that a new mutant allele differs from its ancestor	Gamma (shape parametr = 2)	1.00E-02	9.00E-01

The analysis of mismatch distribution showed that the observed distribution of pairwise differences among the haplotypes from Lineage 1 (comprising haplotype 1 and its derivatives, see Fig. 2) did not differ significantly from the distribution expected under the sudden expansion model (P = 0.56) (Fig. S2) The time of the expansion was estimated at 81,000 years ago (95% CI, 59 000–120 000) based on the mode of mismatch distribution $\tau = 0.457$ (95% CI, 0.330–0.676).

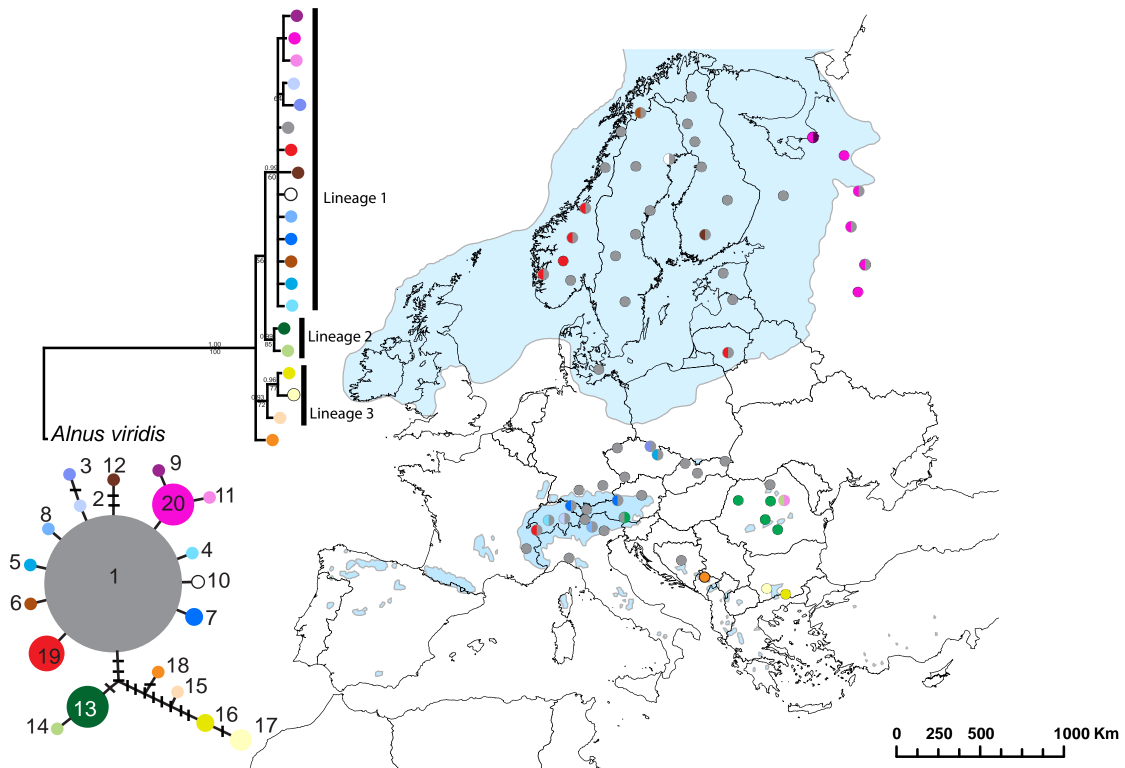


Fig. 2 Analysis of cpDNA (*ndh*F-*rpl*32, *psb*J-*pet*A, 3′rps16-5′trnK) haplotypes of *Alnus incana*. The 20 cpDNA haplotypes are represented by colours and named as in Table 1. (a) Geographic distribution of haplotypes across sampled populations. Pie charts represent haplotype presence, obtained after sequencing one to three individuals per population. Haplotype 15, unique to the Caucasus, is not indicated in the map. (b) Chloroplast DNA haplotype network (lines represent single nucleotide substitutions, and bars indicate missing haplotypes (extinct or not found). Circle sizes are proportional to the number of sequences obtained for each haplotype. (c) Fifty per cent majority-rule consensus tree of the Bayesian phylogenetic analysis; numbers above branches are Bayesian posterior probabilities (only values > 0.9 are indicated); numbers below branches are bootstrap supports (in percentage, only values > 50 are indicated) from the maximum parsimony analysis.

Population genetic diversity and structure

Genetic diversity

We identified 193 alleles at eighteen microsatellite loci, with an average of 10.7 alleles per locus. Summary statistics for genetic variability are shown in Table 1. For the 51 populations with at least 16 samples, the number of alleles (A) was 4.3 ± 0.6 (average ± standard deviation), gene diversity (H_S) was 0.479 ± 0.040, and allelic richness (R_S) was 3.905 ± 0.451 on average across markers. Heterozygote deficit was non-significant with a very low inbreeding coefficient [$f(F_{IS})$ = 0.007 ± 0.073]. Only four populations out of 51 were not in Hardy-Weinberg equilibrium (Table 1). These populations were distributed randomly across the range of *A. incana* (Table 1).

To understand the population genetic structure further, we divided the distribution range of *A. incana* into six geographical areas (see Table 1 for exact population distribution of individual groups and Fig. 1 for their geographical position). When we compared individual population genetic characteristics (see Table 4) among six predefined geographic areas, significant differences emerged among all of them, excluding inbreeding $f(F_{IS})$ and fixation $\theta(F_{ST})$ coefficients (Table 4). Populations in the Carpathians reached the highest values of individual population genetic characteristics, followed by populations growing in the Balkan Peninsula and eastern Europe (Table 4).

Table 4 Statistical comparison of allelic richness (R_S), observed heterozygosity (H_O), gene diversity (H_S), inbreeding coefficient $f(F_{IS})$ and levels of differentiation among populations $\theta(F_{ST})$ for *Alnus incans* populations in six geographically defined areas (see Table 1 and Fig. 1): Balkan region (BA), Carpathians (CA), Alps (AL), Central Europe (CE), Fennoscandia (SC) and Baltic and Northeastern Europe (EE). Probability values for differences among individual areas are for two-sided t-tests after 10,000 permutations. The analyses were performed using FSTAT software (Goudet, 1995).

Diversi	BA	CA	AL	CE	SC	EE	*P*
R_S	4.165	4.386	3.499	3.662	3.729	4.192	0.0034
H_O	0.492	0.527	0.429	0.448	0.463	0.488	0.0007
H_S	0.499	0.537	0.443	0.458	0.463	0.493	0.0010
$f(F_{IS})$	0.015	0.019	0.031	0.021	0.002	0.010	0.8144
$\theta(F_{ST})$	0.033	0.033	0.056	0.038	0.032	0.007	0.4859

Populations occurring in Fennoscandia and the Alps that did not exist throughout the LPG because of the ice sheet exhibited the lowest values of population genetic characteristics (Table 4). Populations growing in Central Europe exhibit low genetic variability compared to Fennoscandian and Alpine populations (Table 4).

The map of interpolated allelic richness (R_S) and gene diversity (H_S) shows that both parameters reach their highest values in the Carpathians and their lowest values in the westernmost Alps and the northern part of Fennoscandia (Fig. 3).

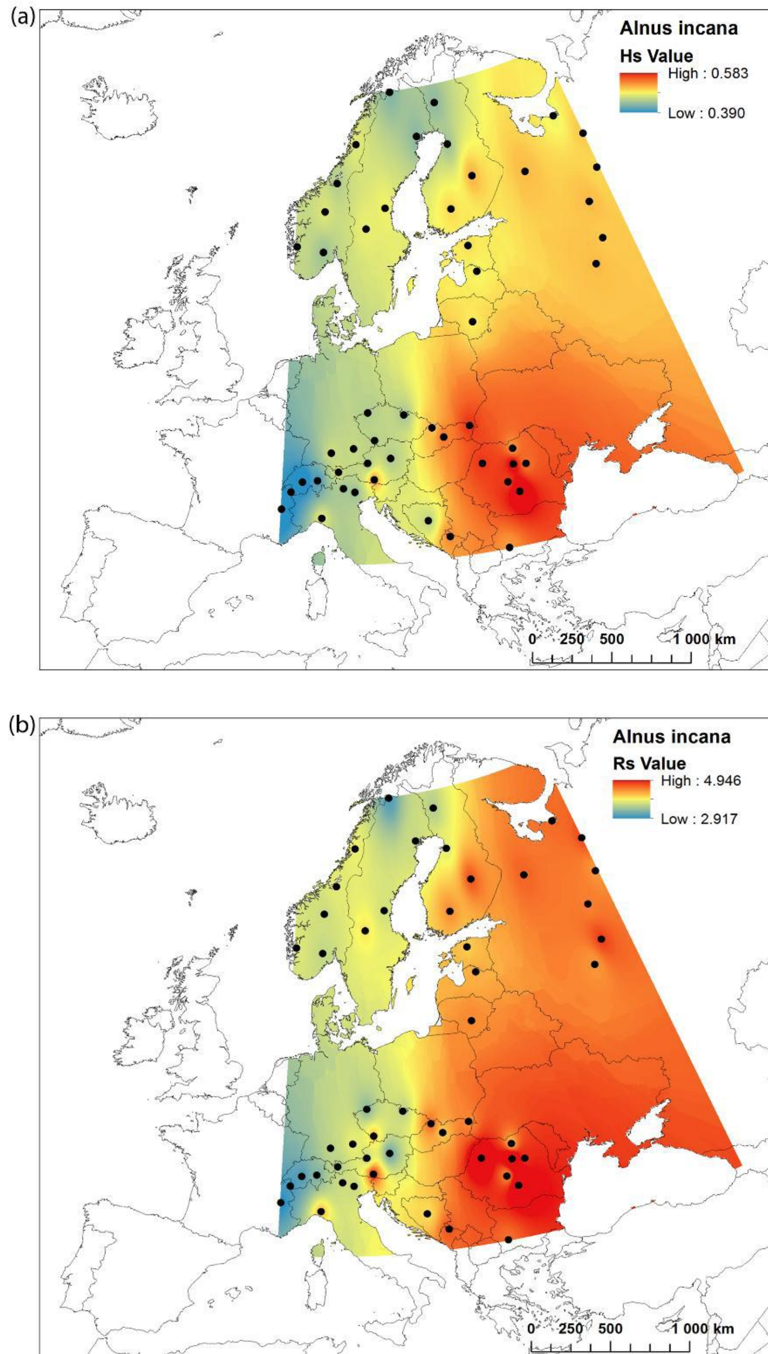


Fig. 3 Maps of overall (a) allelic richness (R_S) and (b) gene diversity (H_S).

Population structure

While the analysis of similarity coefficients indicated that two, three, four and six clusters best explained the genetic structuring of *A. incana* populations (Fig. S3), ΔK

indicated that only two, three and four clusters are informative (Fig. S4). Results of Bayesian clustering were further interpreted only for K = 2–4 (Fig. S5). Populations assigned to two clusters corresponding to a group of populations occurring generally in Western Europe (comprising Western, Southern and Central Europe) and a group of populations occurring in Southeastern Europe, i.e. the Carpathians and the Balkan Peninsula (Fig. S5).

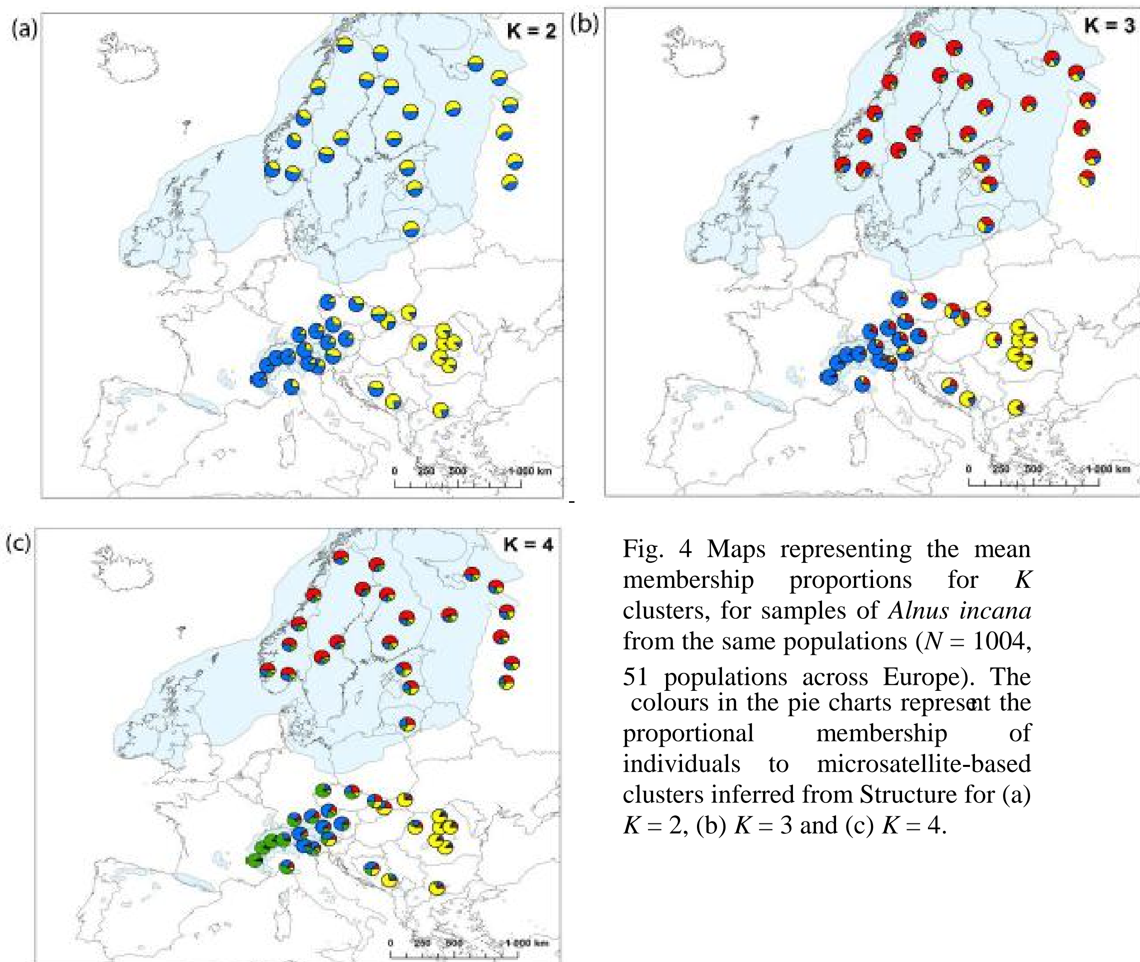


Fig. 4 Maps representing the mean membership proportions for K clusters, for samples of *Alnus incana* from the same populations (N = 1004, 51 populations across Europe). The colours in the pie charts represent the proportional membership of individuals to microsatellite-based clusters inferred from Structure for (a) K = 2, (b) K = 3 and (c) K = 4.

The rest of the populations sampled in Fennoscandia and Baltic and Northeastern Europe combine the two clusters to different extents (Fig. 4a). Fennoscandia was apparently not colonized from the east, but is a mixture of clusters occurring in the southern part of the *A. incana* distribution range (Fig. 4a).Assignment to three clusters provides us with the same picture, clearly separating Western Europe from the Carpathians and the Balkan Peninsula (Fig. 4b). A high proportion of genetic cluster depicted in red (Fig. 4b) is present in northern and eastern Europe and also within populations located in Central Europe. Assignment to four clusters provided more information concerning further division of Western Europe into the areas of the western and eastern Alps (Fig. 4c).

Models of population expansion and estimation of demographic parameters

The ABC approach did not support the presence of three refugial areas, i.e. the Balkan Peninsula, the Carpathians and the Alps (scenario 1.1), and gave priority to one refugium situated in the Balkan Peninsula. Hence, scenario 1.2 in which populations in the Alps and the Carpathians were derived from Balkan populations in Holocene received the highest support (Table 2, S1, 2). Central-European populations originated from the Alpine populations, i.e. scenario 2.5 (Table 2, S1, 2). Eastern Europe cannot be considered an effective glacial refugium based on the ABC approach, i.e. scenario 3.1 (Table 2, S1, 2). Instead, Central Europe gave rise to populations in Eastern Europe (scenario 3.6) (Table 2, S1, 2). It follows that Fennoscandia was not colonized from Eastern Europe as in the case of *Picea abies*, but from Central Europe, i.e. scenario 4.4 (Tables 2, S1, 2). The most probable colonization scenario for European *A. incana* populations is shown in Fig. 1.

Discussion

Colonization history of Alnus species based on pollen and macrofossil data

Our inability to precisely distinguish *A. glutinosa* and *A. incana* in palaeoecological records and the prevailing concept of eastern colonization of Europe by *A. glutinosa* (Huntley & Birks, 1983; King & Ferris, 1998) significantly complicate the interpretation of postglacial colonization of *A. incana*. Although fossil pollen data are scarce, mainly for the period of the LPG (there is only one locality in the Pyrenees; Fig. S6), there are relatively many localities dated from the Late Glacial and the early Holocene indicating the presence of *Alnus* species in all three classical southern refugia, i.e. the Iberian, Italian and Balkan peninsulas (Fig. S6) (Douda *et al.*, 2014). Macrofossils are also very scarce, and we have only few records supporting the presence of *Alnus* close to the Scandinavian ice sheet around the end of the last Ice Age which would indicate the presence of *Alnus* relatively high up north (Fig. S6). Macrofossil data, among others, suggest a possible colonization scenario of Fennoscandia from the area of the northeastern refugium located in the Russian plains. However, due to their scarcity, macrofossils do not provide us with information from

other areas, and although the existence of a refugium in eastern Europe cannot be ruled out, *Alnus* species could have survived the last Ice Age even in some other region or regions where their macrofossils are yet to be discovered.

Haplotype diversity and main glacial refugia

The more conservative cpDNA showed a clear pattern of distribution of haplotype diversity of *A. incana* in Europe, similar to that found in many other species (King & Ferris, 1998; Palmé & Vendramin, 2002; Petit *et al.*, 2002; Grivet & Petit, 2003; Hampe *et al.*, 2003; Heuertz *et al.*, 2004, 2006; Cheddadi *et al.*, 2006; Magri *et al.*, 2006; Höhn *et al.*, 2009; Liepelt *et al.*, 2009, Cornille *et al.*, 2013), i.e. haplotypes concentrated in Southeastern Europe are highly divergent and differ from haplotypes found in the rest of Europe. Significantly lower haplotype divergence was found in the Alpine region (Fig. 2), not forming a separate cluster like populations in Southeastern Europe. Populations in western Russia have their own haplotype (20) derived from and closely related to the most common one (1) (Fig. 2). Surprisingly, haplotypes present in western Russia did not colonize Fennoscandia, but the haplotype found in the Alps (19) was present in several populations currently growing in Norway and Lithuania. The emerging pattern based on cpDNA variation does not resemble the results of Lagercrantz & Ryman (1990) and Tollesfrud *et al.* (2009) for *Picea abies,* who documented an expansion of Russian populations to Fennoscandia after the glacial retreat. Even though *A. incana* is a boreal tree species, it probably took a different route when it colonized Europe in the postglacial period, and most of its populations are derived from populations surviving the LPG in the south. We can draw two important conclusions from the cpDNA variation presented here: (1) Fennoscandia was not colonized from eastern Europe like in the case of *Picea abies* (Tollesfrud *et al.*, 2009), and (2) putative glacial refugia situated in the Balkan Peninsula and partly in the Carpathians have nothing to do with northerly distributed populations.

Parducci *et al.* (2012), on the basis of analyses of macrofossils and cpDNA, showed that *Picea abies* might have survived the LPG in northwestern Scandinavia. They determined two haplotypes, A and B. While haplotype B was found outside of Scandinavia and appeared to have colonized this area from a refugium located in

Russia, haplotype A appeared to have spread from western Scandinavia. Hence, after the glacial retreat, colonization of Scandinavia allegedly took place from both directions, i.e. from local western stands that mixed with the colonization wave coming from the east. Even though we do not have such detailed data as Parducci *et al.* (2012), who analysed ancestral DNA from lake sediments dating back 10,300 years, we detected a very similar geographic distribution of one specific cpDNA halotype (i.e. 19) in western Scandinavia (Fig. 2), the only exception being that we have found this haplotype also in the Alps and in Lithuania. We would like to point out the possibility that this pattern determined for *A. incana* (i.e. colonization of western Scandinavia from populations that survived the LPG in Southern Europe) might have also been followed by *Picea abies*. The absence of *Picea abies* haplotype A in Southern Europe can easily be explained by its possible rarity and low probability of being captured even by a very dense sampling design. The possible rarity of haplotype A in large extant populations of *Picea abies* in the southern part of the species' distribution range might be yet another argument against the presence of refugia in western Scandinavia besides those already raised by Birks *et al.* (2012).

The haplotype distribution also suggests other interesting connections. The first one is the occurrence of the Carpathian haplotype 13 in the eastern Alps. This is also particularly visible in the microsatellite data, as population 6 contained individuals assigned to both the Carpathian and the Alpine cluster (see Fig. 4). We can only speculate about the possibility that a connection existed between populations in the Alps and the Carpathians across the Hungarian Lowland throughout the Ice Age. Of course, there is also the possibility that some individuals have been recently introduced from the Carpathians by man. The second connection is the occurrence of haplotype 11 derived from the western Russian haplotype 20 in one Carpathian population (Fig. 4). Again, a connection probably existed between Carpathian and eastern-European populations during the Ice Age across the Ukrainian and Belarusian plains. All abovementioned connections point to a relatively continuous distribution range of *A. incana* south and east to the ice sheet during the LPG.

Our data clearly show that haplotypes distributed in Southeastern Europe have never participated in the Holocene postglacial recolonization of the European continent. The reason behind the so commonly observed pattern might reside in the presence of various barriers preventing population expansion from the south, isolation

of these southern populations from northern ones, penetrability of the early Holocene landscape, floodplain development, individual species' dispersal ability, weak ecological adaptation of southern populations to conditions in the north or competition from other species.

Analysis of demographic history indicated significant demographic expansion in Northern and Central Europe about 81,000 years ago with a confidence interval of 59,000–120,000 which roughly falls within the last Weichselian glacial period (10,000–115,000 years ago) and predates the LPG. During that period, *Alnus incana*, as a typical boreal tree species, find suitable conditions for expansion and establishment in the territory of Central Europe that was not covered by the ice sheet. After the LPG, populations of the species were confined to the south. Later, at the beginning of the Holocene they colonized northern parts of Europe previously covered by the ice sheet. Hence, variation in cpDNA does not reflect Holocene migration. Rather, it is related to the last Ice Age preceding the LPG. This also implies lower resolution of cpDNA markers that generally inform us more about the Pleistocene than the Holocene.

The widely distributed haplotype 1 illustrates the situation clearly. We can speculate that haplotype 1 occurred or evolved in some population situated close to northern border of the distribution range throughout the last glacial period. In the early Holocene, populations on the edge of the distribution range possibly started to spread north, and haplotype 1 could be part of the expansion wave. It is important to note that the expansion to, for example, Fennoscandia was quite fast because it happened over a mere 3000 years, as has been documented by Douda *et al.* (2014, see Figs. 5F, G herein). Recent theoretical studies show that the genetic patterns resulting from range expansions can be very different from patterns generated by stationary distributions and closely dependent on the details of the recolonization process (Austerlitz *et al.*, 2000; Klopfstein *et al.*, 2006). If haplotype 1 was present in northern populations during the glacial time, then the present pattern of distribution should support the conclusions of Edmonds *et al.* (2004) and Klopfstein *et al.* (2006). They studied the fate of mutations in the course of range expansion and concluded that range expansions have had a strong impact on genetic diversity of individual species. Hence, mutations arising in populations at the edge of an expanding range can surf on the wave of advance, and thus reach a larger spatial distribution and a much higher frequency than

would be expected in stationary populations. However, these new mutations usually do not travel at all and remain rare or get lost by genetic drift. Successful "surfing mutations" can reach very high frequencies and eventually occupy large areas. This scenario, even though it can be seen as pure speculation, might explain the distribution of haplotype 1, which covers most of Central and Northern Europe to where it could have expanded from some other area, partly obscuring our ability to map the position of effective glacial refugia by studying variation in slowly mutating plastid markers such as cpDNA.

Postglacial migration routes established on variation in microsatellite markers

Highly variable SSRs markers are partly congruent with the pattern established on the basis of plastid markers. Based on the two genetic clusters defined by Structure, we show that only two clearly differentiated populations can be distinguished in Europe: (i) one in Southeastern Europe spanning the area of the Balkan Peninsula, the Carpathian Mountains and the Dinaric Alps and (ii) one in Western Europe covering the Alps and Central Europe. Northern Europe, including Scandinavia and adjacent western Russia, does not form a genetically separate entity. The same results were obtained for three and four genetic clusters defined by Structure, which stressed the importance of Central Europe for recolonization of Fennoscandia and Eastern Europe (Fig. 4).

As is evident from Fig. 4, and is also strongly supported by the ABC analysis (Table 2, Fig. 1), populations growing in Central Europe today belong to a genetic cluster that is very common in Northern and Eastern Europe, but very rare in Southern and Southeastern Europe. Based on the combined results of the Structure and ABC analyses, there is evidence that populations currently occurring in the Sudeten Mountains, the westernmost Carpathians and in the eastern Alps represent cryptic refugia once situated in lowland regions of Central Europe and possibly also in microenvironmentally favourable pockets in lower parts of mountain ranges during the LPG. The same pattern has been found by Tollesfrud *et al.* (2009) for *Picea abies*. In contrast to *A. incana*, Tollesfrud *et al.* (2009) did not document colonization of Northern Europe from the area of Central Europe. From this point of view, populations growing nowadays in mountains of Central Europe might surprisingly be some kind

of glacial relicts derived from populations that survived the LPG somewhere in the lowlands of Central Europe.

The highest genetic diversity has been found in Southeastern Europe and also in Eastern Europe (Fig 3), especially in the Carpathians, where the refugium of *A. incana* should be situated. However, this refugium should be regarded as non-effective, and the high genetic diversity of extant populations is probably due to continuously large populations and is not a sign of a refugium. This is also supported by the results of our cpDNA analysis, which clearly separated Southeastern Europe from the rest of the territory, probably indicating large populations that were significantly differentiated and not connected with northern populations.

Comparison with other boreal tree species is somewhat problematic because there are no studies, except that the one dealing with *Picea abies*, that cover the distribution range of the study species and, at the same time, have the same distribution pattern as *A. incana*. However, three tree species have been studied using different molecular markers, namely *Betula pendula* (Maliouchenko *et al.*, 2007), *Pinus sylvestris* (Naydenov *et al.*, 2007; Pyhäjärvi *et al.*, 2008; Vidyakin *et al.*, 2012) and *Salix caprea* (Palmé *et al.*, 2003, 2004). These studies support the presence of genetically rich populations in the area north of the Alps and confirmed our concept of an open taiga dominated by *Larix*, *Pinus*, *Picea* and *Betula* suggested by palaeoecological studies from Eastern Europe (Jankovská & Pokorný, 2008; Kuneš *et al.*, 2008; Binney *et al.*, 2009). Another important outcome of these studies is the lack of a phylogeographic structure, which might be a consequence of multiple factors: (i) presence of intermediate latitude refugia with large population sizes during the LPG, (ii) rapid recolonisation and dispersal ability, (iii) a high mutation rate, and (iv) extensive hybridisation with other species (Palmé *et al.*, 2003), especially *Salix* and *Betula*. In the case of *Pinus sylvestris*, Vidyakin *et al.* (2012) studied the variability of the first intron of the *nad7* gene and hypothesized that populations of *P. sylvestris* in northeastern Russia and Fennoscandia originated from different glacial refugia, i.e. that Fennoscandia was not colonized from the northeastern Russian glacial refugium, but from an isolated microrefugium in eastern Fennoscandia or adjacent areas of European Russia. This pattern resembles the one determined for *A. incana* in this study. However, the lack of resolution in the marker used, which has the same haplotype in Eastern Russia and Western Europe (Pyhäjärvi *et al.*, 2008) or in Finland

and in Novosibirsk (Vidyakin *et al.*, 2012), highlights the necessity to develop a different genetic marker which would clarify the position of glacial refugia of *P. sylvestris*.

Conclusions

Our study of the colonization history of *A. incana* in Europe reveals new perspectives on the history of boreal tree species and answers some of the questions concerning the relationships between the past geographic distribution of this species and the modern distribution of its genetic diversity. Both chloroplast and nuclear markers indicate that Eastern Europe was not an effective glacial refugium serving as a source for the colonization of Fennoscandia. Instead, Fennoscandian populations resemble populations currently growing in Central Europe (Fig. 1). Our results highlight the importance of cryptic refugia in Central Europe, which should be considered as the main source area from which *A. incana* colonized Northern Europe, confirming the existence of relatively northern refugia of boreal trees in Europe. By contrast, populations that survived the last glacial period in Southeastern Europe did not spread into Central and Northern Europe.

Acknowledgements

We wish to thank Eva Ničová and Martina Hadincová for their help in the lab and Fred Rooks for helping us with the language of the manuscript. Jana Doudová kindly helped obtain cpDNA sequences. We are also grateful to Adam Konečný for fruitful discussions about ABC modelling. We also thank everyone who collected samples of *Alnus incana* and *A. viridis* for us, especially Filip Kolář, Martin Lepší, Vít Dvořák and Milan Štech. This study was supported by grant P504/11/0402 from the Grant Agency of the Czech Republic and by the long-term research development project RVO 67985939.

References

Austerlitz F, Mariette S, Machon N, Gouyon P-H, Godelle B. 2000. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* **154:** 1309–1321.

Beaumont MA, Zhang WY, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics* **162:** 2025–2035.

Bennett KD, Tzedakis PC, Willis KJ. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* **18:** 103–115.

Binney HA, Willis KJ, Edwards ME, Bhagwat SA, Anderson PM, Andreev AA, Blaauw M, Damblon F, Haesaerts P, Kienast F, Kremenetski KV, Krivonogov SK, Lozhkin AV, MacDonald GM, Novenko EY, Oksanen P, Sapelko TV, Väliranta M, Vazhenina L. 2009. The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews* **28:** 2445–2464.

Birks HH, Giesecke T, Hewitt GM, Tzedakis PC, Bakke J, Birks HJB. 2012. Comment on "Glacial survival of boreal trees in northern Scandinavia". *Science* **338:** 742–742.

Bryja J, Smith C, Konečný A, Reichard M. 2010. Range-wide population genetic structure of the European bitterling (*Rhodeus amarus*) based on microsatellite and mitochondrial DNA analysis. *Molecular Ecology* **19:** 4708–4722.

Chakraborty R, Jin L. 1993. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, eds. *DNA fingerprinting: state of the science*. Birkhäuser Basel: Springer, pp. 153–175.

Cheddadi R, Vendramin GG, Litt T, François L, Kageyama M, Lorentz S, Laurent J-M, De Beaulieu J-L, Sadori L, Jost A, Lunt D. 2006. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography* **15:** 271–282.

Cornille A, Giraud T, Bellard C, Tellier A, Cam BL, Smulders MJM, Kleinschmit J, Roldan-Ruiz I, Gladieux P. 2013. Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Molecular Ecology* **22:** 2249–2263.

Cornuet JM, Ravigné V, Estoup A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1. 0). *BMC Bioinformatics* **11**: 401.

Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin J-M, Balding DJ, Guillemaud T, Estoup A. 2008. Inferring population history with DIYABC: a user-friendly approach to Approximate Bayesian Computation. *Bioinformatics* **24:** 2713–2719.

Cruzan MB, Templeton AR. 2000. Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution* **15:** 491–496.

de Lafontaine G, Amasifuen Guerra CA, Ducousso A, Sánchez-Goñi M-F, Petit RJ. 2014b. Beyond skepticism: uncovering cryptic refugia using multiple lines of evidence. *New Phytologist* **204:** 450–454.

de Lafontaine G, Ducousso A, Lefevre S, Magnanou E, Petit RJ. 2013. Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology* **22:** 4397–4412

de Lafontaine G, Guerra CAA, Ducousso A, Petit RJ. 2014a. Cryptic no more: soil macrofossils uncover Pleistocene forest microrefugia within a periglacial desert. *New Phytologist* **204:** 715–729.

Douda J, Čejková A, Douda K, Kochánková J. 2009. Development of alder carr after the abandonment of wet grasslands during the last 70 years. *Annals of Forest Science* **66:** 712.

Douda J, Doudová J, Drašnarová A, Kuneš P, Hadincová V, Krak K, Zákravský P, Mandák B. 2014. Migration patterns of subgenus *Alnus* in Europe since the Last Glacial Maximum. A systematic review. *PLoS ONE* **9:** e88709.

Douda J. 2010. The role of landscape configuration in plant composition of floodplain forests across different physiographic areas. *Journal of Vegetation Science* **21:** 1110–1124.

Drašnarová A, Krak K, Vít P, Doudová J, Douda J, Hadincová V, Zákravský P, Mandák B. 2014. Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genetics & Genomes* **10:** 865–873.

Edmonds CA, Lillie AS, Cavalli-Sforza LL. 2004. Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 975–979.

Ehrich D, Gaudeul M, Assefa A, Koch M, Mummenhoff K, Nemomissa S, Intrabiodiv Consortium, Brochmann C. 2007. Genetic consequences of Pleistocene range shifts: Contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* **16:** 2542–2559.

El Mousadik A, Petit RJ. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92:** 832–839.

Estoup A, Jarne P, Cornuet JM. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11:** 1591–1604.

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.

Excoffier L, Estoup A, Cornuet JM. 2005. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* **169**: 1727–1738.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Garza JC, Williamson E. 2001. Detection of reduction in population size using data from microsatellite DNA. Molecular Ecology **10:** 305–318.

Gavin DG, Fitzpatrick MC, Gugger PF et al. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* **204**: 37–54.

Giesecke T, Bennett KD. 2004. The Holocene spread of *Picea abies* (L.) Karst. in Fennoscandia and adjacent areas. *Journal of Biogeography* **31**: 1523–1548.

Goudet J. 1995. Fstat version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Graur D, Li WH. 1999. Fundamentals of molecular evolution. Sinauer & Associates: Sunderland, Massachusetts.

Grivet D, Petit R. 2003. Chloroplast DNA phylogeography of the hornbeam in Europe: Evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics* **4:** 47–56.

Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Hampe A, Arroyo J, Jordano P, Petit RJ. 2003. Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. – *Molecular Ecology* **12:** 3415–3426.

Heuertz M, Carnevale S, Fineschi S, Sebastiani F, Hausman JF, Paule L, Vendramin GG. 2006. Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. *Molecular Ecology* **15:** 2131–2140.

Heuertz M, Fineschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste V, Hardy OJ, Vekemans X, Vendramin GG. 2004. Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology* **13:** 3437–3452.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Höhn M, Gugerli F, Abran P, Bisztray G, Buonamici A, Cseke K, Hufnagel L, Quintela-Sabarís C, Sebaastiani F, Vendramin GG. 2009. Variation in the chloroplast DNA of Swiss stone pine (*Pinus cembra* L.) reflects contrasting post-glacial history of populations from the Carpathians and the Alps. *Journal of Biogeography* **36:** 1798–1806.

Huntley B, Birks HJB. 1983. An atlas of past and present pollen maps for Europe: 0–13000 years ago. Cambridge: Cambridge University Press.

Ingvarsson PK, Ribstein S, Taylor DR. 2003. Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. *Molecular Biology and Evolution* **20:** 1737–1740.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801–1806.

Jankovská V, Pokorný P. 2008. Forest vegetation of the last full-glacial period in the Western Carpathians (Slovakia and Czech Republic). *Preslia* **80:** 307–324.

King A, Ferris C. 1998. Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology* **7:** 1151–1161.

Klopfstein S, Currat M, Excoffier L. 2006. The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* **23:** 482–490.

Kuneš P, Pelánková B, Chytrý M, Jankovská V, Pokorný P, Petr L. 2008. Interpretation of the last-glacial vegetation of eastern-central Europe using modern analogues from southern Siberia. *Journal of Biogeography* **35:** 2223–2236.

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

Latałowa M, van der Knaap WO. 2006. Late Quaternary expansion of Norway spruce *Picea abies* (L.) Karst. in Europe according to pollen data. *Quaternary Science Reviews* **25:** 2780–2805.

Lepais O, Bacles CFE. 2011. De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. *Journal of Heredity* **102:** 627–632.

Liepelt S, Cheddadi R, de Beaulieu J-L, Fady B, Gömöry D, Hussendörfer E, Konnert M, Litt T, Longauer R, Terhürne-Berson R, Ziegenhagen B. 2009. Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) – A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology* **153:** 139–149.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latałowa M, Litt T, Paule L, Roure JM, Tantau I, van der Knaap WO, Petit RJ, de Beaulieu J-L. 2006. A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist* **171:** 199–221.

Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, 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.

Müller K. 2005. SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* **4:** 65–69.

Naydenov K, Senneville S, Beaulieu J, Tremblay F, Jean Bousquet J. 2007. Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology* **7:** 203.

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng HG, Bakker E, Calabrese P, Gladstone J, Goyal R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao KY, Kalbfleisch T, Schulz V, Kreitman M, Bergelson J. 2005. The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol* **3:** e196.

Nylander JAA. 2004. MrModeltest v2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.

Palmé AE, Semerikov V, Lascoux M. 2003. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity* **91:** 465–474.

Palmé AE, Su Q, Palsson S, Lascoux M. 2004. Extensive sharing of chloroplast haplotypes among European birches indicates hybridization among *Betula pendula*, *B. pubescens* and *B. nana*. *Molecular Ecology* **13:** 167–178.

Palmé AE, Vendramin GG. 2002. Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology* **11:** 1769–79.

Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y, Edwards ME, Andersen K, Rasmussen M, Boessenkool S, Coissac E, Brochmann C, Taberlet P, Houmark-Nielsen M, Larsen NK, Orlando L, Gilbert MTP, Kjær KH, Alsos IG, Willerslev E. 2012. Glacial survival of boreal trees in northern Scandinavia. *Science* **335:** 1083–1086.

Pascual M, Chapuis MP, Mestres F, Balanya J, Huey RB, Gilchrist GW, Serra L, Estoup A. 2007. Introduction history of Drosophila subobscura in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16:** 3069–308.

Petit J, Csaikl UM, Cottrell J, van Dam B, Deans JD, Dumolin-Lapegue S, Fineschi S, Finkeldey R, Gillies A, Glaz I, Lowe AJ, Goicoechea PG, Jensen JS, Ko AO, Munro RC, Olalde M, Madsen SF, Ma G. 2002. Chloroplast DNA

variation in European white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* **156:** 5–26.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Pyhäjärvi T, Salmela MJ, Savolainen O. 2008. Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genetics & Genomes* **4:** 247–254.

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9197–9201.

Rogers AR. 1995. Genetic evidence for a Pleistocene population explosion. *Evolution* **49:** 552–569.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137–138.

Salzburger W, Ewing GB, Haeseler A. 2011. The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology* **20:** 1952–1963.

Schneider S, Excoffier L. 1999. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152:** 1079–1089.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* **94:** 275–288

Simmons MP, Ochoterena O. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49:** 369–381.

Štorchová H, Hrdličková R, Chrtek Jr J, Tetera M, Fitze D, Fehrer J. 2000. An improved method for DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon* **49:** 79–84.

Swofford DL. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sunderland, Mass.: Sinauer.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cossons J-F. 1998. Comparative phylogeography and postaglacial colonization routes in Europe. *Molecular Ecology* **7:** 453–464.

Taberlet P. 1998. Biodiversity at the intraspecific level: The comparative phylogeographic approach. *Journal of Biotechnology* **64:** 91–100.

Tallantire PA. 1992. The alder [*Alnus glutinosa* (L.) Gaertn.] problem in the British Isles: a third approach to its palaeohistory. *New Phytologist* **122:** 717–731.

Tollefsrud MM, Kissling R, Gugerli F, Johnsen Ø, Skrøppa T, Cheddadi R, van der Knaap WO, 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.

Tollefsrud MM, Sønstebø JH, Brochmann C, Johnsen Ø, Skrøppa T, Vendramin GG. 2009. Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity* **102:** 549–562.

Tzedakis PC, Emerson BC, Hewitt GM. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology and Evolution* **28:** 696–704.

Vidyakin AI, Semerikov VL, Polezhaeva MA, Dymshakova OS. 2012. Spread of mitochondrial DNA Hhplotypes in population of Scots pine (*Pinus sylvestris* L.) in Northern European Russia. *Russian Journal of Genetics* **48:** 1267–1271.

Villesen P. 2007. FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes* **7:** 965–968.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Graphical representation of the 21 competing scenarios used in the ABC analyses.

Fig. S2 Distribution of pairwise differences among 14 haplotypes belonging to Lineage 1

Fig. S3 Estimation of the number of genetic clusters following the method of Nordborg *et al.* (2005).

Fig. S4 Estimation of the number of genetic clusters following the method of Evanno et al. (2005).

Fig. S5 Population structure of *Alnus incana* estimated in STRUCTURE based on 1004 individuals from 51 localities.

Fig. S6 Pollen percentages and macrofossils of *Alnus incana* at selected sites since the Last Glacial Maximum based on data from Douda et al. (2014).

Table S1 Model checking by deviations of summary statistics for the observed data form the posterior predictive distribution of the most likely scenario of population origin.

Table S2 Parameter estimates of the most-likely demographic scenario for the origin of populations of *Alnus incana* in different regions.

Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture

Alena Havrdová[1, 2,*], Jan Douda[1, 2], Karol Krak[1, 2], Petr Vít[1, 2], Věroslava Hadincová[1], Petr Zákravský[1] and Bohumil Mandák[1, 2, *]

[1] Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, CZ-252 43, Průhonice, Czech Republic

[2] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha 6 – Suchdol, CZ-165 21, Czech Republic

[*]Corresponding author: havrdovaa@fzp.czu.cz, mandak@ibot.cas.cz

Running title: *Alnus glutinosa* postglacial colonization

Abstract

Genetic admixture is supposed to be an important trigger of species expansions because it can create the potential for selection of genotypes suitable for new climatic conditions. Up until now, however, no continent-wide population genetic study has performed a detailed reconstruction of admixture events during natural species expansions. To fill this gap, we analysed the postglacial history of *Alnus glutinosa*, a keystone species of European swamp habitats, across its entire distribution range using two molecular markers, cpDNA and nuclear microsatellites. CpDNA revealed multiple southern refugia located in the Iberian, Apennine, Balkan and Anatolian Peninsulas, Corsica and North Africa. Analysis of microsatellites variation revealed three main directions of postglacial expansion: 1) from the northern part of the Iberian Peninsula to Western and Central Europe and subsequently to the British Isles, 2) from the Apennine Peninsula to the Alps, and 3) from the eastern part of the Balkan Peninsula to the Carpathians followed by expansion towards the Northern European plains. This challenges the classical paradigm that most European populations originated from refugial areas in the Carpathians. It has been shown that colonizing lineages have met several times and formed secondary contact zones with unexpectedly high population genetic diversity in Central Europe and Scandinavia. On the contrary, limited genetic

admixture in southern refugial areas of *A. glutinosa* renders rear-edge populations in the Mediterranean region more vulnerable to extinction due to climate change.

Keywords: Approximate Bayesian computation, black alder, climate change, ice ages, phylogeography, temperate tree

Introduction

Genetic admixture, i.e. the mixture of previously long-isolated lineages after they come into secondary contact, creating novel genetic variation, is supposed to be an important trigger of species invasions (e.g. Durka *et al.* 2005; Genton *et al.* 2005; Marrs *et al*. 2008; Chun *et al.* 2009; Mandák *et al.* 2013; Keller *et al.* 2014) as well as natural range expansions (Colosimo *et al.* 2005; Mullen & Hoekstra 2008; De Carvalho *et al.* 2010; Rius & Darling 2014). Refugia that existed during the Late Pleniglacial (LPG), an interval of the most extreme glacial conditions which lasted from 15000 to 24000 ka BP (Tzedakis *et al.* 2013), provided suitable conditions for the survival of many species. Postglacial expansions from such refugia then determined the present-day distributions of these species in Europe. It appears that species expansion after the warming in the early Holocene was accompanied by hybridization between different genetic lineages within secondary contact zones located outside the glacial refugia (Taberlet *et al.* 1998; Hewitt 1999; Hewitt 2000). These hybridization events could result in continent-wide admixture zones with high within-population genetic diversity (Petit *et al.* 2003; Rius & Darling 2014). Petit *et al.* (2003) showed, using chloroplast DNA variation of 22 widespread European trees, that despite the presence of highly divergent populations in Mediterranean refugial areas, most diverse populations occur at intermediate latitudes in Central Europe. More recent studies that analysed population genetic admixture directly, agree that increasing allelic richness of plants and animals in recently colonized areas is the result of amalgamation of different genetic lineages coming from separate glacial refugia (Sakaguchi *et al.* 2011; van Els *et al.* 2012; Potter *et al*. 2012; Lait & Burg 2013). Up until now, however, no continent-wide population genetic study has performed a detailed reconstruction of the origin of secondary contact zones using demographic simulations.

In Europe, survival of temperate trees during the LPG in the three southern peninsulas (i.e. Iberian, Apennine and Balkan) is largely supported by fossil and molecular data (e.g. Palmé & Vendramin 2002; Grivet & Petit 2003; Tzedakis *et al.* 2013). It has been shown, however, that most of the southern areas were not the origins for postglacial expansions (but see Petit *et al.* 2002; Heuertz *et al.* 2004a). The northern edges of the southern peninsulas and surrounding regions such as southern France, the eastern foothills of the Alps and the Carpathians, have been proposed to be important source areas (King & Ferris 1998; Palmé & Vendramin 2002; Grivet & Petit 2003; Magri *et al.* 2006; Cornille *et al.* 2013). Post-glacial range expansions typically create clines of decreasing genetic diversity with increasing distance from refugia (Stamford & Taylor 2004; Muller *et al.* 2008). Founder effects and bottlenecks associated with postglacial dispersal may also have caused reductions in genetic diversity in contemporary populations of organisms (Hewitt 1996). However, genetic admixture might have combined diversity established by repeated founder events, increasing the genetic diversity outside refugial areas and probably enhancing the success of colonizing populations (Kolbe *et al.* 2004; Keller & Taylor 2010; Rius & Darling 2014).

Moreover, recent palaeoecological and molecular evidence (Willis *et al.* 2000; Willis & Van Andel 2004; Magri *et al.* 2006; Birks & Willis 2008; Binney *et al.* 2009; De Lafontaine *et al.* 2013; De Lafontaine *et al.* 2014) points to possible survival of many temperate trees during the entire last glaciation in micro-environmentally favourable pockets at higher latitudes than had previously been expected (i.e. concept of cryptic refugial areas; Stewart & Lister 2001). Nevertheless, the reliability of such evidence has been repeatedly questioned (Carcaillet & Vernet 2001; Tzedakis *et al.* 2013; Huntley 2014), and the concept of cryptic refugia is still under active debate, awaiting to be supported or falsified on the basis of available data.

Even though different European tree species have rather specific postglacial migration histories, we can find several general patterns. First, while many tree species colonized most northern areas from Southeastern Europe, populations that survived the LPG period in the Apennine and Iberian Peninsulas remained trapped there (King & Ferris 1998; Grivet & Petit 2003; Hampe *et al.* 2003; Magri *et al.* 2006). During the expansion from single refugia, the loss of genetic diversity with

increasing distance from refugial areas is hypothesized to be a result of repeated founding of new populations (Hewitt 1999; Hewitt 2000). Second, for some tree species multiple source refugia were found to be equally important (Petit *et al.* 2002; Heuertz *et al.* 2004a; Liepelt *et al.* 2009; Cornille *et al.* 2013). In these cases, lineages starting their expansion in different refugial areas could meet in secondary contact zones and mix extensively. Moreover, Aguinagalde *et al.* (2005) found that the level of present-day cpDNA genetic structure is significantly influenced by the type of distribution (i.e. boreal-temperate or boreal). However, the relationship between life-history traits and genetic differentiation in maternally inherited markers is weaker, especially when phylogenetic effects are controlled for.

To shed light on the postglacial history of temperate tree species in Europe, we studied the genetic diversity, migration patterns and lineage admixture in populations of *Alnus glutinosa* (Fig. 1), a keystone species of riparian and water-logged habitats (McVean 1953; Douda *et al.* 2009; Douda 2010). It has recently been shown that *A. glutinosa* occurs in two ploidy levels, diploid and tetraploid (Lepais *et al.* 2013; Mandák *et al.*, in review). Tetraploid populations have been found to form two geographically very well delimited clusters located in the Iberian Peninsula and in the Dinaric Alps (Mandák *et al.*, in review; Fig. 1). The high evolutionary and conservation value of these marginal, rear-edge relict populations has been emphasized (Lepais *et al.* 2013). Generally, understanding the past history and postglacial migration pattern of *A. glutinosa* populations may help understand the resistance and resilience of wetland forest habitats to global climate change (Erwin 2008; Garssen *et al.* 2015). In this context, lineage admixture can play a fundamental role because it can increase the fitness of populations and trigger species expansions (Keller & Taylor 2010; Keller *et al.* 2014; Rius & Darling 2014).

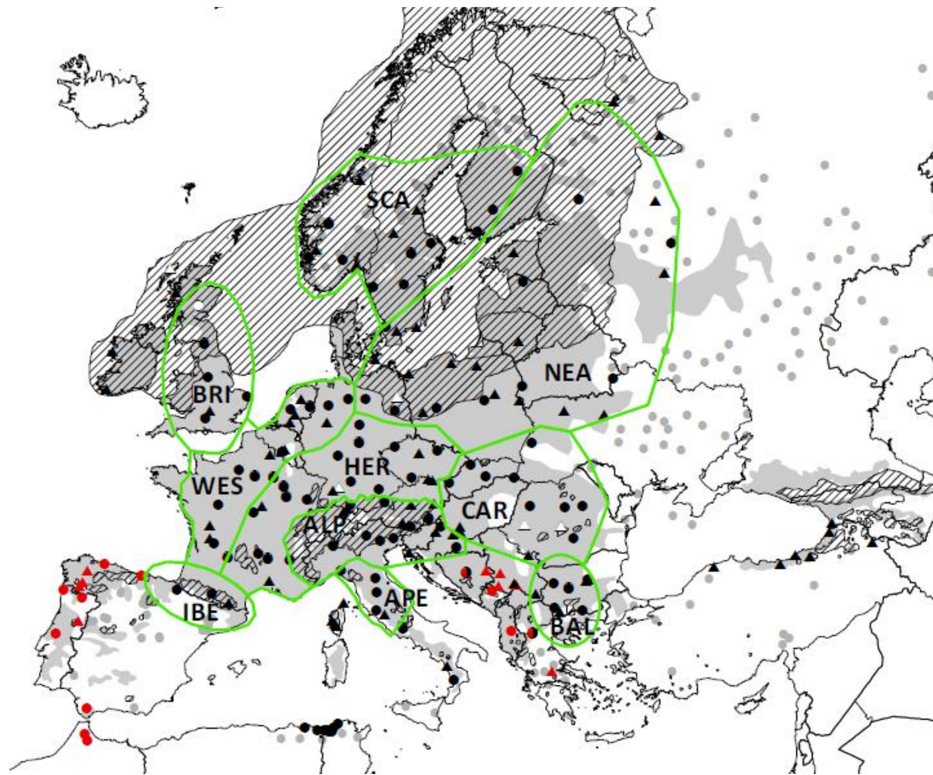


Fig. 1: Localities of *Alnus glutinosa* and main geographical regions recognized in this study. The current distribution of *A. glutinosa* (according to EUFORGEN 2009, www.euforgen.org) is shaded, and the extent of glaciers in the LPG is hatched. Black symbols – diploid populations, red symbols – tetraploid populations, black-and-red symbols – mixed-ploidy populations (di-tri-tetra), white symbols – unknown ploidy level; populations with detailed sampling for microsatellite analysis are marked by circles. Ploidy levels according to Mandák*et al.,* in review (for methods see Appendix S1). European regions are highlighted by green (i.e. IBE – Iberian Peninsula, APE – Apennine Peninsula, BAL – Balkan Peninsula, HER – Hercynian Mountains and Massif Central, ALP – Alps, CAR – Carpathians, BRI – British Isles, SCA – Scandinavia, NEA – Baltic and Northeastern European plains and WES – Western European plains). For Structure analyses of diploid populations, the IBE region was included as a part of the WES region.

We tested two alternative scenarios with different consequences for possible lineage admixture. The first scenario supposed the existence of source refugia for postglacial expansion exclusively in Southeastern Europe and an east-to-west decline in population genetic diversity corresponding with the direction of European colonization. This scenario follows Huntley & Birks (1983), who placed the main source refugia for *A. glutinosa* expansion after the LPG in the Eastern Alps, the Carpathians and the Ukrainian lowlands, as well as King & Ferris (1998), who accepted this east-west migration pattern. King & Ferris (1998) revealed 13 cpDNA haplotypes of *A. glutinosa*, mainly associated with Southern European eninsulas. They suggested that two of these haplotypes colonized

Northern and Central Europe from LPG refugia located in the Carpathians. The second scenario supposed the existence of multiple source refugial areas of *A. glutinosa* in Europe, at least in Western and Eastern Europe, probably accompanied by multiple secondary contact zones and lineage admixture. Several studies have uncovered valuable clues supporting this scenario: (1) The colonization of the British Isles originated in the Western European refugium (Bush *et al.* 1987; Chambers & Elliott 1989; Bennett & Birks 1990; Douda *et al.* 2014) rather than in Eastern Europe. (2) Scandinavia might have been colonized from Northeastern European populations, but there is also the possibility that eastern populations could have spread west and mixed with western populations before expansion north (Bush *et al.* 1987; Douda *et al.* 2014).

To assess the two scenarios, we aimed to reveal postglacial migration patterns of the temperate tree *A. glutinosa* by analyzing populations across its whole range by studying two molecular markers, cpDNA and microsatellites. Our main questions were: (1) Do the migration patterns of *A. glutinosa* correspond to the scenario of expansion from Southeastern Europe or from multiple Southern European refugia? (2) Which routes did *A. glutinosa* follow during postglacial colonization? (3) Did different genetic lineages meet in secondary contact zones and give rise to genetically admixed populations?

Materials and Methods

Study species

Alnus glutinosa (L.) Gaertn. (Betulaceae) is a wind-pollinated, self-incompatible, keystone tree species of riparian and water-logged habitats (McVean 1953; Douda *et al.* 2009; Douda 2010). *A. glutinosa* is distributed in lowlands and middle altitudes across the European continent except in the extreme north (Fig. 1). It extends as far as West Siberia and the mountains of Western Turkey and North Africa (McVean 1953). Its seeds are dispersed by water over longer distances and also locally by wind (McVean 1953). *A. glutinosa* is a long-lived tree (c. 100–120 years) that reproduces between the ages of 10 and 20 years (McVean 1953). It has recently been shown that *A. glutinosa* occurs in two ploidy levels, diploid and

tetraploid (Lepais *et al.* 2013; Mandák *et al.*, in review). Tetraploid populations have been found to form two geographically very well delimited clusters located in the Iberian Peninsula and in the Dinaric Alps (Mandák *et al.*, in review; Fig. 1).

Sampling and DNA extraction

In total, 190 populations of *A. glutinosa* from natural forests were sampled between 2011 and 2013 throughout most of the species' distribution range (Fig. 1 and Table S1, Supporting information). All samples were stored in silica gel. Before DNA analyses, the ploidy level of most collected individuals was checked by flow cytometry (Mandák *et al.*, in review; for methods see Supporting information S1). We applied a distinct sampling strategy for each type of molecular marker: (1) **Chloroplast DNA** (cpDNA) was analysed in samples from all populations, ideally in three samples per population, with 0.5 km distance between samples (Table S1, Supporting information). Fruits originating from 54 out of 190 populations were obtained from the International Alder Seed Bank at the Research Institute for Nature and Forest in Belgium (marked with KF in Table S1, Supporting information) and grown in the experimental garden of the Institute of Botany, Czech Academy of Sciences, Průhonice, Czech Republic (49.991667, 14.566667, *ca.* 320 m above sea level). The seeds were germinated in 5 × 5 cm bedding cells with homogenous garden compost and later moved to 19 × 19 × 19 cm (6.9 L) pots filled with common garden substrate. (2) **Detailed sampling for microsatellites** (SSR) analyses was carried out in 90 populations (Table S1, Supporting information). These populations were at least 100 km apart. In each population, leaves were collected along a linear transect from 20 trees at least 50 m apart (in some cases fewer trees were available, see Table S1, Supporting information).

DNA for microsatellite analysis was extracted using the DNeasy 96 Plant Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. For cpDNA analysis DNA was extracted using the method of Štorchová *et al.* (2000). The quality and yield of isolated DNA was checked on 1% agarose gels, and the DNA concentration was then evaluated using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA).

cpDNA analysis

The *ndh*F-*rpl*32, *psb*J-*pet*A and *3'rps*16-*5'trn*K intergenic spacers were amplified using the primers published by Shaw *et al.* (2007). The PCRs were performed in 25 µl containing 1x Plain PP Mastermix (TopBio, Prague, Czech Republic), 0.2 mM of each primer and a few nanograms of genomic DNA. The cycling conditions were: 95°C for 3 min, followed by 40 cycles of 95°C for 30s, 49°C (*ndh*F-*rpl*32, *psb*J-*pet*A) or 52°C (*3'rps*16-*5'trn*K) for 30s and 72°C for 2 min. The reactions were completed by a final elongation step at 72°C for 15 min. The PCR products were checked on 1% agarose gels and sent to Macrogen (Amsterdam, the Netherlands) for sequencing. Sequencing of both DNA strands was performed using the original PCR primers. For the *psb*J-*pet*A region, an additional sequencing reaction using the primer cp6-310F (5'-TCTGTTCCTTGATAGTATCTGTCG-3') was performed to overcome the difficulties stemming from the presence of a long poly-C stretch.

Microsatellite analysis

We analysed nuclear genetic variation at 19 nuclear microsatellite loci. These loci have been cross-amplified from closely related species by Drašnarová *et al.* (2014) (Multiplex 1: A2, A7, A10, A22, A26, A35, A37, A38) or developed specifically for *A. glutinosa* by Lepais & Bacles (2011) (Multiplex 2: Ag1, Ag5, Ag9, Ag10, Ag13, Ag20, Ag23, Ag25, Ag27, Ag30, Ag35). For detailed characteristics of microsatellite loci, see Lepais & Bacles (2011) and Drašnarová *et al.* (2014).

PCRs were performed in a total volume of 5 µl using 1× Qiagen Multiplex PCR Master Mix (Qiagen), 0.05–0.8 µM of each primer and 1 µl of template DNA (20–25 ng/µl). Reactions were run on a Master Cycler Pro (Eppendorf, Hamburg, Germany) thermal cycler using the following conditions for Multiplex 1: 15 min of denaturation at 95°C, followed by 40 cycles at 94°C for 30 s, with annealing temperature at 58°C for 30 s, extension at 72°C for 1 min, and a final extension at 72°C for 10 min and Multiplex 2: 5 min of denaturation at 95°C, followed by 30 cycles at 95°C for 30 s, with annealing temperature at 58°C

for 3 min, extension at 72°C for 30 s and a final extension at 60°C for 30 min. One microliter of PCR product (ten times diluted) was mixed with 0.1 µl of GeneScan-500 LIZ internal size standard (Applied Biosystems, Foster City, CA, USA) and 12 µl of Hi-Di formamide (Applied Biosystems) and electrophoresed using an ABI PRISM 3130 Genetic Analyser (Applied Biosystems). Allele size was determined using GeneMarker version 2.4.0 (SoftGenetics, State College, PA, USA). For each marker, automatic allele calling was implemented by defining bins with subsequent manual correction. A single-locus genotype was coded as missing data point after at least two amplification failures. Micro-Checker (Van Oosterhout *et al.* 2004) applied to diploid populations found no evidence of large allele dropout for any of the loci. The analysis pointed to 6.14% of population/marker combinations suggesting the presence of null alleles involving mostly loci A37, Ag13 and Ag14. The null allele frequency averaged over loci was very low, 3.2% per population at maximum which could only slightly bias classical estimates of population differentiation (Chapuis & Estoup 2007). The mean genotyping error rate was not assessed for the present study, but based on previous studies, it was low: 1.3% for Multiplex 1 (Drašnarová *et al.* 2014) and 0.24% for Multiplex 2 (Lepais *et al.* 2013).

Statistical analysis

cpDNA

The sequences were proofread using Chromas Lite 2.01 (Technelysium Pty. Ltd., Brisbane, Queensland, Australia) and aligned manually in BioEdit 7.4.0.1 (Hall 1999). Sequences of the three cpDNA regions were combined into a single dataset using FaBox (Villesen 2007). Mononucleotide repeats were excluded due to a potentially high level of homoplasy (Ingvarsson *et al.* 2003) before processing the data further. Indels were coded by the simple gap coding method (Simmons & Ochoterena 2000) as implemented in SeqState 1.4.1 (Müller 2005). A chloroplast DNA haplotype network was constructed using TCS 1.21 (Clement *et al.* 2000). A considerable part of the observed genetic variation was represented by insertions and deletions. However, TCS cannot work with the presence/absence matrix created for indels and appended by SeqState to the sequence alignment

file. To include this information in the analysis, the 0/1 characters were replaced manually by A/T in the nexus input file. The analysis was then run with default conditions (i.e. with gap characters treated as missing data), so the software could analyse information on indels as A/T variation, but the original gap characters (represented by "-" in the sequence alignment) were not taken into account. This approach allowed the incorporation and equal weighting of indels with different lengths in the analysis. In order to identify groups of related haplotypes, phylogenetic analyses were performed. Each haplotype was represented by one sequence, and sequences of *Alnus viridis* from the locality Zlatá Koruna, Czech Republic (48.851488, 14.337864) were included as the outgroup. Bayesian analysis using MrBayes v 3.2.2 (Ronquist & Huelsenbeck 2003) and maximum parsimony (MP) analysis using PAUP* 4.0.b10 (Swofford 2002) were performed for this purpose. For the Bayesian approach, the data were divided into two parts (nucleotide sequence and presence/absence matrix for coded indels) prior to the analysis. The F81+I+G model was selected as the appropriate model of nucleotide substitution by HLR tests as implemented in MrModeltest 2.3 (Nylander 2004) for the sequence data. The default model settings were kept for the indel data. Two replicate analyses with four chains each were computed for 1.6 million generations, sampling every 1,000-th generation. After this number of runs the average standard deviation of split frequencies reached a value lower than 0.01, indicating that convergence was reached. The first 400 trees per run (25%) were discarded as burn-in, and the remaining 2,402 trees were used to construct a consensus tree. MP analysis was performed using heuristic search with 100 replicates of random sequence addition and TBR branch swapping. Bootstrap analysis with 1,000 replicates was performed to evaluate the support of the resulting clades.

Microsatellites

In order to quantify levels of genetic variation within populations, the number of alleles (N_A), rarefied allelic richness (A_r), Nei gene diversity corrected for sample size (H_e, Nei 1978) and the inbreeding coefficient (F_{IS}) were calculated in the software SPAGedi (Hardy & Vekemans 2002), which can deal with different

ploidy levels (di-, tri- and tetra-ploids in our case). To account for differences in ploidy level, we computed allelic richness and gene diversity also for haploid genomes (hA_r and hH_e). The significance of the inbreeding coefficient was tested by 999 permutations.

We used the individual-based Bayesian clustering method implemented in Structure 2.3.3 (Pritchard *et al.* 2000) to investigate population subdivision. Ten independent analyses were carried out for K = 1–30 using a burn-in period of 25,000 steps followed by 100,000 Markov Chain Monte Carlo iterations. An admixture ancestry model and a Correlated Allele Frequency model were used for all computations. The trees showing the relationships between clusters was constructed using MEGA v. 6.0 (Tamura *et al.* 2013), based on a pairwise matrix of the net nucleotide distances computed in Structure (Pritchard *et al.* 2010).

Two datasets were assembled. (1) All populations sampled in detail for microsatellite analysis (i.e. 90 diploid and tetraploid populations). Due to allele copy ambiguity in this dataset which encompassed polyploid individuals, determination of precise genotype and allele frequencies was impossible. Here we employed coding of microsatellite data using the Polysat package (Clark & Jasieniuk 2011) in R 3.1.1 (R Core Team 2014) for further analyses in Structure. Polyploid (ambiguous) genotypes were coded as follows: in cases of two observed alleles A and B, the genotype was coded as A, B, B, B. In cases of three observed alleles A, B and C, the genotype was coded as A, B, C, C. Diploid individuals were coded according to their two alleles followed by a symbol denoting missing alleles, i.e. A, A, –9, –9 for homozygotes and A, B, –9, –9 for heterozygotes. We accounted for genotypic ambiguity in polyploid individuals by using the Recessive Allele model (Falush *et al.* 2007) with no prior population information (Hubisz *et al.* 2009).

(2) European diploid populations (i.e. 68 diploid populations from Europe) provided us with codominant data with no need to use the Recessive Allele model. However, due to a weak population structure characterized by very low F_{ST} values (0.0613), we used a model which incorporated *a priori* sampling location information (Hubisz *et al.* 2009), i.e. a “locprior” model. This improved model has the advantage of allowing cryptic structures to be detected at a lower level of

divergence and does not bias towards detecting structure spuriously when none is present; it especially helps in situations when the standard structure models do not provide a clear signal of structure (Hubisz *et al.* 2009). Nine groups of populations were used as priors when populations were grouped into regions differing in environmental conditions during the LPG and the Holocene, and the nature of expansion according to results of palaeoecological analysis (Douda *et al.* 2014), i.e. Apennine Peninsula (APE), Balkan Peninsula (BAL), Hercynian Mountains and Massif Central (HER), the Alps (ALP), the Carpathians (CAR), British Isles (BRI), Scandinavia (SCA), Baltic and Northeastern European plains (NEA) and Western European plains (WES) (Fig. 1 and Table S1, Supporting information). Two alternative methods were used to explore the true number of *K*: (1) similarity of Structure runs (Nordborg *et al.* 2005) and (2) delta *K* (ΔK) (Evanno *et al.* 2005). Both were calculated using the Structure-sum script (Ehrich *et al.* 2007) in R 3.1.1 (R Core Team 2014). *K* values with high similarity among independent Structure runs and high ΔK values that capture most of the structure in the data and seem to be biologically relevant were considered optimal. Alignment of cluster assignments across replicate analyses was then conducted in CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and subsequently visualized using DISTRUCT 1.1 (Rosenberg *et al.* 2007).

The Approximate Bayesian Computation (ABC) method implemented in DIYABC v. 2.0.4 (Cornuet *et al.* 2014) was used to explore likely effective refugia and migration routes and to estimate associated demographic parameters of European diploid populations sampled in detail for microsatellite analysis. Moreover, the ABC approach can reveal admixture among colonizing populations and it was shown that it leads to accurate parameter estimates when using population admixture models (Excoffier *et al.* 2005). Due to computational limitations and an infinity of possible scenarios when numerous populations were considered, inferences were based on ten geographically delimited groups corresponding to those previously used as priors in Structure analysis (Fig. 1 and Table S1, Supporting information). In addition, the group comprising populations in the Western European plains (WES) was divided into Iberian populations (IBE) located in a putative glacial refugium in the north–east part of the Iberian

Peninsula and the remaining populations belonging to the WES group. Hence, the delimited groups of populations represented southern refugia (IBE, APE and BAL), putative cryptic northern refugia (WES, ALP, CAR, NEA and HER) and newly colonized areas after the withdrawal of the ice sheet (SCA and BRI). In total, 24 demographic scenarios in seven steps were compared to test the origin of each group except for IBE, APE and BAL which are considered as ancestral southern refugia (Fig. 2). We tested the origin of each group based on adjacent groups and tried to simplify the models as much as possible because increasing the complexity of the model can result in poor estimation of parameters (Bertorelle *et al.* 2010). Demographic scenarios were described by divergence times in generations (t1, t1a, t2, t2a), effective population size of putative ancestral (NA), standing (N1, N2, N3) and founding (N1a, N2a, N3a) populations (specific values of prior distributions are presented in Fig. 2). Prior distributions for effective population size of founding populations [10–10000] and putative ancestral populations [10–100000] were broad so as to explore a wide range of population sizes. Prior distributions for divergence times (in generations) were chosen to represent last glaciations [1000–10000], early-Holocene expansion [500–1000] and Holocene expansion [10–1000]. The prior distribution for admixture rate was as broad as possible [0.001–0.999]. Default values were used for genetic parameters (for specific values see Fig. 2), assuming a Generalized Stepwise Mutation model (Estoup *et al.* 2002) in which a mutation increases or decreases the number of repeated motifs by one or several units. The observed and simulated genetic datasets were summarized using the mean number of alleles, mean size variance and M index (Excoffier *et al.* 2005) for each population and for each pair of populations. The classification index (Rannala & Mountain 1997), F_{ST} (Weir & Cockerham 1984) and shared allele distance (Chakraborty & Jin 1993) were computed only among populations. One million simulations were run for each scenario. Subsequently, 1% of the simulated datasets closest to the real genetic dataset were used to estimate posterior probabilities (with 95% confidence intervals) for each scenario using logistic regression (Cornuet *et al.* 2008). Posterior parameter distributions were estimated from 1% of the closest datasets

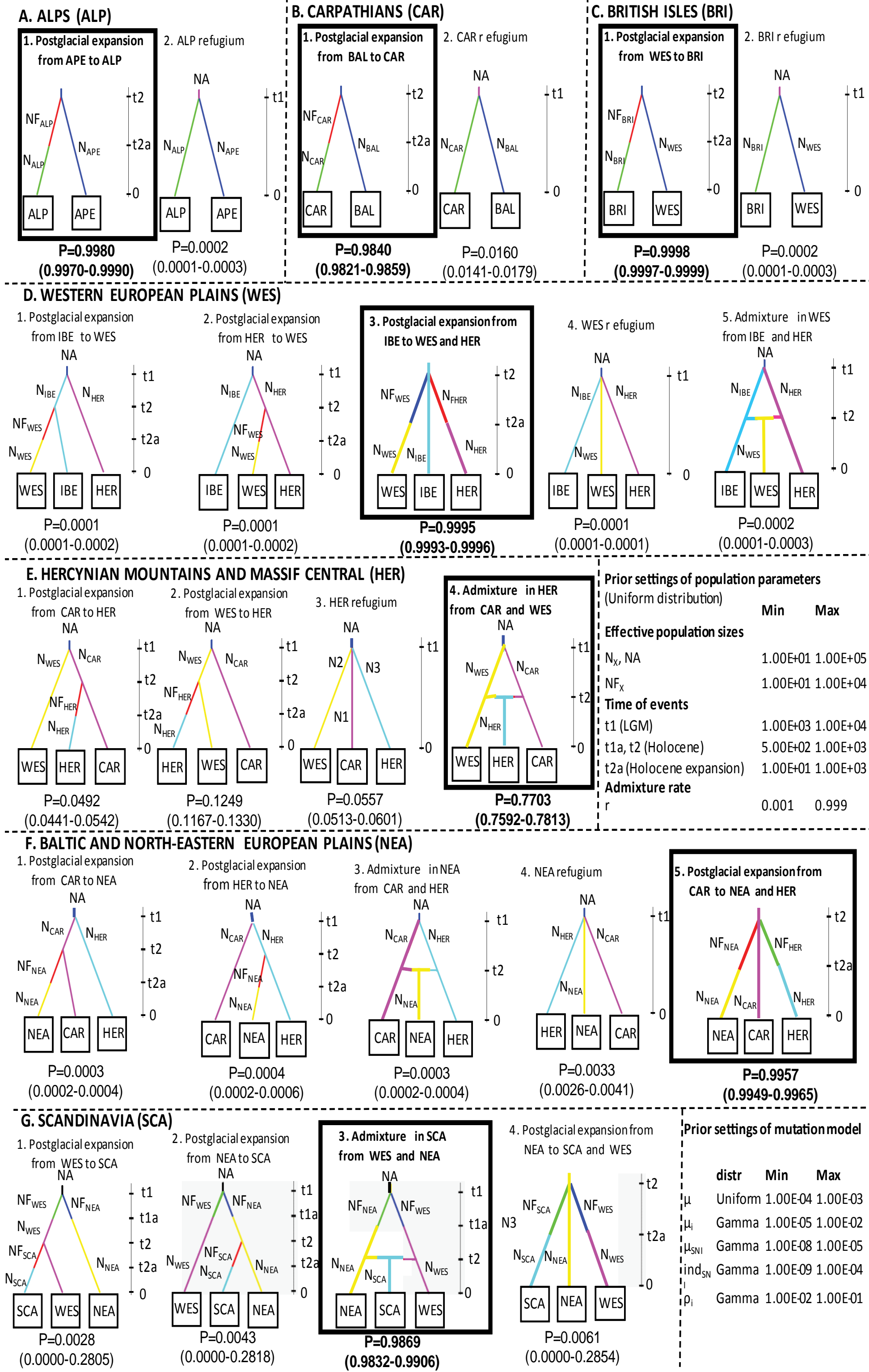


Fig. 2: Graphical representation of 24 competing scenarios tested in seven steps (A-G) by approximate Bayesian computation. NA, N_X and NF_X refer to effective sizes of putative ancestral, standing and founding populations, and t1, t1a, t2 and t2a to divergence times (prior settings of population parameters and mutation models are mentioned within the Figure). Posterior probabilities (P) of the scenarios and 95% confidence intervals of P (in brackets) computed using a logistic regression estimate are given under each scenario. The most probable scenario for each step is framed by black rectangle.

simulated according to the most likely scenario for each step. To evaluate the level of confidence to which we can trust the previous analysis, we simulated 500 pseudo-observed datasets (pods) with each scenario using the mode of posterior distributions as demographic parameters. Each of these datasets consisted of pseudo-observed data, which was analysed using all the simulated datasets previously obtained for ABC estimation with the actually observed data. Parameter values for scenarios were drawn from previously used prior distributions. Subsequently we calculated the false negative rate as the fraction of pods generated under the focal scenario that support other scenarios and the false positive rate as the fraction of pods generated under all the other scenarios that support the focal scenario (Bertorelle *et al.* 2010). We performed a model checking analysis by comparing the first two axes from a PCA of observed summary statistics with those obtained from 1000 simulations based on the posterior predictive distribution of the best fitting model (Cornuet *et al.* 2010). Lastly, we measured the discrepancy between a model parameter posterior combination and a real dataset by considering various sets of quantities. Firstly, we calculated mean relative bias (Bias) and the square root of the relative mean square error (RRMSE) that depend, respectively, on the sum of differences and on the sum of squared differences, between the 1000 estimates of each parameter obtained from the pseudo-observed dataset and the respective modes estimated from the observed data (‘true values’). A value of 0 means that the mode estimated the parameter with no bias. Positive and negative values reflect biases towards overestimation and underestimation, respectively. We also calculated the factor 2 statistic, defined as the proportion of the 1000 estimates lying between half and double of the true value, and the 50% and 90% coverage, defined as the proportion of times the true value falls within the 50% and 90% credible interval of the 1000 estimates. During this analysis, posterior distributions of the same parameters were used. The data were thus simulated from the posterior predictive distribution (Cornuet *et al.* 2014).

Results

Phylogenetic relationships and geographic distribution of chloroplast haplotypes

A total of 43 haplotypes and nine haplogroups showed strong geographic structure across the distribution range of *A. glutinosa* (Fig. 3 and Table S1, Supporting information). Southern parts of the distribution range of *A. glutinosa* are characterized by high cpDNA variation with distinct regional haplogroups. Four of these groups belong to tetraploid populations in Morocco (Group III), the Iberian Peninsula (Group II) and the western Balkan Peninsula (Groups VIII, IX). Distinct diploid haplogroups occur in Romania (Group V), Bulgaria (Group VII), the Anatolian Peninsula (Group VI) and central Mediterranean area (Group I); the last group connects populations in southern Italy and Corsica with populations in northern Tunisia and Algeria. In contrast, a single diploid haplogroup (Group IV) covers the rest of Europe, with a predominant occurrence of haplotype 1 present in 69% of populations. Grouping of seven haplotypes (9, 24, 25, 30, 35, 38 and 39) was not supported by phylogenetic analysis (Fig. 3D).

Population genetic diversity and structure inferred from microsatellite markers

We identified 306 alleles at 19 microsatellite loci, with an average of 16.11 alleles per locus across all populations (Table S1, Supporting information). Mean allelic richness (hA_r) and gene diversity (hH_e) were 3.62 and 0.56 (Table S1, Supporting information), respectively, reaching lowest values in diploid southern populations especially in North Africa and the Iberian and Apennine Peninsulas. Higher genetic diversity was recorded in tetraploid populations (Fig. 4 and Table S1, Supporting information) but surprisingly also in Northern European diploid populations (Fig. 4). The mean value of F_{IS} for diploid populations was 0.040 and in most cases F_{IS} per population were not significantly different from zero (Table S1, Supporting information).

In the combined Bayesian clustering analysis of diploid and tetraploid populations, ΔK and the similarity among Structure runs indicated that two and four clusters best explained the genetic structuring of *A. glutinosa* populations (Fig. S1 and Fig. S2, Supporting information). Although populations were

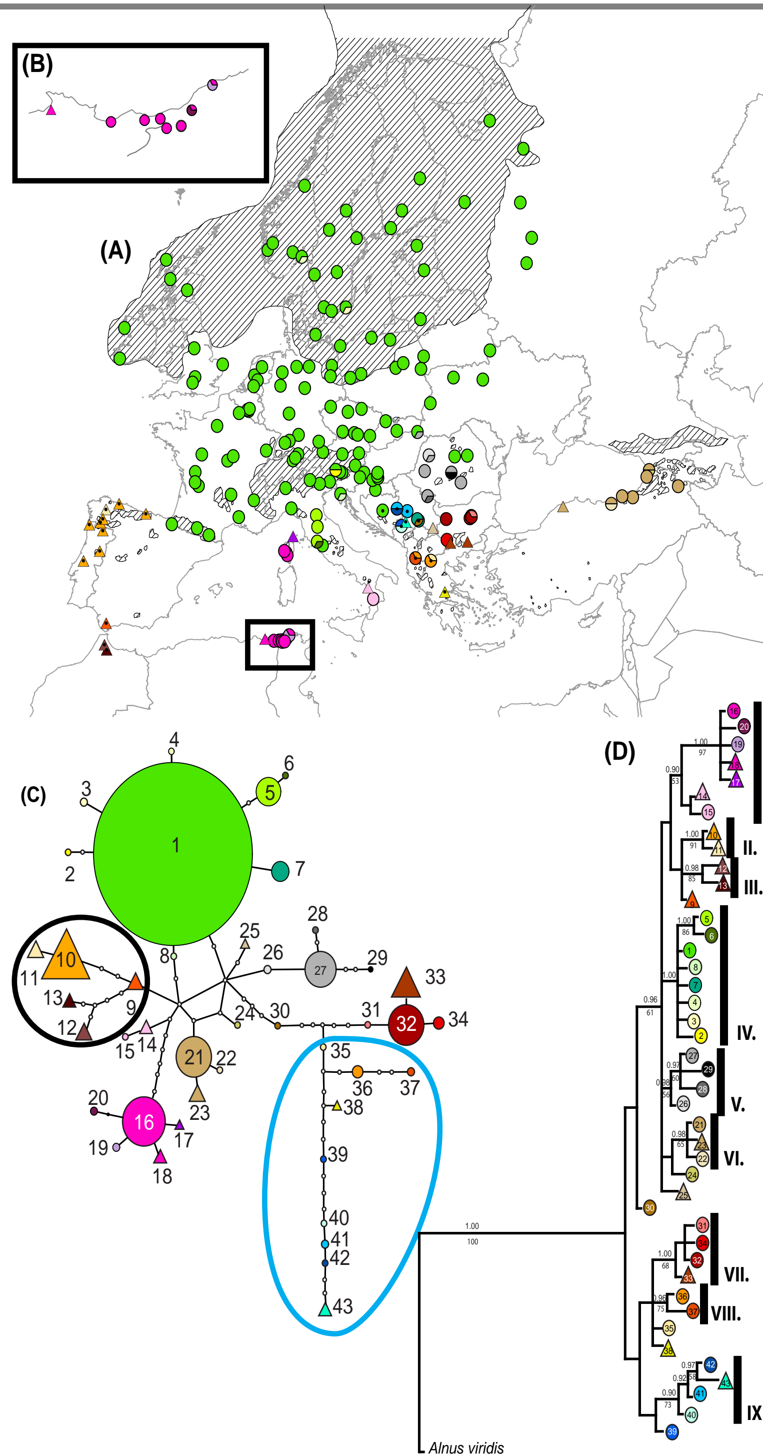


Fig. 3: *Alnus glutinosa* chloroplast haplotypes. Individual cpDNA haplotypes are represented by differently coloured circles or triangles. **(A)** Distribution of haplotypes across all sampled populations and **(B)** across Algerian and Tunisian populations; pie charts represent haplotype proportions of one to three analysed individuals per population; black dots in the centre of symbols indicate tetraploid populations. **(C)** cpDNA haplotype network; lines represent single mutational steps; small open circles indicate missing haplotypes (extinct or unsampled); the size of each circle and triangle is proportional to the frequency of the particular haplotype; tetraploids are outlined by the black (Iberian and Moroccan) and blue (western Balkan) line. **(D)** Fifty per cent majority-rule consensus tree of the Bayesian phylogenetic analysis; numbers above branches are Bayesian posterior probabilities, and numbers below branches indicate bootstrap support (in percentages) from the maximum parsimony analysis.

assigned to two clusters corresponding to the division between Iberian plus North-African tetraploids and the rest of the populations, assignment to four clusters provided more information concerning the distribution of ploidy levels in the dataset (Fig. S2, Supporting information). Those four groups were Iberian-Moroccan tetraploids, western Balkan tetraploids, North-African diploids and European diploids (Fig. 5A).

When analyzing European diploid populations alone, the number of clusters that best explained the population genetic structuring based on the ΔK approach was two, three and six (Fig. S3A, Supporting information), and based on similarity among Structure runs the number of clusters was two, three, four and six (Fig. S3B, Supporting information). Two clusters distinguished populations in the Alps, the Apennine Peninsula and western parts of Slovenia and Croatia from populations in the rest of Europe (Fig. S4 and Fig S5A, Supporting information). Assignment to three and four clusters differentiated populations in Western,

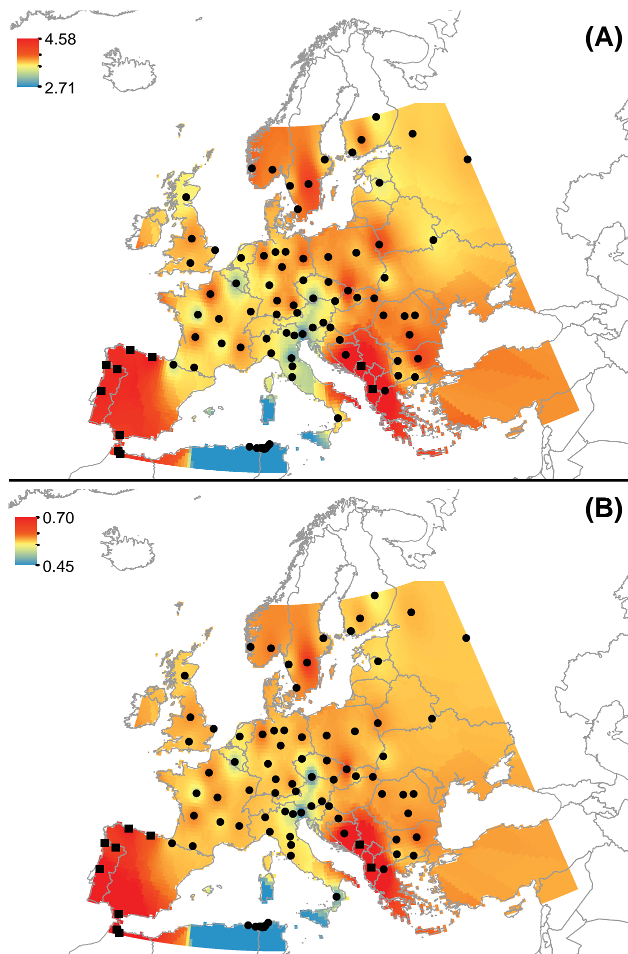


Fig. 4: Maps of allelic richness and gene diversity for all *Alnus glutinosa* populations. (A) allelic richness (hA_r) and **(B)** gene diversity(hH_e) corrcted for sample size and ploidy level calculated from microsatellite data. Black circles – diploid populations, black rectangles – tetraploid populations.

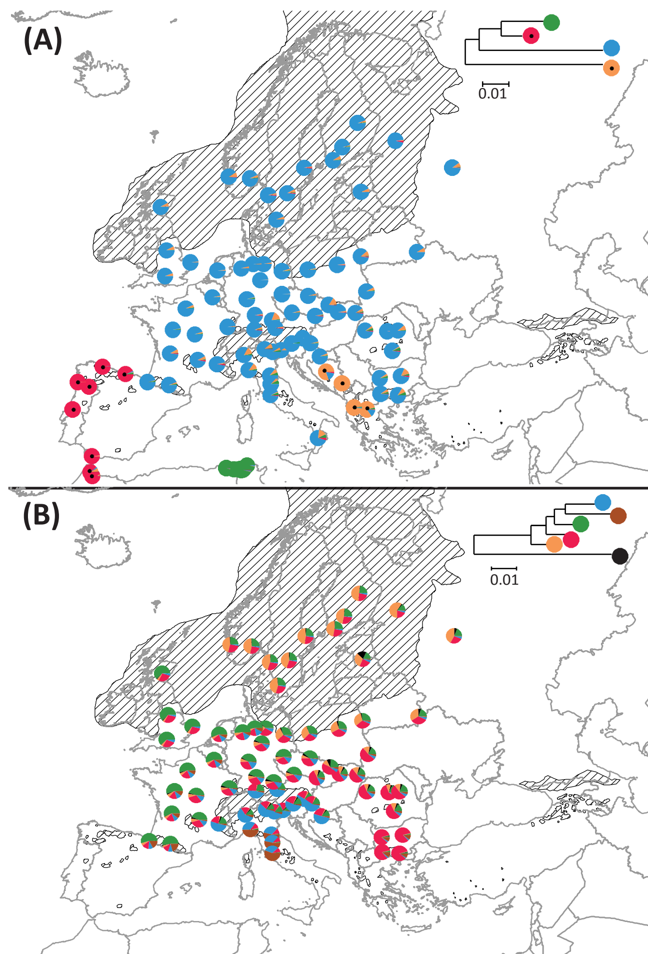


Fig. 5: Regional genetic structure of *Alnus glutinosa* populations as inferred by Structure. Individual assignment to (A) four clusters for all (diploid and tetraploid) 90 populations and (B) six clusters for 68 European diploid populations are visualized as pie charts. Each population was partitioned into several coloured segments proportionally to its membership in a given cluster; black dots in the centre of symbols indicate tetraploid populations; the trees showing the relationships between clusters are based on a pairwise matrix of the net nucleotide distances.

Eastern and Northern Europe, and in the Apennine Peninsula and surrounding regions (Fig. S4 and Fig. S5B,C, Supporting information). Assignment to six clusters more clearly differentiated Apennine and Balkan populations (Fig. 5B and Fig. S4, Supporting information).

Inferring population history using approximate Bayesian computation (ABC)

The ABC approach enabled testing several possible scenarios (Table 1 and Fig. 2) and provided us with detailed information about possible postglacial colonization routes and the origin of secondary contact zones between different genetic lineages (Fig. 6). Populations in all three Southern European refugia (IBE, APE and BAL) served as sources for the colonization of the rest of Europe (Fig. 6). From the northern part of the Iberian Peninsula, the expansion of *A. glutinosa* to Western Europe (WES) and Central Europe (HER) happened around 9270 yr BP

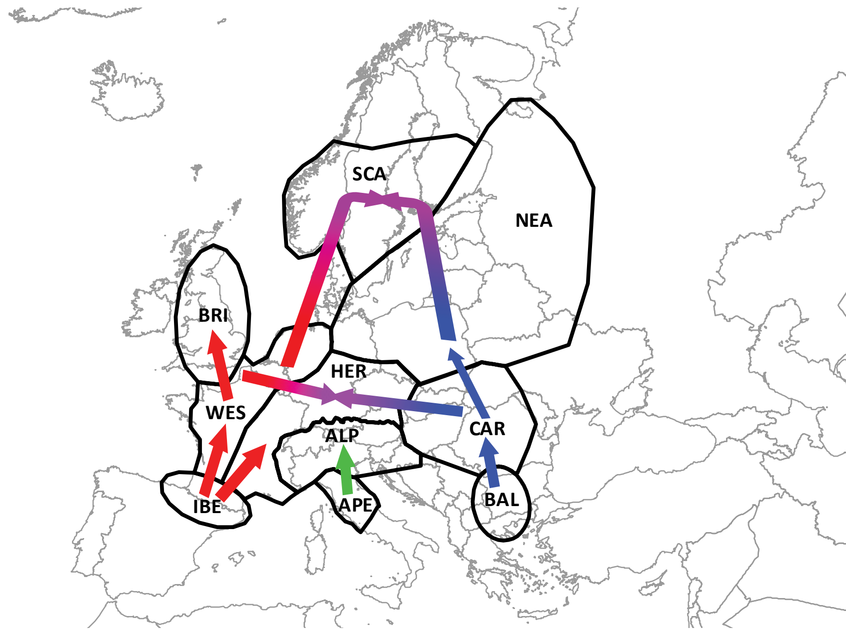


Fig. 6: Postglacial migration routes of *A. glutinosa* as inferred by approximate Bayesian computation. Most likely scenarios of postglacial migration into Alps (ALP), Carpathians (CAR), British Isles (BRI), Western European plains (WES), Hercynic Mountains and Massif Central (HER), Baltic and Northeastern European plains (NEA) and Scandinavia (SCA), supported by analyses based on approximate Bayesian computation. The arrows indicate the most likely migration routes with highest posterior probability values; connections between two arrows indicate admixture of lineages.

[7635–13830] (median estimates and 5-95% quantiles of the posterior distribution, using generation time 15 years), and the subsequent colonization of the British Isles (BRI) from western populations took place around 8700 yr BP [7575–13155]. Populations in the Alps (ALP) originated from the Apennine Peninsula (APE), and expansion time was estimated to be around 8895 yr BP [7590–13485]. Eastern Balkan populations (BAL) expanded to the Carpathians (CAR) around 10455 yr BP [7965–25050] and the expansion continued to Northeastern (NEA) Europe around 8565 yr BP [7560–12990]. Scandinavian populations have been established by admixture of colonizing lineages from Western (WES) and Northeastern (NEA) Europe (median estimates of admixture coefficient 0.486 [0.160–0.826]) estimated to have occurred around 9240 years BP [7650–12855]. When inferring the origin of populations growing in the area of the Hercynian Mountains and Massif Central, admixture from Western and Carpathian populations (median estimates of admixture coefficient 0.411 [0.114–0.784]) was estimated to have happened around 10590 yr BP [7800–14400].

Table 1: Scenarios tested by the ABC approach and their confidence. Origin of seven population group was tested by the ABC approach (i.e. ALP – Alps, CAR – Carpathians, BRI – British Isles, WES – Western Europe, HER – Hercynian Mountains and Massif Central, NEA – Baltic and Northeastern European plains and SCA – Scandinavia) when three group were considered as source refugia (i.e. IBE – Iberian Peninsula, APE – Apennine Peninsula and BAL – Balkan Peninsula). Three types of scenarios were tested: population group was established by expansion (→), admixture (+) or served as refugium (for detailed graphical representation of scenarios see Figure 2). The best scenario according to posterior probability is in grey. Confidence in scenarios was described by both a false positive and false negative rate for logistic regression.

	False negative	False positive
A. ALP		
APE → ALP	0.03	0.07
ALP refugium	0.08	0.04
B. CAR		
BAL → CAR	0.14	0.10
CAR refugium	0.09	0.13
C. BRI		
WES → BRI	0.05	0.09
BRI refugium	0.09	0.05
D. WES		
IBE → WES	0.12	0.10
HER → WES	0.07	0.08
IBE → WES and HER	0.09	0.17
WES refugium	0.20	0.10
IBE + HER = WES	0.15	0.15
E. HER		
CAR → HER	0.07	0.07
WES → HER	0.06	0.08
HER refugium	0.06	0.05
CAR + WES = HER	0.12	0.11
F. NEA		
CAR → NEA	0.11	0.08
HER → NEA	0.06	0.08
CAR + HER = NEA	0.13	0.11
NEA refugium	0.09	0.06
CAR → NEA and HER	0.06	0.14
G. SCA		
WES → SCA	0.01	0.01
NEA → SCA	0.02	0.04
WES + NEA = SCA	0.05	0.02
NEA → SCA and WES	0.00	0.02

ABC modelling did not provide us with a firm estimation of the exact time of expansion and admixture events because confidence intervals were wide and overlapped extensively. Nevertheless, model checking and testing the precision of estimating parameters confirmed the good quality of the best models in each step (Table S2 and Table S3, Supporting information). Specifically, it was shown that the 95% CI of the probability (P) of the most likely scenario never overlapped with

those of competing scenarios (Fig. 2). We also computed false positive and false negative rates and found that our method selected the true scenario with high confidence and markedly low false positive rates (Table 1).

Discussion

Thanks to the combination of two genetic markers with different variability and mode of inheritance (cpDNA and microsatellites), we were able to reconstruct detailed patterns of postglacial colonization of the temperate tree *Alnus glutinosa*. It also allowed us to challenge the classical paradigm that most European populations originated from refugial areas in the Carpathians (Huntley & Birks 1983; King & Ferris 1998). Moreover, we found strong evidence for the existence of secondary contact zones in Central and Northern Europe, where populations accumulated high genetic diversity thanks to admixture of genetic lineages from Eastern and Western Europe.

Haplotype divergence confirms multiple southern refugia, including one in the Carpathians

In southern areas, several divergent haplogroups and high haplotype diversity of *A. glutinosa* populations (Fig. 3 and Table S1, Supporting information) point to the existence of previously mentioned refugia in the Iberian, Apennine and Balkan Peninsulas, Corsica, North Africa and the Anatolian Peninsula (King & Ferris 1998; Lepais *et al.* 2013). Such genetic structure is typical of populations exposed to prolonged isolation under highly variable climatic conditions (Nieto Feliner 2014). The differentiation of individual haplogroups might be very ancient and reflect past biogeographic events (Petit *et al.* 2005; Médail & Diadema 2009). Haplogroup I comprising Corsican, north African and south Italian populations of *A. glutinosa* (Fig. 3), could be the result of migrations during the Messinian salinity crisis, 5.7–5.3 Myr BP, when Africa and Europe were connected for the last time (Ketmaier & Caccone 2013). Similar scenarios have been proposed for other plants (Petit *et al.* 2005) and animals (van der Made *et al.* 2006).

Recently, it was found that some southern populations located in Morocco (Lepais *et al.* 2013), the Iberian and the western Balkan Peninsula are tetraploid

(Mandák *et al.*, in review). They are probably autopolyploids with no indication that *Alnus incana*, the most related species, was involved in their evolutionary history (Mandák *et al.*, in review). These tetraploid populations possess specific haplotypes (Fig. 3) that indicate unique and very relict populations situated in the putative main glacial refugia. This means that neither the southern part of the Iberian Peninsula, nor the western part of the Balkan Peninsula served as effective refugial areas for the northward postglacial expansion of the species (Mandák *et al.*, in review).

Apart from classical Balkan refugia, we found specific haplogroups (Group V) pointing to the existence of refugial populations in the Carpathians (Fig. 3). The Carpathians have usually been hypothesized to be the starting point of postglacial colonization (King & Ferris 1998; Cornille *et al.* 2013) and survival of *A. glutinosa* during the LPG in this area is also supported by palaeoecological records (Douda *et al.* 2014). However, up to now, there has been no strong evidence pointing to LPG survival of temperate trees at latitudes higher than 45° (Tzedakis *et al.* 2013), therefore we cannot exclude the possibility that the current northern edge of this haplogroup's distribution in East Slovakia is the result of postglacial colonization rather than local survival of the glacial period.

Population history indicates that several genetic lineages colonized Northern Europe

The presence of one haplogroup represented mainly by one haplotype in the northern part of Europe (Fig. 3) may reflect postglacial expansion from the Carpathians (King & Ferris 1998), but it may also correspond to fragmentation of the continual *A. glutinosa* range during cold phases of the last glacial period (Douda *et al.* 2014). Similarly, some tree species which most likely survived the last glacial period in the northern part of Europe, such as *Betula pendula*, *B. pubescens*, *Populus tremula* and *Salix caprea* (Palmé *et al.* 2003a,b; Maliouchenko *et al.* 2007; Fussi *et al.* 2010), exhibit a low level of phylogeographic structure. Microsatellites provided a more detailed subdivision of recolonized areas compared to cpDNA (Fig. 5). The clear differentiation of populations evidenced by specific assignments to individual Structure clusters indicates that all three southern peninsulas served

as source refugial areas for postglacial colonization (Fig. 5B). The ABC approach, testing several possible scenarios (Table 1 and Fig. 2), selected the most probable directions in which lineages migrated and detected secondary contact zones between previously separated lineages (Fig. 6). The three main directions of postglacial expansion supported by our ABC analyses were: 1) from the northern part of Iberian Peninsula to Western and Central Europe and subsequently to the British Isles, 2) from the Apennine Peninsula to the Alps and 3) from the eastern part of the Balkan Peninsula to the Carpathians followed by expansion towards the Northern European plains (Fig. 6). During the colonization, two secondary contact zones were established: (i) Migration routes from Western Europe on the one hand and the Carpathians on the other met in Central Europe (HER), and (ii) Scandinavian populations originated from admixture of Western (WES) and Eastern (NEA) European populations (Fig. 6). The Carpathians were not directly confirmed to be a refugium (i.e. ancestral population) by the ABC analysis, but the wide confidence interval of the expansion time from the Balkan Peninsula to the Carpathians [7965–25550] points to the possibility that the Balkan population was the source for the colonization of the Carpathians in the Pre-LPG period. Postglacial migration from all three southern peninsulas has been supported for other tree species, namely *Fraxinus excelsior* (Heuertz *et al.* 2004a) and *Quercus* sp. (Petit *et al.* 2002), but also for animals such as *Mus musculus* (Rajabi-Maham *et al.* 2008) and *Erinaceus* spp. (Santucci *et al.* 1998). In many other organisms, and probably also in *A. glutinosa*, migration from the Apennine Peninsula might have been limited due to the barrier of the Alps (Taberlet *et al.* 1998; Hewitt 1999; Palmé & Vendramin 2002; Grivet & Petit 2003; Magri *et al.* 2006).

In contradiction to previously published data (Huntley & Birks 1983; King & Ferris 1998), our results suggest that multiple refugia served as sources for the postglacial colonization of Europe by *A. glutinosa*. We can therefore reject the suggestion proposed by King & Ferris (1998) that postglacial colonization of Europe happened from a single refugial area located in the Carpathians. The delayed expansion in Western Europe was therefore rather caused by factors other than distant position of glacial refugia in Eastern Europe (Huntley & Birks 1983; King & Ferris 1998). The same conclusion was reached in our palaeoecological study (Douda *et al.* 2014), which proposed delayed expansion in Western European

regions as a result of unfavourable climatic conditions. In addition, the delayed migration could also have been caused by competitive interactions with other species, as has been shown in the case of beech, the delayed migration of which was caused by competitive interactions with fir (Watson 1996). Recently, multiple mechanistic models applied to palaeoecological data found that plant competition among *Alnus*, *Betula*, *Quercus* and *Pinus* provided a better explanation for the observed population dynamics than growing season temperature or N availability during the Holocene period in Western Europe (Jeffers *et al.* 2014).

Multiple secondary contact zones triggered continent-wide lineage admixture

Genetic admixture seems to be the most important process during colonization of *A. glutinosa* attenuating the influence of repeated founder events which typically resulted in low population-genetic diversity. This genetic admixture maintained or even increased genetic diversity during recolonization of Central and Northern Europe by *A. glutinosa* (Fig. 4 and Fig. 6). If we consider that only diploid European populations of *A. glutinosa* were involved in postglacial colonization, genetic diversity generally remained equal or higher outside refugial areas, surprisingly reaching the highest values in Scandinavia (Fig. 4). The ABC approach identified genetic admixture between western and eastern lineages from Central Europe and Scandinavia as the most likely scenario describing the origin of populations in these regions (Fig. 2 and Fig. 6). Putative admixture zones in Central Europe, harbouring higher intra-population genetic diversity than in areas further south, have been ascertained for several woody species (Petit *et al.* 2003). A recent study on the postglacial colonization of *Taxus baccata* (Mayol *et al.* 2015) shows, using the ABC approach, that secondary contact between its eastern and western migration wave resulted in an admixture zone with high genetic diversity in Central Europe. On the other hand, high genetic diversity in Scandinavian populations is unique among European temperate tree species and has otherwise been observed only for boreal trees (Palmé *et al.* 2003a; Tollefsrud *et al.* 2009). Thus, multiple secondary contact zones between different lineages of *A. glutinosa* colonizing Northern Europe from southern refugial populations appear to be the main reason for high

genetic diversity of northern populations as a consequence of continent wide-admixture recorded outside of refugial areas of *A. glutinosa*.

According to recent findings, genetic admixture is a common feature of range expansion and probably increases the success of colonizing populations (Rius & Darling 2014). Genetic admixture is expected to create potential for selection of adaptive dispersal traits during expansion (Kolbe *et al.* 2004; Drake 2006; De Carvalho *et al.* 2010; Keller & Taylor 2010). In addition, several other processes maintaining high level of genetic diversity during postglacial colonization could be involved: (i) High genetic diversity can be maintained thanks to delayed reproduction, which allows multiple introductions of new seeds and therefore a large increase in the number of initial founders before reproduction begins (Austerlitz *et al.* 2000); (ii) adult trees might serve as pollen traps, and long-distance gene flow via pollen dispersal can be an important source of genetic diversity in wind-pollinated species (Latta & Mitton 1999; Liepelt *et al.* 2002; Richardson *et al.* 2002; Heuertz *et al.* 2004b); (iii) according to a simulation study, high genetic diversity can be preserved when populations migrate in wide corridors and genotypes are being reshuffled behind the migration front (Bialozyt *et al.* 2006) and (iv) the difference in genetic diversity between Scandinavia and the rest of Europe could also be due to a stronger reduction in recent population sizes in the latter due to higher anthropogenic activities.

Wetland habitats are strongly endangered by recent climate changes, which influence the water supply and flooding regime, altering the distributions of many species (Erwin 2008; Garssen *et al.* 2015). Our study brings evidence that populations of *Alnus glutinosa*, a keystone species of European swamp habitats, exhibit high genetic diversity in northerly located populations due to continent-wide admixture of different genetic lineages. On the contrary, southern refugial rear-edge populations of *A. glutinosa* represent current relicts with unique haplotype diversity. Limited genetic admixture resulting from spatial isolation and different ploidy levels of haplotype groups prevents enrichment of populations as was proposed for other European tree species as the result of strong dispersal limitations among isolated Mediterranean populations (Petit *et al.* 2003, 2005; Hampe & Petit 2005; Svenning & Skov 2004). This makes populations of *A. glutinosa* in Mediterranean region, i.e. the Iberian, Apennine and Balkan Peninsulas, Corsica,

North Africa and the Anatolian Peninsula, more vulnerable to extinction due to climate change.

Acknowledgements

We would like to thank Martina Hadincová and Eva Ničová for their help in the lab and Fred Rooks for helping us out with the language of the manuscript. Jana Doudová kindly helped us with obtaining cpDNA sequences. We also wish to thank all those that collected samples of *Alnus* for us, especially Vít Dvořák, Filip Kolář, Boudewijn Michiels and Milan Štech. Serge Müller and Cecile Fanny Emilie Bacles deserve thanks for sharing samples from North Africa. This study was supported by the Grant Agency of the Czech Republic (P504/11/0402) and grant CIGA no. 20124201 from the Czech University of Life Sciences in Prague (http://ga.czu.cz/) and as part of long-term research development project RVO 67985939.

References

Aguinagalde I, Hampe A, Mohanty A, Martín JP, Dumunil J, Petit RJ (2005) Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography*, **32**, 329–339.

Bennett K, Birks H (1990) Postglacial history of alder (*Alnus glutinosa* (L.) Gaertn.) in the British Isles. *Journal of Quaternary Science*, **5**, 123–133.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology*, **19**, 12–20.

Binney HA, Willis KJ, Edwards ME *et al.* (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews*, **28**, 2445–2464.

Birks HJB, Willis KJ (2008) Alpines, trees, and refugia in Europe. *Plant Ecology & Diversity*, **1**, 147–160.

Bush AMB, Hall AR, Bush MB (1987) Flandrian *Alnus*: expansion or immigration? *Journal of Biogeography*, **14**, 479–481.

Carcaillet C, Vernet J-L (2001) Comments on "The full-glacial forests of Central and Southeastern Europe" by Willis et al. *Quaternary Research*, **55**, 385–387.

De Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638–1650.

Clark LV, Jasieniuk M (2011) POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources*, **11**, 562–566.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Colosimo PF, Hosemann KE, Balabhadra S *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*, **307**, 1928–1933.

Cornille A, Giraud T, Bellard C *et al.* (2013) Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Molecular Ecology*, **22**, 2249–2263.

Cornuet J-M, 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.

Cornuet J-M, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *Bioinformatics*, **11**, 1–11.

Cornuet J-M, Santos F, Beaumont M a *et al.* (2008) Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Douda J (2010) The role of landscape configuration in plant composition of floodplain forests across different physiographic areas. *Journal of Vegetation Science*, **21**, 1110–1124.

Douda J, Čejková A, Douda K, Kochánková J (2009) Development of alder carr after the abandonment of wet grasslands during the last 70 years. *Annals of Forest Science*, **66**, 1–13.

Douda J, Doudová J, Drašnarová A *et al.* (2014) Migration patterns of subgenus *Alnus* in Europe since the last glacial maximum: a systematic review. *PloS one*, **9**, e88709.

Drake JM (2006) Heterosis, the catapult effect and establishment success of a colonizing bird. *Biology Letters*, **2**, 304–307.

Drašnarová A, Krak K, Vít P *et al.* (2014) Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genetics & Genomes*, **10**, 865–873.

Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. *Molecular Ecology*, **14**, 1697–1706.

Ehrich D, Gaudeul M, Assefa A *et al.* (2007) Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology*, **16**, 2542–2559.

Van Els P, Cicero C, Klicka J (2012) High latitudes and high genetic diversity: phylogeography of a widespread boreal bird, the gray jay (*Perisoreus canadensis*). *Molecular Phylogenetics and Evolution*, **63**, 456–465.

Erwin KL (2008) Wetlands and global climate change: the role of wetland restoration in a changing world. *Wetlands Ecology and Management*, **17**, 71–84.

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–1604.

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.

Excoffier L, Estoup A, Cornuet J-M (2005) Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, **169**, 1727–1738.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Fussi B, Lexer C, Heinze B (2010) Phylogeography of *Populus alba* (L.) and *Populus tremula* (L.) in Central Europe: secondary contact and hybridisation during recolonisation from disconnected refugia. *Tree Genetics & Genomes*, **6**, 439–450.

Garssen A, Baattrup-Pedersen A, Voesenek LACJ, Verhoeven JTA, Soons MB (2015) Riparian plant community responses to increased flooding: a meta-analysis. *Global Change Biology*, DOI: 10.1111/gcb.12921.

Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Molecular Ecology*, **14**, 4275–4285.

Grivet D, Petit R (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics*, **4**, 47–56.

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

Hampe A, Bairlein F (2000) Modified dispersal-related traits in disjunct populations of bird-dispersed Frangula alnus (Rhamnaceae): a result of its Quaternary distribution shifts? *Ecography*, **23**, 603–613.

Hampe A, Arroyo J, Jordano P, Petit RJ (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, **12**, 3415–3426.

Hardy O, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Heuertz M, Fineschi S, Anzidei M *et al.* (2004a) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology*, **13**, 3437–3452.

Heuertz M, Hausman J-F, Hardy OJ *et al.* (2004b) Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution*, **58**, 976–988.

Hewitt GM (1996) Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

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

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Huntley B (2014) Extreme temporal interpolation of sparse data is not a sufficient basis to substantiate a claim to have uncovered Pleistocene forest microrefugia. *New Phytologist*, **204**, 447–449.

Huntley B, Birks H (1983) *An atlas of past and present pollen maps for Europe: 0–13000 years ago*. Cambridge University Press., Cambridge, UK.

Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: *DNA fingerprinting: state of the science*, pp. 153–175. Birkhäuser,Basel, Switzerland.

Chambers F, Elliott L (1989) Spread and expansion of *Alnus* Mill. in the British isles: timing , agencies and possible vectors. *Journal of Biogeography*, **16**, 541–550.

Chapuis M, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Chun YJ, Nason JD, Moloney KA (2009) Comparison of quantitative and molecular genetic variation of native vs. invasive populations of purple loosestrife (*Lythrum salicaria* L., Lythraceae). *Molecular Ecology*, **18**, 3020–3035.

Ingvarsson P, Ribstein S, Taylor D (2003) Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. *Molecular Biology and Evolution*, **20**, 1737–1740.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jeffers ES, Bonsall MB, Froyd CA, Brooks SJ, Willis KJ (2014) The relative importance of biotic and abiotic processes for structuring plant communities through time. *Journal of Ecology*, **103**, 459–472.

Keller SR, Fields P, Berardi A, Taylor D (2014) Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. *Journal of Evolutionary Biology*, **27**, 616–627.

Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, **23**, 1720–1731.

Ketmaier V, Caccone A (2013) Twenty years of molecular biogeography in the West Mediterranean islands of Corsica and Sardinia: lessons learnt and future prospects. In: *Current Progress in Biological Research* (ed Silva-Opps M), pp. 71–93. InTech, Rijeca. Croatia.

King A, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Kolbe JJ, Glor RE, Rodríguez Schettino L *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

De Lafontaine G, Amasifuen Guerra CA, Ducousso A, Petit RJ (2014) Cryptic no more: soil macrofossils uncover Pleistocene forest microrefugia within a periglacial desert. *The New Phytologist*, 715–729.

De Lafontaine G, Ducousso A, Lefèvre S, Magnanou E, Petit RJ (2013) Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology*, **22**, 4397–4412.

Lait LA, Burg TM (2013) When east meets west: population structure of a high-latitude resident species, the boreal chickadee (*Poecile hudsonicus*). *Heredity*, **111**, 321–329.

Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution*, **53**, 769–776.

Lepais O, Bacles CFE (2011) De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. *The Journal of Heredity*, **102**, 627–632.

Lepais O, Muller SD, Ben Saad-Limam S *et al.* (2013) High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PloS one*, **8**, e75029.

Liepelt S, Bialozyt R, Ziegenhagen B (2002) Wind-dispersed pollen mediates postglacial gene flow among refugia. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 14590–14594.

Liepelt S, Cheddadi R, de Beaulieu J-L *et al.* (2009) Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) — A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139–149.

Van der Made J, Morales J, Montoya P (2006) Late Miocene turnover in the Spanish mammal record in relation to palaeoclimate and the Messinian Salinity Crisis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238**, 228–246.

Magri D, Vendramin GG, Comps B *et al.* (2006) A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *The New Phytologist*, **171**, 199–221.

Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, 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.

Mandák B, Hadincová V, Mahelka V, Wildová R (2013) European invasion of North American *Pinus strobus* at large and fine scales: high genetic diversity and fine-scale genetic clustering over time in the adventive range. *PLoS one*, **8**, e68514.

Marrs RA, Sforza R, Hufbauer RA (2008) Evidence for multiple introductions of *Centaurea stoebe micranthos* (spotted knapweed, Asteraceae) to North America. *Molecular Ecology*, **17**, 4197–4208.

Mayol M, Riba M, González-Martínez SC, Bagnoli F, de Beaulieu J-L, Berganzo E, Burgarella C, Dubreuil M, Krajmerová D, Paule L, Romšáková I, Vettori C, Vincenot L, Vendramin GG (2015) Adapting through glacial cycles: insights from a long-lived tree (*Taxus baccata*). doi: 10.1111/nph.13496

McVean D (1953) Account of *Alnus glutinosa* (L.) Gaertn. for the Biological Flora of the British Isles. *Journal of Ecology*, **41**, 447–466.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.

Mullen LM, Hoekstra HE (2008) Natural selection along an environmental gradient: A classic cline in mouse pigmentation. *Evolution*, **62**, 1555–1570.

Müller K (2005) SeqState: primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics*, **4**, 65–69.

Muller MH, Leppala J, Savolainen O (2008) Genome-wide effects of postglacil colonization in *Arabidopsis lyrata*. *Heredity*, **100**, 47–58.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nieto Feliner G (2014) Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 265–278.

Nordborg M, Hu TT, Ishino Y *et al.* (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biology*, **3**, 1289–1299.

Nylander J (2004) *MrModeltest v2. Program distributed by the author*. Evolutionary Biology Centre, Uppsala University.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Palmé AE, Semerikov V, Lascoux M (2003a) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003b) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology*, **12**, 201–212.

Palmé AE, Vendramin GG (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology*, **11**, 1769–1779.

Petit R, Aguinagalde I, de Beaulieu J-L *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Petit J, Brewer S, Cheddadi R *et al.* (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Petit RJ, Hampe A, Cheddadi R (2005) Climate changes and tree phylogeography in the Mediterranean. *Taxon*, **54**, 877–885.

Potter KM, Jetton RM, Dvorak WS *et al.* (2012) Widespread inbreeding and unexpected geographic patterns of genetic variation in eastern hemlock (*Tsuga canadensis*), an imperiled North American conifer. *Conservation Genetics*, **13**, 475–498.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen X, Falush D (2010) Documentation for structure software: Version 2.3. Available from: http://pritch.bsd.uchicago.edu/structure.html/.

R Core Team (2014) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available from: http://www.r-project.org/.

Rajabi-Maham H, Orth A, Bonhomme F (2008) Phylogeography and postglacial expansion of *Mus musculus* domesticus inferred from mitochondrial DNA coalescent, from Iran to Europe. *Molecular Ecology*, **17**, 627–641.

Rannala B, Mountain J (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9197–9201.

Richardson BA, Brunsfeld SJ, Klopfenstein NB (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology*, **11**, 215–227.

Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution*, **29**, 233–242.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosenberg NA, Ave W, Mi AA (2007) *Distruct: a program for the graphical display of population structure*. Center for Computational Medicine and Biology, Department of Human Genetics, University of Michigan, USA.

Sakaguchi S, Takeuchi Y, Yamasaki M, Sakurai S, Isagi Y (2011) Lineage admixture during postglacial range expansion is responsible for the increased gene diversity of *Kalopanax septemlobus* in a recently colonised territory. *Heredity*, **107**, 338–348.

Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, **7**, 1163–1172.

Shaw J, Lickey E, Schilling E, Small R (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**, 275–288.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, **49**, 369–381.

Stamford MD, Taylor EB (2004) Phylogeographic lineages of Arctic grayling (*Thymallus arcticus*) in North America, divergence, origins and affinities with Eurasian Thymallus. *Molecular Ecology*, **13**, 1533–1549.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, **16**, 608–613.

Svenning J-Ch, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565–573.

Svenning J, Normand S, Skov F (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, **31**, 316–326.

Swofford D (2002) *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sinauer, Sunderland, Massachusetts, USA.

Štorchová H, Hrdličková R, Chrtek J *et al.* (2000) An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon*, **49**, 79–84.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.

Tollefsrud MM, Sønstebø JH, Brochmann C *et al.* (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity*, **102**, 549–562.

Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696–704.

Villesen P (2007) FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes*, **7**, 965–968.

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Willis K, Van Andel T (2004) Trees or no trees? The environments of Central and Eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, **23**, 2369–2387.

Willis KJ, Rudner E, Sümegi P (2000) The full-glacial forests of Central and Southeastern Europe. *Quaternary Research*, **53**, 203–213.

Data Accessibility

All cpDNA sequences were deposited in GenBank under Accession nos. KR085641–KR085769. Files with microsatellite genotypes, input data for Structure, sequence alignment and consensus Bayesian trees are available on DataDryad (doi:10.5061/dryad.g3jc1).

Author Contributions

All authors collected the DNA samples. A.H., K.K. and P.V. performed laboratory analyses. A.H., J.D., K.K. and B.M. realized statistical analyses. A.H., V.H. and K.K. were responsible for figure preparation. A.H. and J.D. wrote the manuscript with contributions from B.M. and K.K. All authors read, edited and approved the manuscript.

Supporting information

Additional supporting information may be found in the online version of this article.

Supplementary information S1: Estimation of DNA ploidy level, genome size and chromosome counts.

Table S1: Locations, haplotype and microsatellite details.

Table S2: Parameter estimates of the most likely scenarios and the bias and precision of the estimations.

Table S3: Model checking of the most likely scenarios.

Figure S1: Estimation of the most probable values of K for all 90 *A. glutinosa* populations.

Figure S2: Histogram of regional genetic structure of all 90 *A. glutinosa* populations as inferred by Structure.

Figure S3: Estimation of the most probable values of K for 68 European *A. glutinosa* diploid populations.

Fig. S4: Histogram of regional genetic structure of 68 European diploid populations of *A. glutinosa* as inferred by Structure.

Fig. S5: Regional genetic structure of 68 European diploid population of *A. glutinosa* as inferred by Structure.

4. PRINCIPAL CONCLUSIONS: Key results about postglacial history of *Alnus glutinosa* and *Alnus incana* in Europe

Thanks to extensive population sampling, testing of hypothesis postulated based on fossil data by molecular data and using two molecular markers with different mode of inheritance and polymorphisms, this project revealed not only the position of glacial refugia of European tree species and discriminate between effective and non-effective ones, but also help infer the main migration routes. This approach enabled us to change some long-lasting paradigms and brought new pieces of knowledge about postglacial colonization of European tree species.

Key results about migration patterns of subgenus Alnus in Europe from palaeoecological records **(Paper I)**

Paloecological records placed ***Alnus*** **LGM refugial areas** to **western Europe, the northern foothills of the Alps, the Romanian Carpathians and a large area of northeastern Europe** where macrofossil remains were found in Poland, Belarus, Lithuania and Latvia. From the southern refugia only **the Pyrenees** were strongly supported.

The expansion of *Alnus* began in the Baltic region and Scandinavia between 11 and 10 cal. kyr BP. The absence of *Alnus* evidence in most of central and northwestern Europe indicates that **populations in northeastern Europe were predominant sources for the colonisation of Scandinavia**. The delayed expansion of *Alnus* in the British Isles between 10 and 8 cal. kyr appears to have originated in a western Europe. Synchronously with the rise of *Alnus* in the British Isles, alders expanded in Hercynian Mountains and subsequently in the Massif Central and the remaining unoccupied area of France between 7 and 6 cal. kyr BP. Source populations are unknown for the *Alnus* expansion in Hercynian Mountains and Massif Central. In northern areas, the *Alnus* expansion shows a scale-dependent pattern characterised by a synchronous increase of *Alnus* within individual regions. In contrary, in southern Europe *Alnus* expanded at individual sites rather than synchronously in whole regions.

The major limit of this study is **inability to precisely distinguish *A. glutinosa* and *A. incana* in palaeoecological records**. Based on palaeoecological data, we cannot choose the true scenarios for individual species, i.e. the possibility that the observed pattern of migration routes was followed by *A. incana* more likely than by *A. glutinosa*

or that the overall picture established on the basis of fossil data is a mixture of postglacial colonization histories of both species. Moreover, due to the scarcity of macrofossils, they do not make the information about position of refugia and migration routes more accurate.

Key results about the distribution and origins of cytotypes of Alnus glutinosa **(Paper III)**

In this study, we for the first time report **the occurrence of tetraploid populations of *A. glutinosa* in Europe** (Fig. 10). The distribution of tetraploids is far from random, as it forms two geographically very well delimited populations. The first is situated in the **Iberian Peninsula, extending to North Africa**. The second one lies in the **Dinaric Alps, extending to south-western Greece**. Both tetraploid populations are probably of **autopolyploid origin** with no indication that *A. incana* has been involved in their evolutionary history.

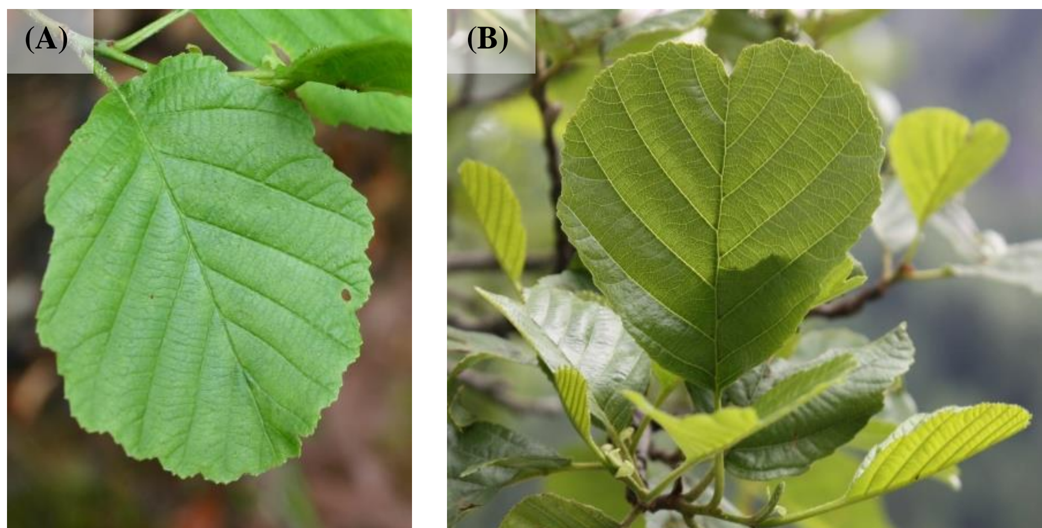


Fig. 10: Leaves of **(A)** tetraploid and **(B)** diploid *Alnus glutinosa*.

Tetraploids were found in two separate populations located in two important glacial refugia, one in the Iberian Peninsula and the second covering the western part of the Balkan Peninsula (Fig. 11). Hence, these areas could not have served as effective refugia for *A. glutinosa* diploids growing in the rest of Europe. Diploid populations located in southern European peninsulas that might be taken into consideration for northward expansion are located only along the border between the Iberian Peninsula and Europe, i.e. in the Pyrenees, in the eastern part of the Balkan Peninsula, in the

Apennine Peninsula or in more northern refugial areas such as the Carpathians, northern foothills of the Alps, western Europe and northeastern Europe.

By combining data from ecological niche models and Bayesian analyses of microsatellite data, it is possible to postulate two contrasting scenarios for each area. While some populations in the Iberian Peninsula were probably stable over a very long period of time, tetraploid populations in the Dinaric Alps withdrew during glacial times and expanded to new suitable localities in interglacial times.

Fig. 11: Canyon of Tara River in Monte Negro where tetraploid populations of *Alnus glutinosa* growing.

Key results about postglacial history of Alnus incana in Europe **(Paper IV)**

The more conservative cpDNA showed a clear pattern of distribution of haplotype diversity of *A. incana* in Europe when **haplotypes concentrated in Southeastern Europe are highly divergent and differ from haplotypes found in the rest of Europe**. Even though *A. incana* is a boreal tree species, it probably took a different route than other boreal trees when it colonized Europe in the postglacial period. Specifically, most European populations of *A. incana* are derived from populations surviving the LPG in southern parts while population from northeastern Europe were not involved.

Fig. 12: Populations of *A. incana* from refugial area in the Carpathians, river Trotus near the town Agăș in Romania.

Based on the combined results of the Structure and ABC analyses applied to microsatellite data, there is evidence that populations currently occurring in the Sudeten Mountains, the westernmost Carpathians and in the eastern Alps represent (Fig. 12 and 13A) **cryptic refugia once situated in lowland regions of Central Europe** and possibly also in microenvironmentally favourable pockets in lower parts of mountain ranges during the LPG. This cryptic **refugium in Central Europe should be considered as the main source area from which *A. incana* colonized Northern Europe** (Fig. 13B), confirming the existence of relatively northern refugia of boreal trees in Europe. By contrast, populations that survived the last glacial period in Southeastern Europe did not spread into Central and Northern Europe.

Fig. 13: Forests with populations of *A. incana* from **(A)** refugial area in the Alps, the Swiss Alps along Le Giffre River near the town Sixt-Fer-à-Cheval in Switzerland and **(B)** area colonized after ice-sheet retreat, above the Skånevik fjord near the town Etne in Norway.

Key results about postglacial history of Alnus glutinosa in Europe **(Paper V)**

Analysis of **cpDNA** distinguished several divergent haplogroups and high haplotype diversity of *A. glutinosa* populations in southern areas point to the existence **refugia in the Iberian, Apennine and Balkan Peninsulas, Corsica, North Africa and the Anatolian Peninsula** (Fig. 14 and 15). Some of these refugial populations located on Iberian and Balkan Peninsula were of tetraploid origin (Fig. 15 B). Moreover, we found specific haplogroup pointing to the existence of **refugial populations in the Carpathians**.

Microsatellites provided a more detailed subdivision of recolonized areas compared to cpDNA and indicated that **all three southern peninsulas served as source refugial areas** for postglacial colonization. The ABC approach, testing several possible scenarios, selected three **the most probable directions of postglacial expansion**: **1)** from the Iberian Peninsula to Western Europe and subsequently to the British Isles, **2)** from the Apennine Peninsula to the Alps and **3)** from the Balkan Peninsula to the Carpathians followed by expansion towards the Northern European plains.

Fig. 14: *A. glutinosa* populations from the refugial populations on the Balkan Peninsula, river Bělica near the town Veliko Tarnovo in Bulgaria.

During the colonization, **two secondary contact zones** were established: (i) Migration routes from Western Europe on the one hand and the Carpathians on the other met in Central Europe, and (ii) Scandinavian populations originated from admixture of Western and Eastern European populations.

Fig. 15: Forests with populations of *A. glutinosa* in refugial areas **(A)** in Parco regionale de la Vena along the Torrente River in Italy where diploids growing and **(B)** along the Coura River near the town Vilares in Portugal where tetraploids growing.

Multiple secondary contact zones between different lineages of *A. glutinosa* colonizing Northern Europe from southern refugial populations appear to be the main

reason for **high genetic diversity of northern populations** as a consequence of **continent wide-admixture** recorded outside of refugial areas of *A. glutinosa*.

New insights into postglacial history of European woody species

Studying two European alder species showed us the differences between postglacial histories of temperate (*A. glutinosa*) and boreal (*A. incana*) tree on one hand and bring new insights into postglacial history of European tree species on the other hand. Our study has brought new evidences about northern cryptic refugia and direction of postglacial migration routes, which have not been traditionally reported for European tree species indicating that their postglacial histories could be more complicated than it has been previously thought.

- **Position of refugia**

For both alder species **highly divergent southern refugia**, similar to that found in many other species (King & Ferris 1998; Petit *et al.* 2002; Palmé & Vendramin 2002; Grivet & Petit 2003; Hampe *et al.* 2003; Heuertz *et al.* 2004, 2006; Cheddadi *et al.* 2006; Magri *et al.* 2006; Liepelt *et al.* 2009; Höhn *et al.* 2009; Cornille *et al.* 2013), were detected. Their distribution was more southern for temperate *A. glutinosa* (i.e. Iberian, Appenine and Balkan Peninsulas, Corsica, North Africa and the Anatolian Peninsula) than for boreal *A. incana* (Balkan Peninsula, the Carpathians and the Alps). The differentiation of individual haplogroups might be very ancient and reflect past biogeographic events (Petit et al. 2003; Médail & Diadema 2009) such as connection of Africa and Europe during the Messinian salinity crisis. (5.3–5.7 Myr BP) (Ketmaier & Caccone 2013). Hence **cpDNA markers generally inform us more about the Pleistocene than the Holocene**.

We found that **some southern populations situated in the putative main glacial refugia are of tetraploid origin**. The distribution of tetraploids forms two geographically very well delimited populations. The first is situated in the Iberian Penisula extending to North Africa, where tetraploids have previously been reported by Lepais *et al.* (2013). The second one lies in the Dinaric Alps, extending to south-western Greece. They are probably autopolyploids with no indication that *A. incana*, the most related species, was involved in their evolutionary history. In our view, the distribution of cytotypes has been formed by past climatic changes during glacial and interglacial times. Nowadays, pure tetraploid populations occur almost exclusively at

the bottoms of deep valleys suggesting that polyploids have a wider spectrum of tolerance and are adapted to ecological conditions not suitable for diploids (Levin 1983).

Moreover, we found indices about **existence of cryptic refugia** for both species. For ***A. glutinosa***, specific haplogroup and specific assignment to microsatellite clusters pointed to the existence of refugial populations in **the Carpathians** which was supported by palaeoecological records. **From genetic data** this is unique **evidence pointing to existence of northern refugia for temperate trees.** Previously only few another examples were proposed, i.e. existence of microrefugia with distinct microsatellite structure in southwestern France for *Fagus sylvatica* (De Lafontaine *et al.* 2013) and Hungarian refugium for *Frangula alnus*, the area with presence of four different cpDNA haplotypes (Hampe *et al.* 2003). However, according to Tzedakis *et al.* (2013), there has been no strong evidence pointing to LPG survival of temperate trees at latitudes higher than 45°, therefore we cannot exclude the possibility that the current northern edge of the Carpathians haplogroup's distribution is the result of postglacial colonization rather than local survival of the glacial period.

For ***A. incana***, the microsatellite data pointed to the refugium in **Central Europe** where populations belong to a genetic cluster that is very common in Northern and Eastern Europe, but very rare in Southern and Southeastern Europe. The same pattern has been found by Tollesfrud *et al.* (2009) for *Picea abies*. Other boreal tree species studied by different molecular markers, namely *Betula pendula* (Palmé *et al.* 2003b; Maliouchenko *et al.* 2007), *Betula pubescens* (Maliouchenko *et al.* 2007), *Pinus sylvestris* (Cheddadi *et al.* 2006; Naydenov *et al.* 2007; Pyhäjärvi *et al.* 2008) and *Salix caprea* (Palmé *et al.* 2003a), showed the presence of genetically rich populations with the lack of a phylogeographical structure in the area north of the Alps supporting the existence of northern refugia which could be fragmented into several isolated populations.

Refugium in northeastern Europe proposed for *Alnus* by palaeoecological records (Huntley & Birks 1983) was not confirmed for *A. glutinosa* by molecular data. For *A. incana*, unique haplotype was detected in this area, but there are several another possibilities about its origin than the existence of northeastern refugium. This area could be colonized from the Carpathians where closely related haplotype occur or from more eastern part of *A. incana* distribution range which was unsampled in our study. We did not confirm the location of more northern refugium in Scandinavia as well.

This refugium was proposed by Parducci *et al.* (2012) who suggested that *Picea abies* might have survived the LPG in northwestern Scandinavia when specific mtDNA haplotype was recorded (Parducci *et al.* 2012) We detected a very similar geographic distribution of one specific cpDNA haplotype of *A. incana*, the only exception being that we have found this haplotype also in the Alps and in Lithuania. We would like to point out the possibility that this pattern determined for *A. incana* might have also been followed by *Picea abies*.

To sum up, apart from traditional southern refugia revealed by cpDNA, microsatellites supported the existence of cryptic refugia in the Carpathians for *A. glutinosa* and in Central Europe for *A. incana*. Moreover, we for the first time reported the occurrence of tetraploid populations of *A. glutinosa* in Europe situated in the putative main glacial refugia.

- **Migration routes**

Migration routes of temperate tree *A. glutinosa* and boreal tree *A. incana* differ significantly. While *A. glutinosa* started its postglacial expansion from all three southern refugia, *A. incana* expanded exclusively from Central Europe.

Our results showed that multiple **refugia served as sources for the postglacial colonization of Europe by *A. glutinosa***. Specifically, one route led from the Iberian Peninsula to Western Europe and subsequently to the British Isles, the second one from the Apennine Peninsula to the Alps and the third one from the Balkam Peninsula to the Carpathians followed by expansion towards the Northern European plains. Colonizing lineages have met several times and formed secondary contact zones in Central Europe and Scandinavia.

Similarly, multiple postglacial migration routes were followed by trees *Fraxinus excelsior* (Heuertz *et al.* 2004) and *Quercus* sp. (Petit *et al.* 2002), and by animals *Mus musculus* (Rajabi-Maham *et al.* 2008) and *Erinaceus* spp. (Santucci *et al.* 1998). It allowed us to challenge the classical paradigm that most European populations of *A. glutinosa* originated from sole refugial areas in the Carpathians (Huntley & Birks 1983; King & Ferris 1998). Postglacial migration by this, so called Balkan route, has been traditionally reported for many other temperate tree species such as *Carpinus betulus*, *Frangula alnus*, *Malus sylvestris* and *Populus nigra* (King & Ferris 1998; Grivet & Petit 2003; Hampe *et al.* 2003; Cottrell *et al.* 2005; Cornille *et al.* 2013) However, it is possible that this conclusion is the result of using less variable cpDNA

and that the migration route from all three southern peninsulas could be common for more trees than have been previously mentioned.

For *A. incana* **Central Europe served as effective refugium for postglacial recolonization of Fennoscandia and Eastern Europe**. Therefore, *A. incana* did not follow the model established for *Picea abies* (Tollefsrud *et al.* 2008), the boreal tree species with very similar current distribution range, and Fennoscandian populations are not derived from Eastern-European ones. The pattern of postglacial colonization by *Pinus sylvestris* based on analysis of mitochondrial DNA (mtDNA) rather resembles the one determined for *A. incana*. It was proposed that populations of *P. sylvestris* in northeastern Russia and Fennoscandia originated from different glacial refugia (Vidyakin *et al.* 2012) and that refugial populations at mid-northern latitudes contributed significantly to the recolonization of northern Europe (Naydenov *et al.* 2007). However, the lack of resolution in mtDNA marker enables several alternative interpretations. It was proposed that Fennoscandia was colonized from the microrefugium in eastern Fennoscandia or adjacent areas of European Russia (Pyhäjärvi *et al.* 2008; Vidyakin *et al.* 2012) but we cannot ruled out that Central Europe or Danube region also participated in northern colonization (Cheddadi *et al.* 2006; Naydenov *et al.* 2007). According to our results southern populations did not served as effective refugia for postglacial expansion of *A. incana* because these areas rather served as source for colonization of Central Europe during the last Weichselian glacial period.

In conclusion, our study of the colonization history of *A. glutinosa* and *A. incana* in Europe reveals new perspectives on the direction of migration of European tree species. Specifically, temperate *A. glutinosa* colonized Europe from multiple southern refugia when boreal *A. incana* expanded from Central Europe towards Fennoscandia and northeastern Europe.

- **Character of migration**

We gained significantly different information about character of migration from individual genetic markers. CpDNA brought only coarse information about character of postglacial migration and the pattern of haplotype distribution is quite similar for both alder species. While haplotypes in Southern Europe are highly divergent, in Central and Northern Europe only one haplotype is widely distributed. We can speculate that this haplotype occurred or evolved in some population situated close to

northern border of the distribution range throughout the last glacial period. In the early Holocene, populations on the edge of the distribution range possibly started to spread north and this haplotype could be part of the expansion wave. It is in concordance with theory that mutations arising in populations at the edge of an expanding range can surf on the wave of advance, and thus reach a larger spatial distribution and a much higher frequency than would be expected in stationary populations (Edmonds *et al.* 2004; Klopfstein *et al.* 2006). Hence, successful "surfing mutations" can reach very high frequencies and eventually occupy large areas.

Microsatellite pattern showed us differences in two alders species pointed to the dissimilarity in character of their migration. Surprisingly, we did not observe typical phenomenon of post-glacial range expansions, i.e. clines of decreasing genetic diversity with increasing distance from effective refugia (Hewitt 1996; Stamford & Taylor 2004; Muller *et al.* 2008). For *A. incana* allelic richness and gene diversity were similar on the way from refugium in Central Europe towards Fennoscandia probably due to colonization in continuously large populations. In contrary, for *A. glutinosa* high genetic diversity of northern populations was detected as a consequence of multiple secondary contact zones between different lineages colonizing Northern Europe from southern refugial populations. Admixture events following formation of secondary contact zones have combined genetic diversity brought by repeated founder events and probably increased success of colonizing populations (Kolbe *et al.* 2004; Keller & Taylor 2010; Rius & Darling 2014). Putative admixture zone in Central Europe, harbouring higher intra-population genetic diversity than in areas further south, have been ascertained for several woody species (Petit et al. 2003). Recently, secondary contact of Eastern and Western migration wave during postglacial colonization of *Taxus baccata* resulted in admixture zone with high genetic diversity in Central Europe was detected by the ABC approach (Mayol *et al.* 2015). On the other hand, high genetic diversity in Scandinavian populations is unique among European temperate tree species and has otherwise been observed only for boreal tree species (Palmé *et al.* 2003a; Tollefsrud *et al.* 2009).

References

Cornille A, Giraud T, Bellard C, Tellier A, Le Cam B, Smulders MJM, Kleinschmit J, Roldan-Ruiz I, Gladieux P (2013) Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Molecular Ecology*, **22**, 2249–2263.

Cottrell JE, Krystufek V, Tabbener HE, Milner AD, Connolly T, Sing L, Fluch S, Burg K, Lefevre F, Achard P, Bordács S, Gebhardt K, Vornam B, Smulders MJM, Vanden Broeck AH, Van SLycken J, Storme V, Boerjan W, Castiglione S, Fossati T, Alba N, Agúndez D, Maestro C, Notivol E, Bovenschen J (2005) Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA. *Forest Ecology and Management*, **206**, 71–90.

Edmonds C, Lillie A, Cavalli-Sforza L (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 975–979.

Grivet D, Petit R (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics*, **4**, 47–56.

Hampe A, Arroyo J, Jordano P, Petit RJ (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, **12**, 3415–3426.

Heuertz M, Carnevale S, Fineschi S Sebastiani F, Hausman JF, Paule L, Vendramin GG (2006) Chloroplast DNA phylogeography of European ashes, *Fraxinus sp.* (Oleaceae): roles of hybridization and life history traits. *Molecular Ecology*, **15**, 2131–2140.

Heuertz M, Fineschi S, Anzidei, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG (2004) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology*, **13**, 3437–3452.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Höhn M, Gugerli F, Abran P, Bisztray G, Buonamici A, Cseke K, Hufnagel L, Quintela-Sabarís C, Sebastiani F, Vedramin GG (2009) Variation in the chloroplast DNA of Swiss stone pine (*Pinus cembra* L.) reflects contrasting post-

glacial history of populations from the Carpathians and the Alps. *Journal of Biogeography*, **36**, 1798–1806.

Hultén E, Fries M (1986) *Atlas of North European vascular plants: north of the Tropic of Cancer I-III*. Koeltz Scientific Books, Königstein.

Huntley B, Birks H (1983) *An atlas of past and present pollen maps for Europe: 0–13000 years ago*. Cambridge University Press., Cambridge.

Cheddadi R, Vendramin GG, Litt T, Francois L, Kageyama M, Lorentz S, Laurent J-M, De Beaulieu J-L, Sadori S, Jost A, Lunt D (2006) Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, **15**, 271–282.

Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, **23**, 1720–1731.

Ketmaier V, Caccone A (2013) Twenty years of molecular biogeography in the West Mediterranean islands of Corsica and Sardinia: lessons learnt and future prospects. In: *Current Progress in Biological Research* (ed Silva-Opps M), pp. 71–93. InTech, Rijeca. Croatia.

King A, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482–490.

Kolbe JJ, Glor RE, Rodríguez Schettino L, Lara Ach, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

De Lafontaine G, Ducousso A, Lefèvre S, Magnanou E, Petit RJ (2013) Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology*, **22**, 4397–4412.

Lepais O, Muller SD, Ben Saad-Limam S, Benslama M, Rhazi L, Belouahem-Abed D, Daoud-Bouattour A, Gammar AM, Ghrabi-Gammar Z, Bacles CFE (2013) High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PloS one*, **8**, e75029.

Levin D (1983) Polyploidy and novelty in flowering plants. *American Naturalist*, **122**, 1–25.

Liepelt S, Cheddadi R, de Beaulieu J-L, Fady B, Gömöry D, Hussendörfer E, Konnert M, Litt T, Longauer R, Terhürne-Berson R, Ziegenhagen B (2009) Postglacial

range expansion and its genetic imprints in *Abies alba* (Mill.) — A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139–149.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latalowa M, Litt T, Paule L, Route JM, Tantau I, Van de Knaap WO, Petit RJ, De Beaulieu J-L (2006) A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *The New Phytologist*, **171**, 199–221.

Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, 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.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.

Mayol M, Riba M, González-Martinéz SC, Bagnoli F, De Beaulieu J-L, Berganzo E, Burgarella C, Dubreuil M, Krajmerová D, Paule L, Romšáková I, Vettori C, Vincenot L, Vendramin GG (in press) Adapting through glacial cycles: insights from a long-lived tree (*Taxus baccata*). *New Phytologist*. doi: 10.1111/nph.13496

Muller M, Leppala J, Savolainen O (2008) Genome-wide effects of postglacil colonization in *Arabidopsis lyrata*. *Heredity*, **100**, 47–58.

Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J (2007) Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology*, **7**, 233.

Palmé AE, Semerikov V, Lascoux M (2003a) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003b) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology*, **12**, 201–212.

Palmé AE, Vendramin GG (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology*, **11**, 1769–1779.

Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y, Edwards ME, Andersen K, Rasmussen M,

Boessenkool S, Coissac E, Brochmann Ch, Taberlet P, Houmark-Nielsen M, Larsen NK, Orlando L, Gilbert MTP, Kjaer KH, Alsos IG, Willerslev E (2012) Glacial survival of boreal trees in northern Scandinavia. *Science*, **335**, 1083–1086.

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

Petit J, Brewer S, Cheddadi R, Coart E, Cottrell J, Csaikl UM, Van Dam B, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JA, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Popescu F, Slade D, Tabbener H, De Vries SGM, Ziegenhagen B, De Beauliue J-L, Kremer A (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Pyhäjärvi T, Salmela MJ, Savolainen O (2008) Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genetics and Genomes*, **4**, 247–254.

Rajabi-Maham H, Orth A, Bonhomme F (2008) Phylogeography and postglacial expansion of *Mus musculus* domesticus inferred from mitochondrial DNA coalescent, from Iran to Europe. *Molecular Ecology*, **17**, 627–641.

Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution*, **29**, 233–242.

Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, **7**, 1163–1172.

Schwabe A (1985) Monographie Alnus incana-reicher Waldgesellschaften in Europa, Variabilität und Ähnlichkeiten einer azonal verbreiteten Gesellschaftsgruppe. *Phytocoenologia*, 13, 197–302.

Stamford M, Taylor E (2004) Phylogeographic lineages of Arctic grayling (*Thymallus arcticus*) in North America, divergence, origins and affinities with Eurasian *Thymallus*. *Molecular Ecology*, **13**, 1533–1549.

Tollefsrud MM, Kissling R, Gugerli F *et al.* (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.

Tollefsrud MM, Kissling R, Gugerli F, Johnsen O, Skroppa T, Cheddadi R, Van der Knaap O, Latalowa M, Terhürne-Berson R, Litt T, Geburek T, Brochmann Ch, Sperisen Ch (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity*, **102**, 549–562.

Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696–704.

Vidiakin AI, Semerikov VL, Polezhaeva MA, Dymshakova OS (2012) Spread of mitochondrial DNA haplotypes in population of scots pine (*Pinus sylvestris* L.) in northern European Russia. *Russian Journal of Genetics*, **48**, 1267–1271.

5. SOUHRN (SUMMARY IN CZECH)

Tento projekt objasnil nejenom, kde byla umístěna glaciální refugia evropských dřevin a která z nich byla efektivní, ale také ukázal, jaké byly hlavní směry postglaciální migrace. Tyto detailní výsledky bylo možné získat díky velkému množství vzorků z celého areálu druhů obou studovaných olší (*Alnus glutinosa* a *A. incana*), testováním hypotéz stanovených na základě fosilních dat pomocí molekulárních analýz a použitím dvou molekulárních markerů lišících se typem přenosu mezi generacemi a jejich proměnlivostí. Tento přístup nám umožnil změnit některé dlouhodobě přetrvávající teorie a přinesl nové poznatky o postglaciální historii evropských dřevin.

Nejdůležitější výsledky o charakteru migrace podrodu Alnus v Evropě na základě paleoekologických dat (Článek I)

Paleoekologická data ukázala, že **glaciální refugia podrodu *Alnus* se nacházela v západní Evropě, severním podhůří Alp, Rumunských Karpatech a v rozlehlých oblastech na severovýchodě Evropy**. Refugia na severovýchodě byla navíc podpořena nálezy makrofosílií v Polsku, Bělorusku, Litvě a Lotyšsku. Z jihoevropských refugií paleoekologická data přesvědčivě potvrdila pouze **refugium na Pyrenejském poloostrově**.

Expanze podrodu *Alnus* v pobaltských zemích a ve Skandinávii začala **před 11 až 10 tisíci lety**. Chybějící důkazy o výskytu olší ve střední a severozápadní Evropě poukazují na to, že hlavní zdrojové populace pro kolonizaci Skandinávie se nacházely v severovýchodní Evropě. Oproti tomu, **západoevropské populace** byly pravděpodobně zdrojem pro opožděnou **expanzi na Britské ostrovy před 10 až 8 tisíci lety.** V tuto dobu olše expandovaly i v **hercynských pohořích a následně ve Francouzském středohoří před 7 a 8 tisíci lety**, přičemž původ zdrojových populací není znám. Zatímco v severských oblastech docházelo k nárůstu populací současně napříč regiony, v jižní Evropě se charakter expanze lišil mezi jednotlivými lokalitami.

Hlavním limitem této studie je nemožnost rozlišit od sebe druhy *A. glutinosa* a *A. incana*. Tudíž nemůžeme na základě paleoekologických dat rozhodnout, kterému z těchto druhů náleží nalezené migrační cesty. Makrofosilní nálezy, které by tento problém mohly vyřešit, jsou bohužel velmi vzácné.

Nejdůležitější výsledky o rozšíření a původu cytotypů druhu Alnus glutinosa **(Článek III)**

V této studii jsme poprvé zaznamenali **výskyt tetraploidních populací druhu *A. glutinosa* v Evropě**. Tyto tetraploidní populace rostou ve dvou geograficky dobře vymezených územích. První z oblastí se nachází **na Iberském poloostrově a severu Afriky**. Druhá oblast se rozkládá **v Dinárských Alpách a na jihozápadě Řecka**. Všechny tetraploidní populace jsou pravděpodobně **autopolyploidního původu** a nevykazují známky toho, že by blízce příbuzný druh *A. incana* hrál roli v jejich evoluční historii.

Výskyt tetraploidních jedinců je vázán na oblasti **dvou důležitých glaciálních refugií** na Iberském a Balkánském poloostrově. Z tohoto důvodu tyto oblasti nemohly sloužit jako efektivní refugia, protože v severnějších částech Evropy rostou pouze diploidní populace. Zdrojové populace pro postglaciální kolonizaci se tedy mohly nacházet v oblastech s výskytem diploidních populací jako jsou Pyreneje, východ Balkánského poloostrova, Apeniny nebo v severnějších oblastech jako jsou Karpaty, severní podhůří Alp, západní nebo severovýchodní Evropa.

Výsledky modelování ekologických nik a Bayesiánských analýz mikrosatelitů ukazují, že **oblasti s výskytem tetraploidních populací se liší svojí historií**. Zatímco populace na Iberském poloostrově pravděpodobně dlouhodobě přežívaly na stejných lokalitách, Balkánské populace jsou spíše výsledkem postglaciální migrace.

Nejdůležitější výsledky o postglaciální historii druhu Alnus incana v Evropě **(Článek IV)**

Konzervativnější chloroplastová DNA ukázala, že **haplotypy v jihovýchodní Evropě jsou vzájemně rozrůzněné** a liší se od haplotypů ve zbývajících částech Evropy. Postglaciální kolonizace boreálního druhu *A. incana* se pravděpodobně lišila od scénářů navrhnutých pro jiné boreální dřeviny. Většina evropských populací druhu *A. incana* byla totiž **kolonizována spíše z jižněji položených oblastí** než ze severovýchodní Evropy.

Na základě analýzy mikrosatelitů usuzujeme, že **populace v Sudetských pohořích, na západě Karpat a východě Alp pocházejí z kryptického refugia**, které se nacházelo ve středoevropských nížinách a pravděpodobně také v nižších polohách horských oblastí s příznivým mikroklimatem. Kryptické refugium ve střední Evropě bylo zdrojem pro postglaciální kolonizaci severní Evropy a je důkazem, že refugia boreálních dřevin se vyskytovala i v severněji položených oblastech Evropy. Oproti tomu **populace, které přežívaly poslední glaciál v jihovýchodní Evropě, nepřispěly ke kolonizaci střední a severní Evropy**.

Nejdůležitější výsledky o postglaciální historii druhu Alnus glutinosa v Evropě (Článek V)

Chloroplastová DNA rozlišila několik vysoce rozrůzněných haploskupin a vysokou haplotypovou diverzitu v jižních oblastech. To poukazuje na **výskyt refugií na Iberském, Apeninském a Balkánském poloostrově, na Korsice, v severní Africe a na Anatolském poloostrově**. Kromě toho jsme objevili důkaz pro existenci **severněji položeného refugia v Karpatech**, kde jsme objevili unikátní haploskupinu.

Analýza mikrosatelitů nám poskytla **detailnější informace o průběhu postglaciální kolonizace** než chloroplastová DNA. Ukázalo se, že **všechna tři jihoevropská refugia přispěla k postglaciální kolonizaci Evropy**. ABC analýza, která proti sobě testovala více možných scénářů, vybrala **nejpravděpodobnější cesty postglaciální kolonizace**: **1)** z Iberského poloostrova do střední a západní Evropy a následně na Britské ostrovy, **2)** z Apeninského poloostrova do Alp a **3)** z Balkánského poloostrova do Karpat a následně do severní Evropy. V průběhu kolonizace vznikly **dvě sekundární kontaktní zóny**: **(i)** migrační proudy ze západní Evropy a z Karpat se setkaly ve střední Evropě a **(ii)** skandinávské populace jsou tvořeny západoevropskými a východoevropskými populacemi. Setkávání různých linií v průběhu postglaciální kolonizace se zdá být hlavním důvodem vysoké genetické diverzity i mimo refugiální oblasti.

6. SUPPORTING INFORMATION

1. Supporting information to the paper Douda J, Doudová J, Drašnarová A, Kuneš P, Hadincová V, Krak K, Zákravský P, Mandák B (2014) **Migration patterns of subgenus *Alnus* in Europe since the Last Glacial Maximum: a systematic review**. *Plos one*, **9**, e88709.

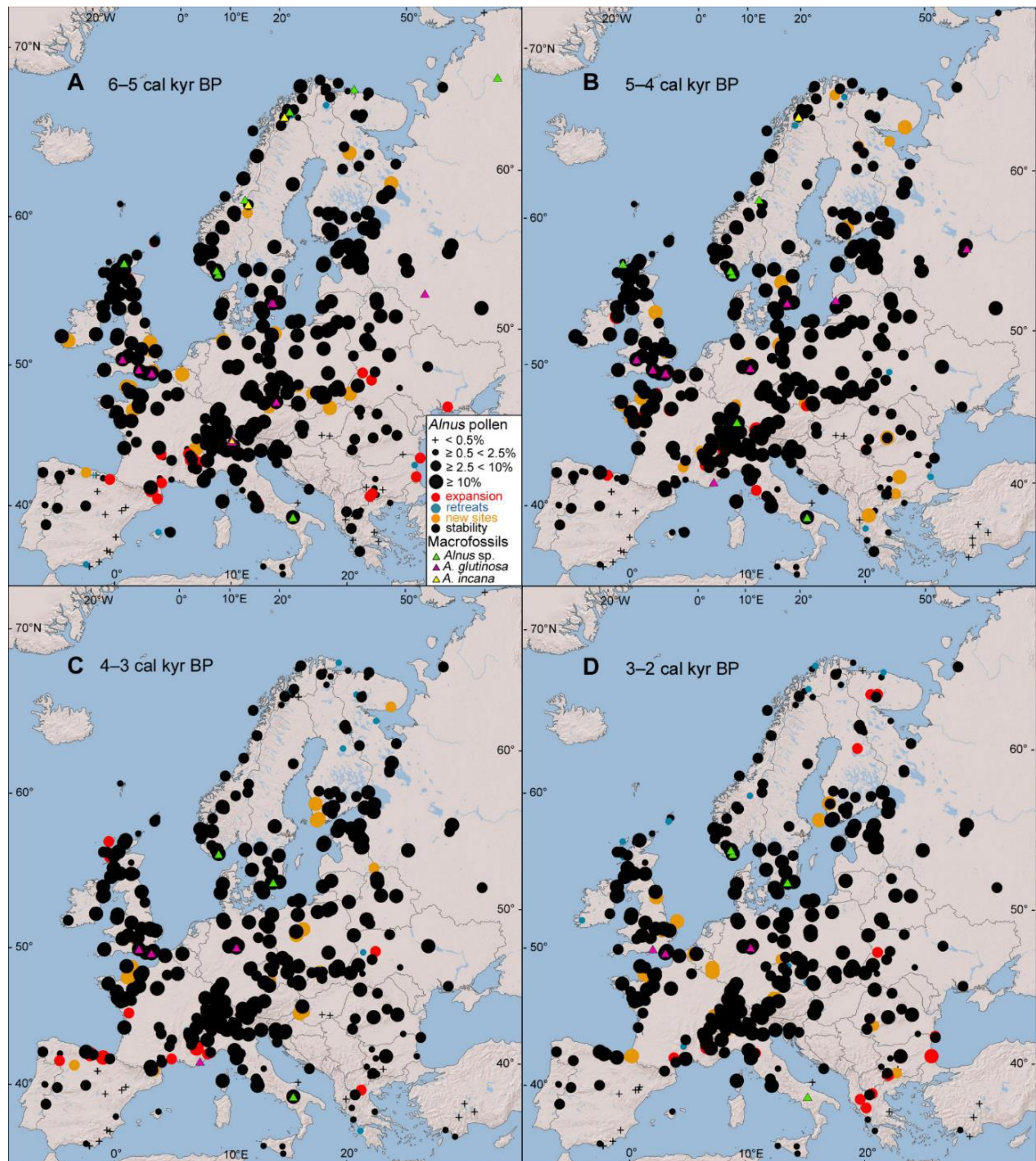


Figure S1. Holocene distribution (6–2 cal. kyr BP) of *Alnus* pollen sites. According to four classes of percentage of *Alnus* pollen and macrofossil remains. The colour of dots indicates changes compared to the previous period; red, expansion, *Alnus* pollen < 2.5% in preceding period; blue, retreat, *Alnus* pollen ≥ 2.5% in preceding period; orange, new pollen sites of *Alnus* pollen ≥ 2.5%, respectively; black, stability; the course of deglaciation (white) and changes in coastline (dot lines).

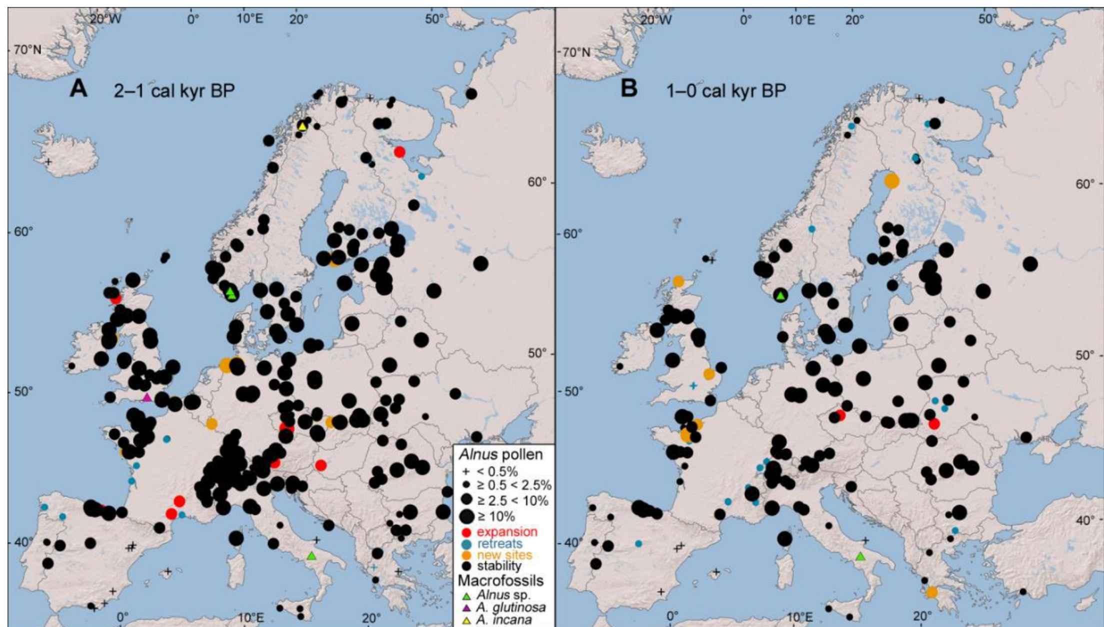


Figure S2. Holocene distribution (2–0 cal. kyr BP) of *Alnus* pollen sites. According to four classes of percentage of *Alnus* pollen and macrofossil remains; for details see Appendix S3 Figure S1.

Table S1, S2 and S3 available online:

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0088709#s5

Table S1. Location of the pollen sites from EPD, PALYCZ and the literature (Lit.).

Table S2. Location of the macrofossil sites from NEMD and the literature (Lit.).

Table S3. References of the pollen and macrofossil sites from EPD, PALYCZ, NEMD and the literature (Lit.).

2. Supporting information to the paper Drašnarová, A., Krak, K., Vít, P., Doudová, J., Douda, J., Hadincová, V., Zákravský, P. & Mandák, B. (2014) **Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana***. *Tree Genetics & Genomes*, 10, 865-873.

Online Resource 1: Characteristics of 39 microsatellite loci from Betulaceae used for cross-amplification on *A. glutinosa* and *A. incana* sorted according to success or failure during optimization (included in final multiplexes or excluded from them)

No	Loci	GenBank no.	Repeats	Species	Size range (bp)	Source
A. Loci included in final multiplexes						
A2	Bp04	AB084474	$(GT)_{12}$... $(GA)_{5}$	BETPLA	151–173	Wu et al. 2002
A6	Bmax624	AB094100	$(TC)_{14}$	BETMAX	202–216	Ogyu et al. 2003
A7	Bmax097	AB094104	$(CT)_{13}$	BETMAX	190–200	Ogyu et al. 2003
A10	L5.5	AF310863	$C_{12}CTCC(CT)_{7}TT(CT)_{5}$	BETPEN	121–146	Kulju et al. 2004
A18	B634	FJ986496	$(AG)_{15}$	CORAVE	218–238	Gürcan and Mehlenbacher 2010
A22	B720	FJ986523	$(AG)_{14}$	CORAVE	159–179	Gürcan et al. 2010
A26	Alma11	n/a	$(CT)_{11}$	ALNMAR	364–374	Lance et al. 2009
A35	n/a	CD276907	$(TC)_{8}$	BETMAX	217–227	Tsuda et al. 2009a
A37	n/a	CD277113	$(TC)_{9}$	BETMAX	272–278	Tsuda et al. 2009b
A38	n/a	CD278280	$(CAA)_{5}$	BETMAX	101–116	Tsuda et al. 2009a
B. Loci excluded because of low quality profiles in multiplexes						
A3	Bp07	AB084475	$(GT)_{10}(GA)_{12}$	BETPLA	135–150	Wu et al. 2002
A4	Bp11	AB084478	$(GT)_{8}$...$(AT)_{4}$	BETPLA	113–119	Wu et al. 2002
A8	L7.1	AF310854	$(CT)_{12}CCTT(CT)_{4}$	BETPEN	146–152	Kulju et al. 2004
A14	AGAG164a	Y08436	$(TA)_{17}$	ALNGLU	215	Gürcan and Mehlenbacher 2010

A16	B617	FJ986490	$(GA)_{15}$	CORAVE	280–298	Gürcan et al. 2010
A17	B619	FJ986491	$(TC)_{21}$	CORAVE	146–180	Gürcan et al. 2010
A19	B664	FJ986510	$(TC)_{21}$	CORAVE	186–216	Gürcan et al. 2010
A20	B702a	FJ986517	$(CT)_{13}CG(CT)_{3}Ns$	CORAVE	280–305	Gürcan et al. 2010
A21	B709	FJ986519	$(GA)_{21}$	CORAVE	219–233	Gürcan et al. 2010
A24	Alma20	n/a	$(CT)_{11}$	ALNMAR	208–230	Lance et al. 2009
A27	Alma5	n/a	$(ATGT)_{9}$	ALNMAR	349–381	Lance et al. 2009
A34	n/a	CD278264	$(GAT)_{7}$	BETMAX	72–75	Tsuda et al. 2009a
A36	n/a	CD277230	$(CCT)_{3}CAG(CTC)_{3}$	BETMAX	388–390	Tsuda et al. 2009a
C. Loci excluded because of low quality profiles in cross-amplification						
A1	Bp01	AB084473	$(GT)_{10}$	BETPLA	152–160	Wu et al. 2002
A5	Bp16	AB084483	$(GT)_{9}...A_{14}$	BETPLA	166–172	Wu et al. 2002
A9	L1.10	AF310856	$(GA)_{4}AA(GA)_{10}$	BETPEN	168–209	Kulju et al. 2004
A11	n/a	AJ490266	$(CT)_{13}Ns(CT)_{13}$	BETPEN	210–224	Gürcan and Mehlenbacher 2010
A12	Bo.F330	AY423611	$(TC)_{14}$	PETPUB	172–210	Truong et al. 2005
A13	n/a	Z72433	$(CA)_{17}T(AT)_{16}$	BETPEN	169–187	Gürcan and Mehlenbacher 2010
A15	AGAG164b	Y08436	$(TA)_{29}$	ALNGLU	168	Gürcan and Mehlenbacher 2010
A23	B793	FJ986555	$(TC)_{16}$	CORAVE	160–176	Gürcan et al. 2010
A25	Alma25	n/a	$(ATGT)_{6}...(GTTT)_{6}...(GTTT)_{5}$ $...(GTTT)_{5}...(GTTT)_{5}$	ALNMAR	307–399	Lance et al. 2009
A28	L7.8	AF310866	$(CT)_{11}GC(AATG)_{2}$	BETPEN	295–307	Kulju et al. 2004
A29	L2.7	AF310850	$(TC)_{8}(TA)_{8}(TG)_{11}TT(TG)_{3}$	BETPEN	141–186	Kulju et al. 2004

A30	L3.4	AF310852	(GTAT)$_3$(GT)$_5$	BETPEN	258–274	Kulju et al. 2004
A32	L5.4	AF310862	(TC)$_{26}$	BETPEN	230–262	Kulju et al. 2004
A33	L022	AF310874	(CT)$_{18}$	BETPEN	172–196	Kulju et al. 2004
A39	L3.1	AF310851	(CT)$_3$CC(CT)$_2$CC(CT)$_{13}$AT(CT)$_5$	BETPEN	219–241	Kulju et al. 2004
A40	L2.3	AF310847	(AG)$_{16}$	BETPEN	198–220	Kulju et al. 2004

References

Gürcan K, Mehlenbacher S (2010) Transferability of microsatellite markers in the Betulaceae.J Am Soc Hortic Sci 135:159–173

Gürcan K, Mehlenbacher SA, Botta R, Boccacci P (2010) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. Tree Genet Genomes 6:513–531. doi: 10.1007/s11295-010-0269-y

Kulju KKM, Pekkinen M, Varvio S (2004) Twenty-three microsatellite primer pairs for *Betula pendula* (Betulaceae). Mol Ecol Notes 4:471–473. doi: 10.1111/j.1471-8286.2004.00704.x

Lance SL, Jones KL, Hagen C, et al. (2009) Development and characterization of nineteen polymorphic microsatellite loci from seaside alder, *Alnus maritima*. Conserv Genet 10:1907–1910. doi: 10.1007/s10592-009-9851-y

Ogyu K, Tsuda Y, Sugaya T, et al. (2003) Identification and characterization of microsatellite loci in *Betula maximowicziana* Regel. Mol Ecol Notes 3:268–269. doi: 10.1046/j.1471-8286.2003.00419.x

Truong C, Palme A, Felber F, Naciri-Graven Y (2005) Isolation and characterization of microsatellite markers in the tetraploid birch, Bet*ula pubescens* ssp. *tortuosa*. Mol Ecol Notes 5:96–98. doi: 10.1111/j.1471-8286.2004.00848.x

Tsuda Y, Ueno S, Ide Y, Tsumura Y (2009a) Development of 14 EST-SSRs for *Betula maximowicziana* and their applicability to related species. Conserv Genet 10:661–664. doi: 10.1007/s10592-008-9608-z

Tsuda Y, Ueno S, Ranta J, et al. (2009b) Development of 11 EST-SSRs for Japanese white birch, *Betula platyphylla* var. *japonica* and their transferability to related species. Conserv Genet 10:1385–1388. doi: 10.1007/s10592-008-9

3. Supporting information to the paper Mandák, B., Vít, P., Krak, K., Trávníček, P., Havrdová, A., Hadincová, V., Zákravský, P., Jarolímová, V., Bacles, C.F.E. & Douda, J. (in press) **Putative glacial refugia inferred from the geographic distribution of *Alnus glutinosa* cytotypes in Europe.** *Annals of Botany.*

Fig. S1 Estimation of the most probable value of *K* for whole dataset containing di-, tri- and tetraploid *A. glutinosa* individuals and *A. incana*. Average similarity coefficients for each K with standard deviations according to Nordborg *et al.* (2005) and Ehrich *et al.* (2007).

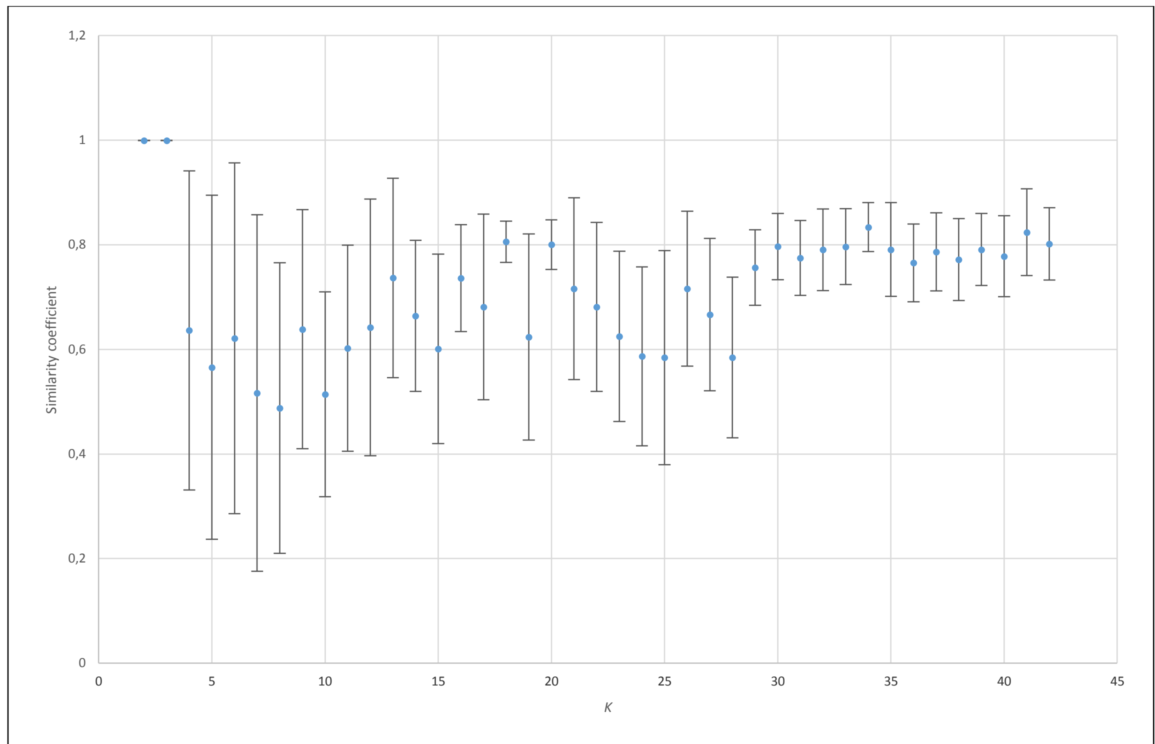


Fig. S2 Estimation of the most probable value of *K* for di-, tri- and tetraploid *A. glutinosa* individuals. Average similarity coefficients for each K with standard deviations according to Nordborg *et al.* (2005) and Ehrich *et al.* (2007).

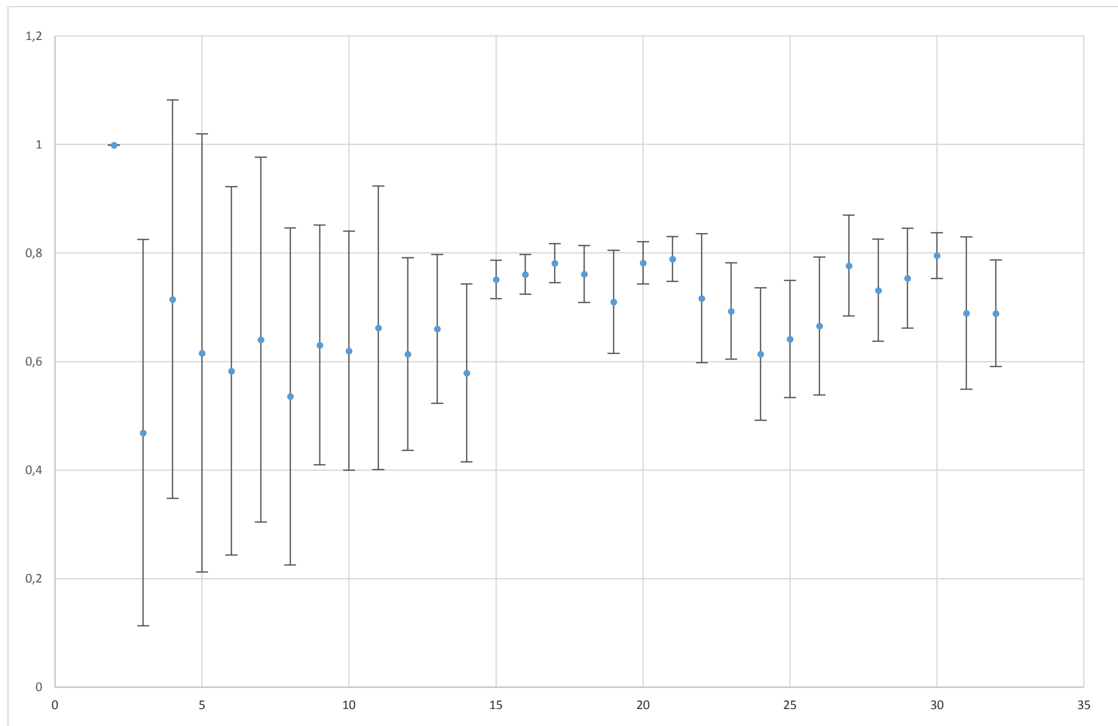
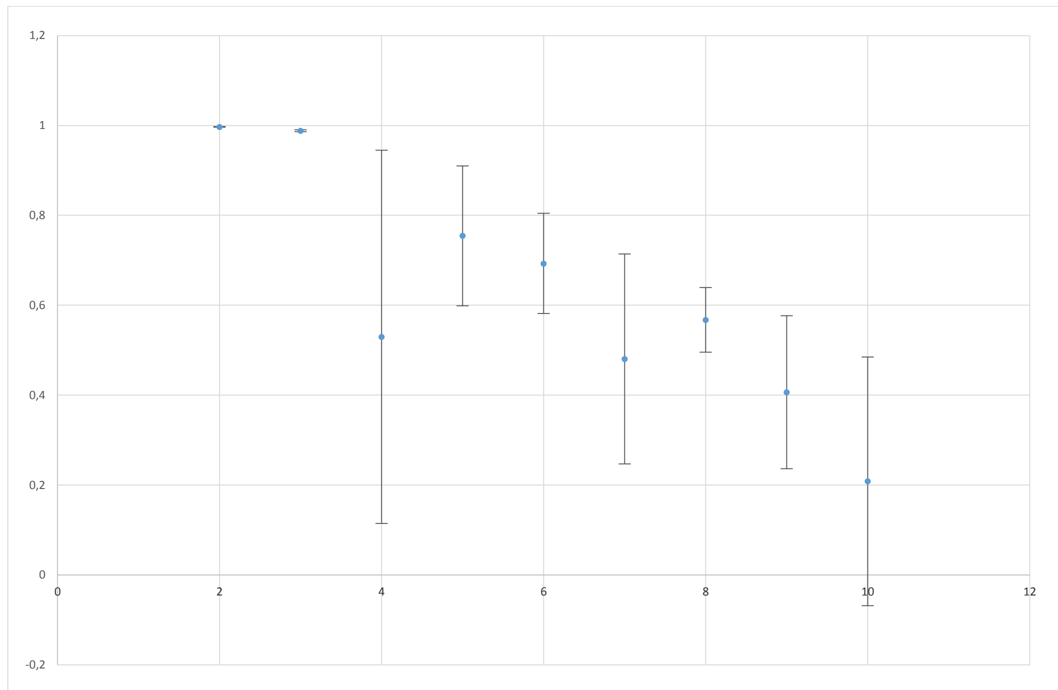


Fig. S2 Estimation of the most probable value of *K* for subset of di-, tri- and tetraploid *A. glutinosa* individuals from Balkan and Austria. Average similarity coefficients for each K with standard deviations according to Nordborg *et al.* (2005) and Ehrich *et al.* (2007).



LITERATURE CITED

Ehrich D, Gaudeul M, Assefa A, *et al.* 2007. Genetic consequences of Pleistocene range shifts: Contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* **16:** 2542–2559.

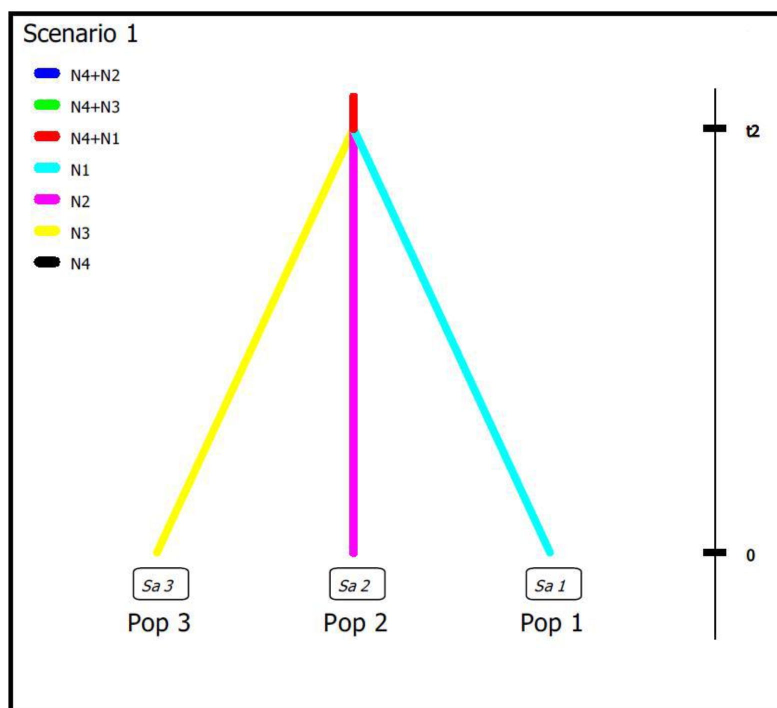
Nordborg M, Hu TT, Ishino Y, *et al.* 2005. The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol* **3:** e196. 10.1371/journal.pbio.0030196.

4. Supporting information to the paper Mandák B, Havrdová A, Krak K, Hadincová V, Vít P, Zákravský P, Douda J (submitted, New Phytologist) **Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation.**

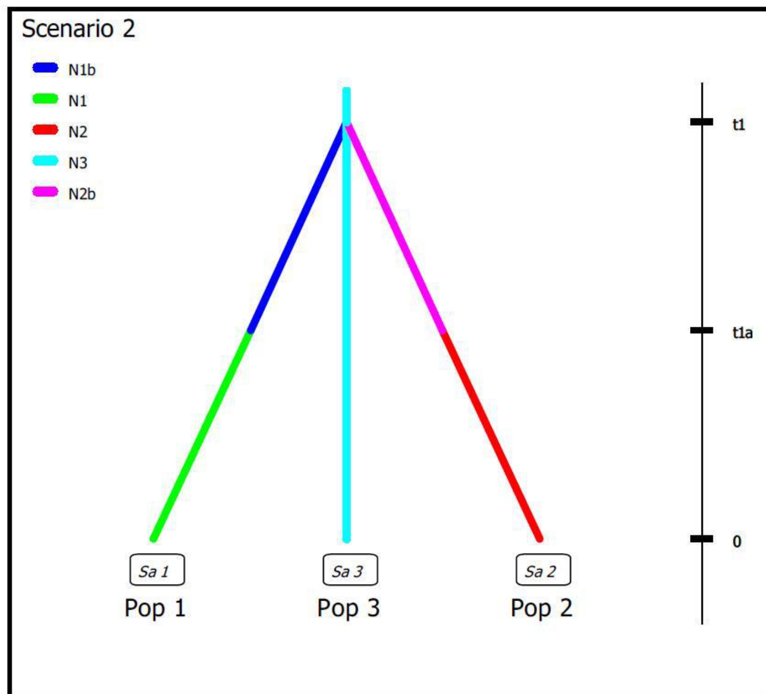
Fig. S1 Graphical representation of the 21 competing scenarios used in the ABC analyses. The origins of populations in the Alps and the Carpathians, Central Europe, Eastern Europe and Fennoscandia were investigated separately. Balkan region (BA), Carpathians (CA), Alps (AL), Central Europe (CE), Fennoscandia (SC) and Baltic and northeastern Europe (EE) (see Table 1 and Fig. 1).

1. Alps and Carpathians (Pop 1 = AL, Pop 2 = CA, Pop 3 = BA)

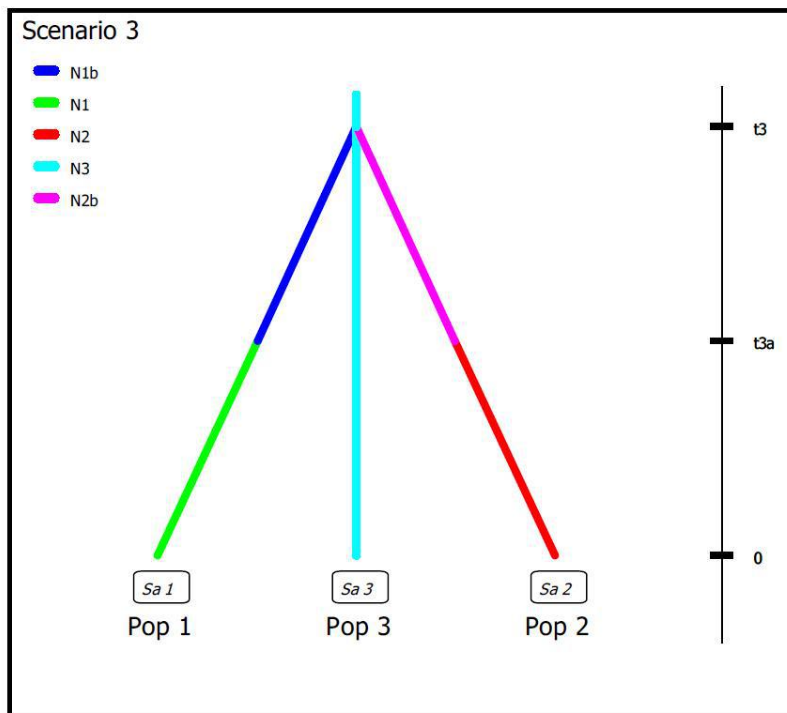
1.1. AL, CA, BA refugium



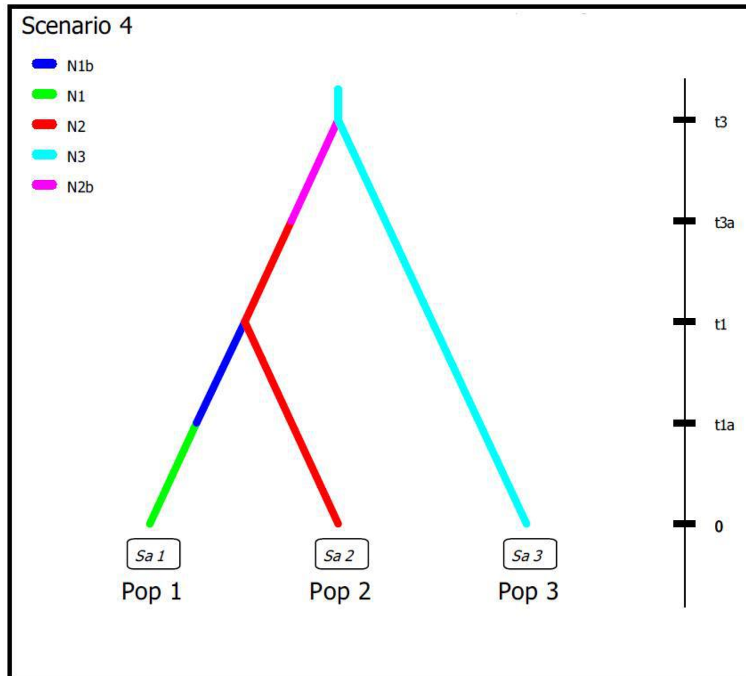
1.2. From BA (Holocene); BA refugium



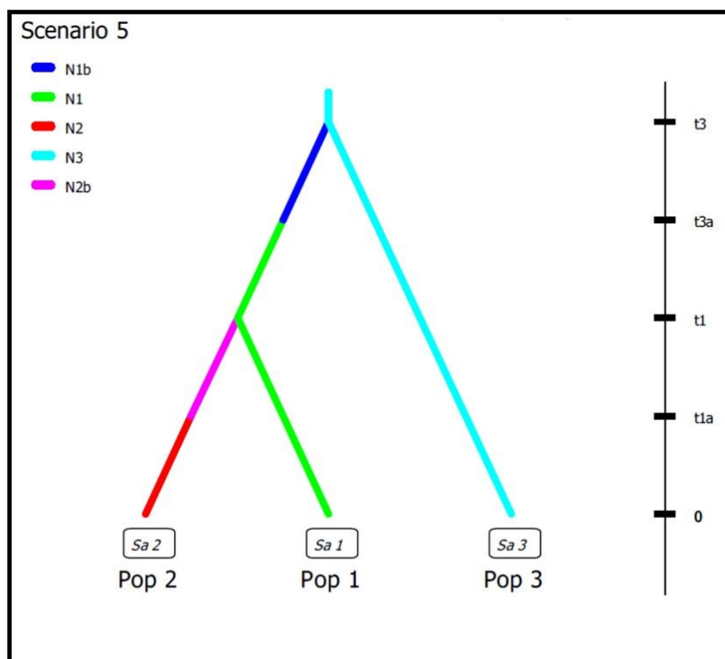
1.3. From BA (Pleistocene); BA refugium; Colonization of AL and CA from BA during Pleistocene (>100,000 generations)



1.4. From BA (Pleistocene); BA refugium; Colonization of CA from BA refugium during Pleistocene (>100,000 generations), colonization of AL from CA during Holocene

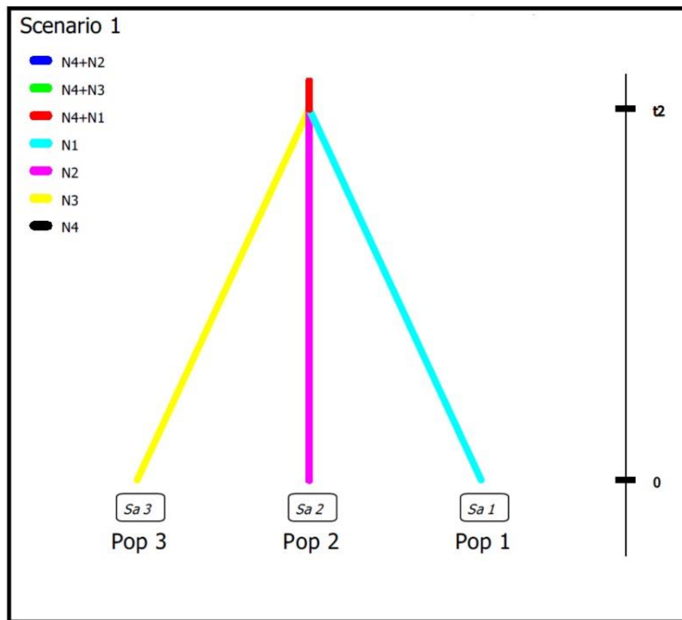


1.5. From BA (Pleistocene); BA refugium; Colonization of AL from BA refugium during Pleistocene (>100,000 generations), colonization of CA from AL during Holocene

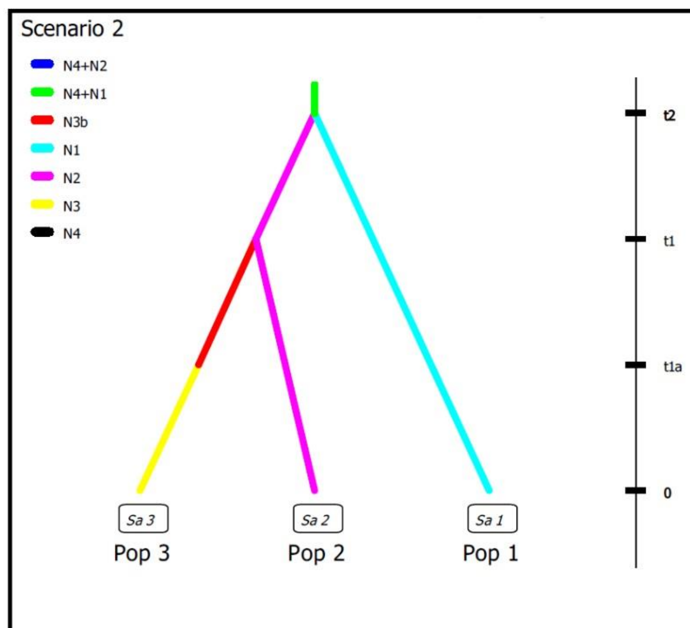


2. Central Europe (Pop 1 = AL, Pop 2 = CA, Pop 3 = CE)

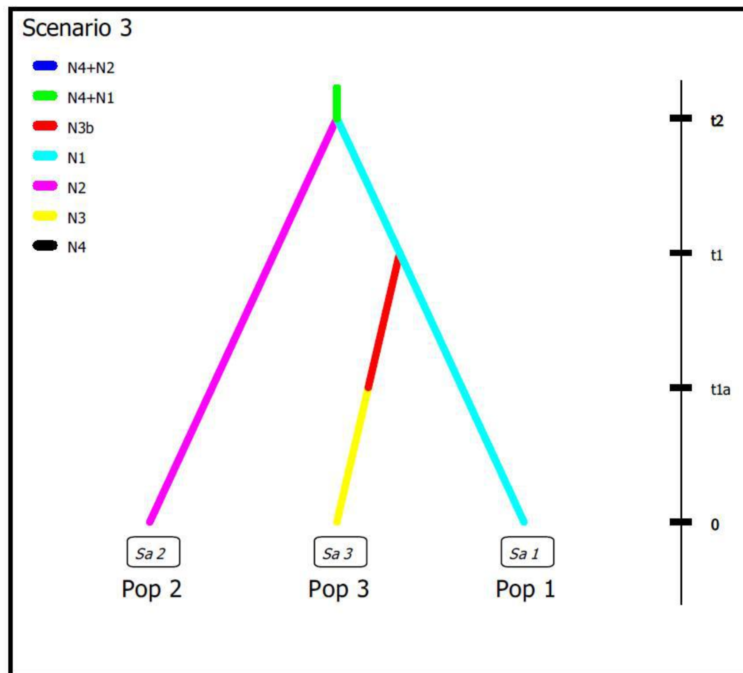
2.1. AL, CA, CE refugium



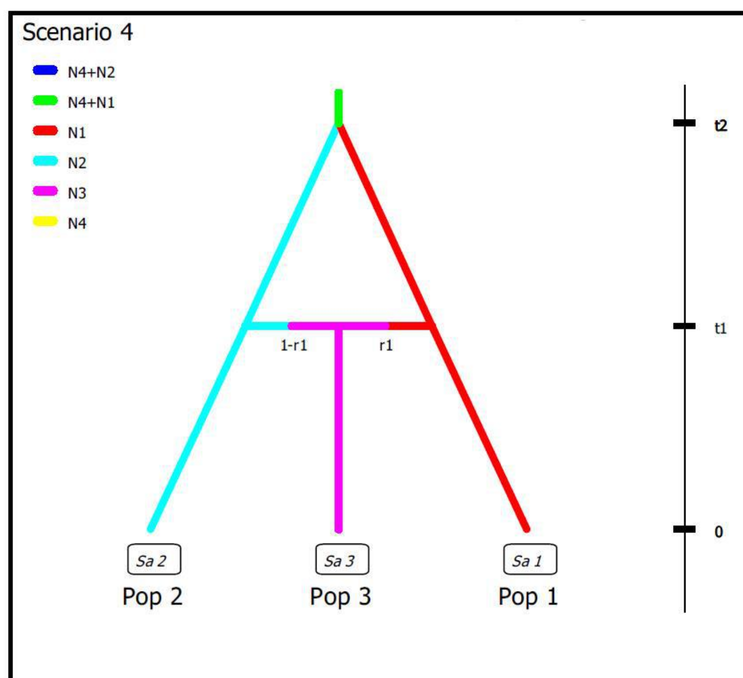
2.2. From CA; CA and AL refugium



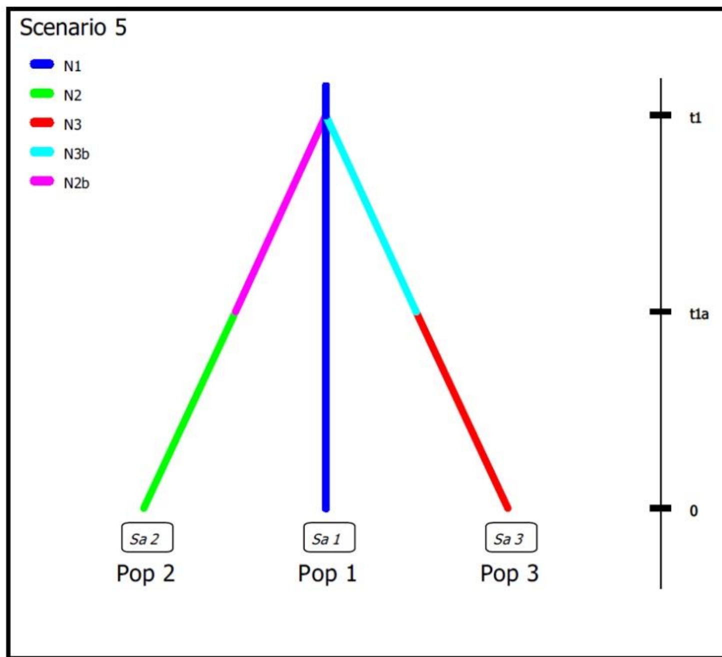
2.3. From AL; AL and CA refugium



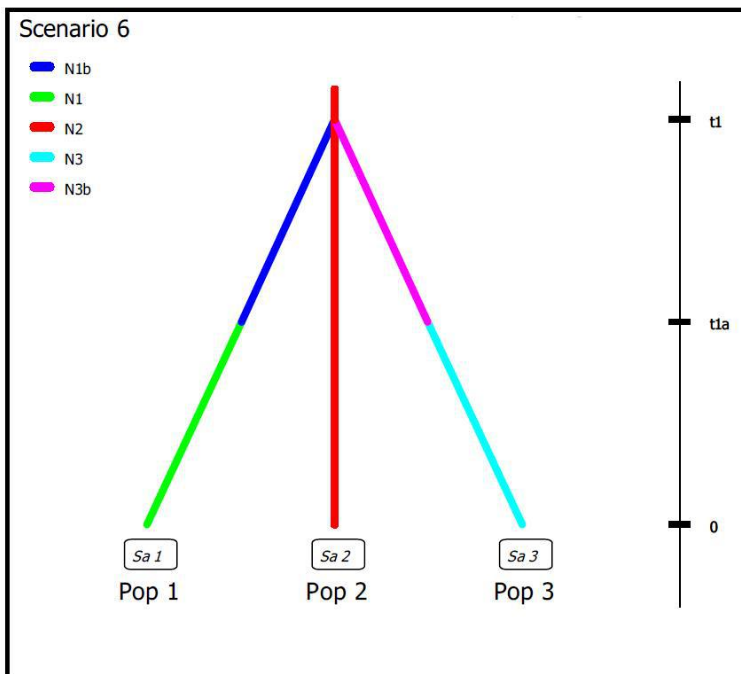
2.4. Admixture of AL and CA



2.5. From AL; only AL refugium

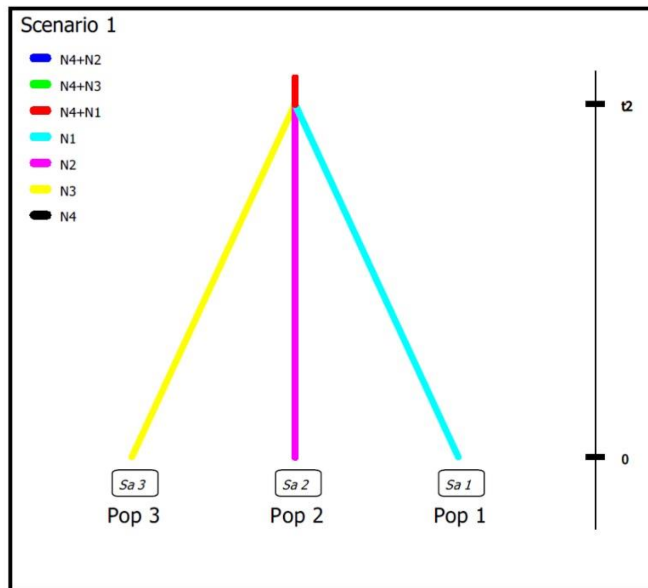


2.6. From CA; only CA refugium

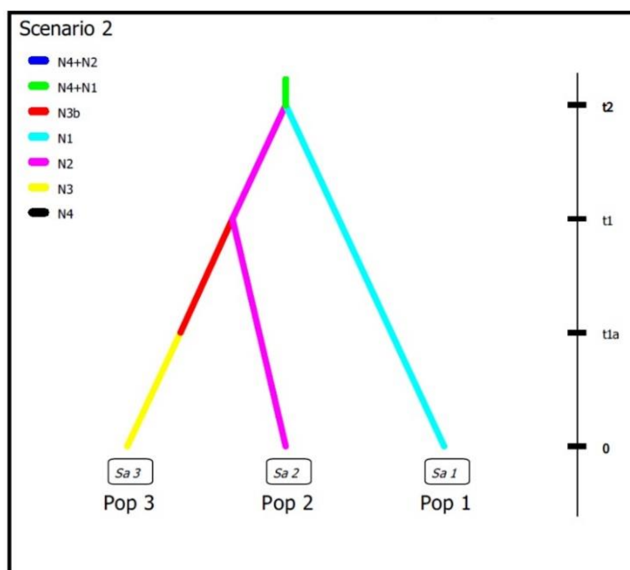


3. Eeastern Europe (Pop 1 = CA, Pop 2 = CE, Pop 3 = EE)

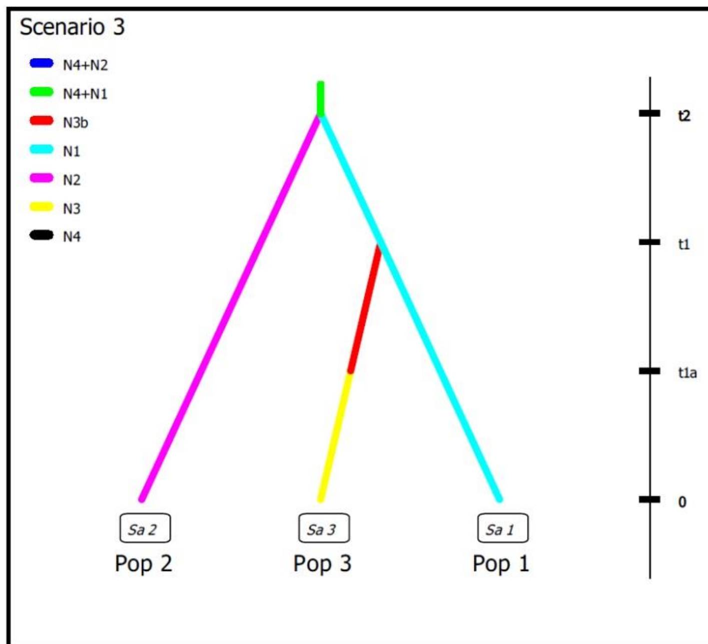
3.1. EE, CA, CE refugium



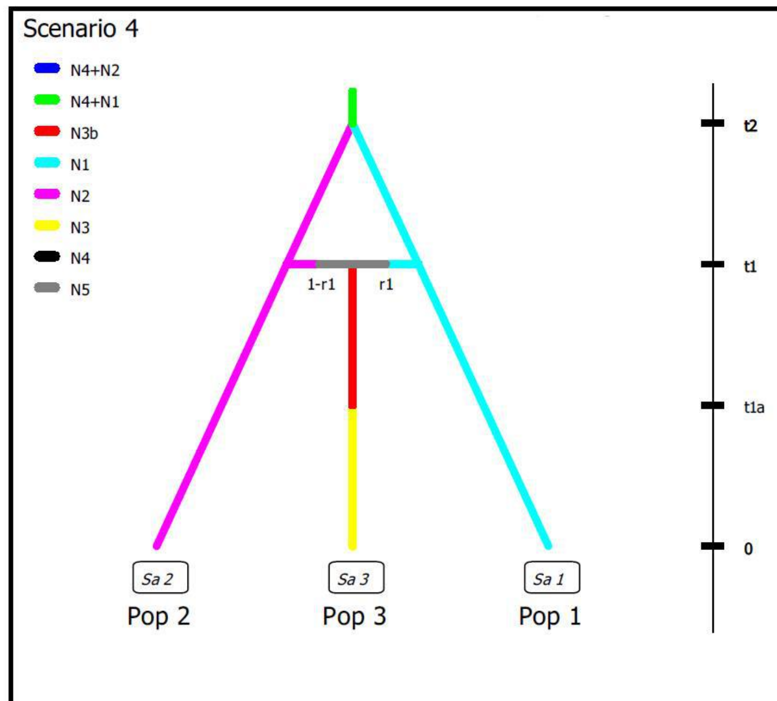
3.2. From CE; CA and CE refugium



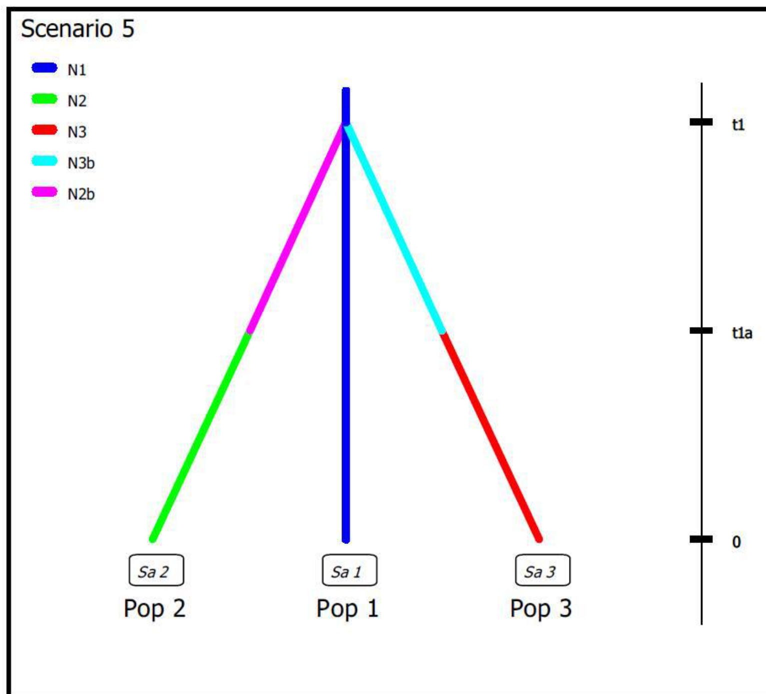
3.3. From CA; CA and CE refugium



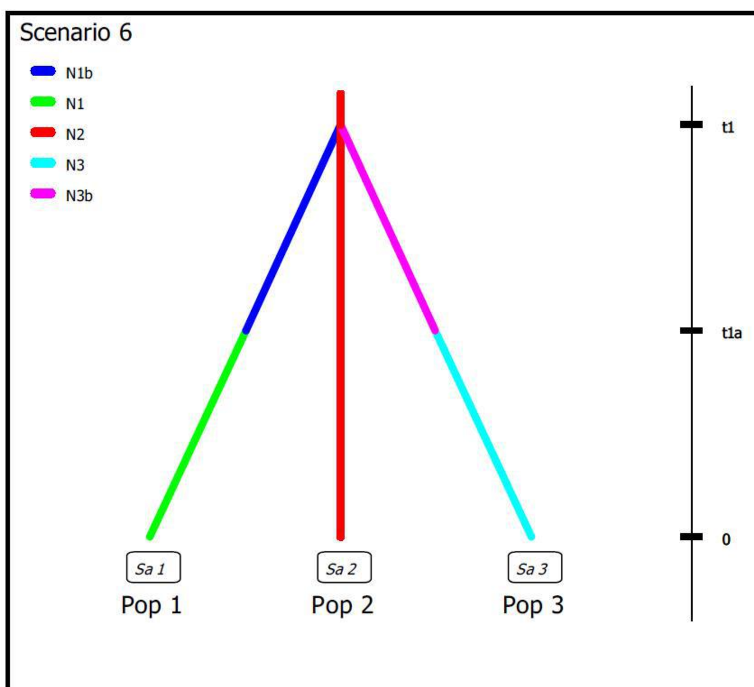
3.4. Admixture of CE and CA



3.5. From CA; only CA refugium

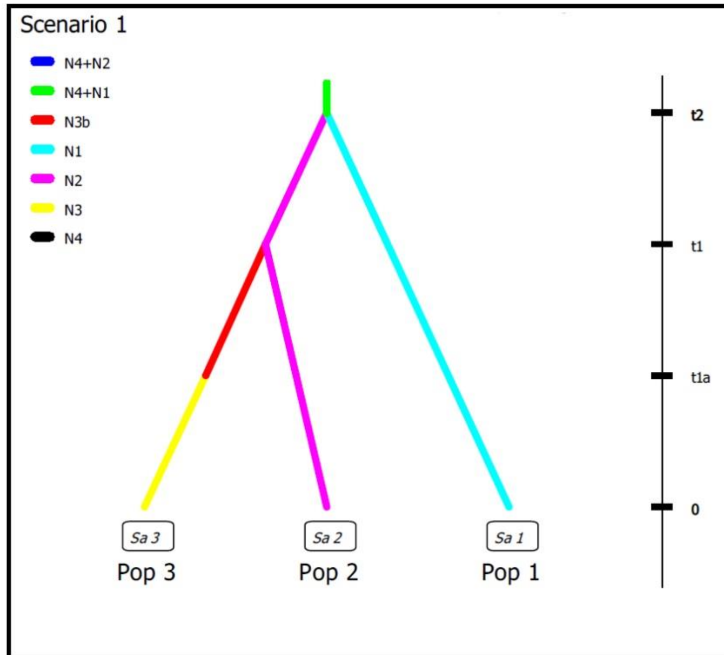


3.6. From CE; only CE refugium

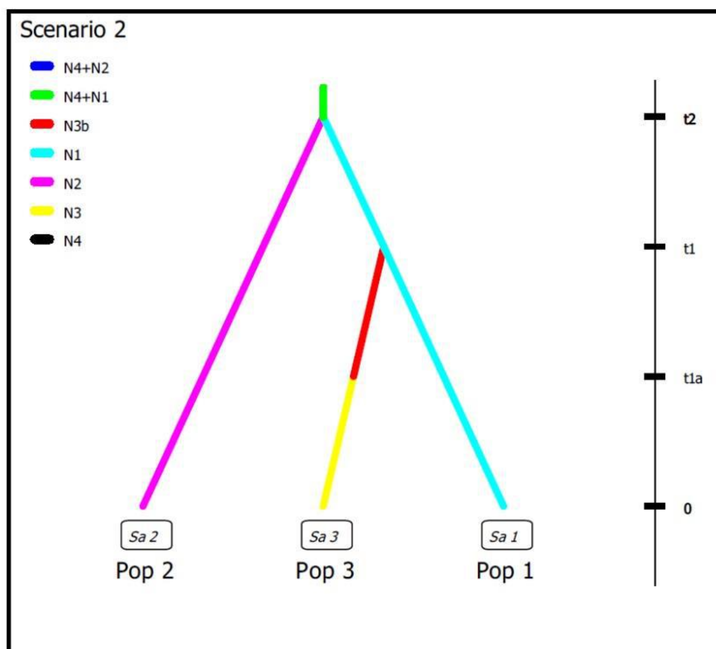


4. Fennoscandia (Pop 1 = CE, Pop 2 = EE, Pop 3 = SC)

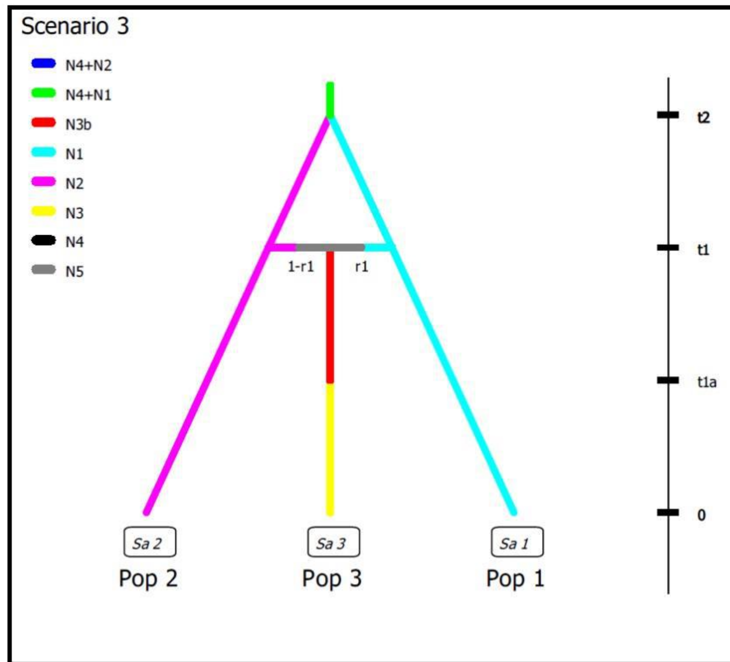
4.1. From EE; EE and CE refugium



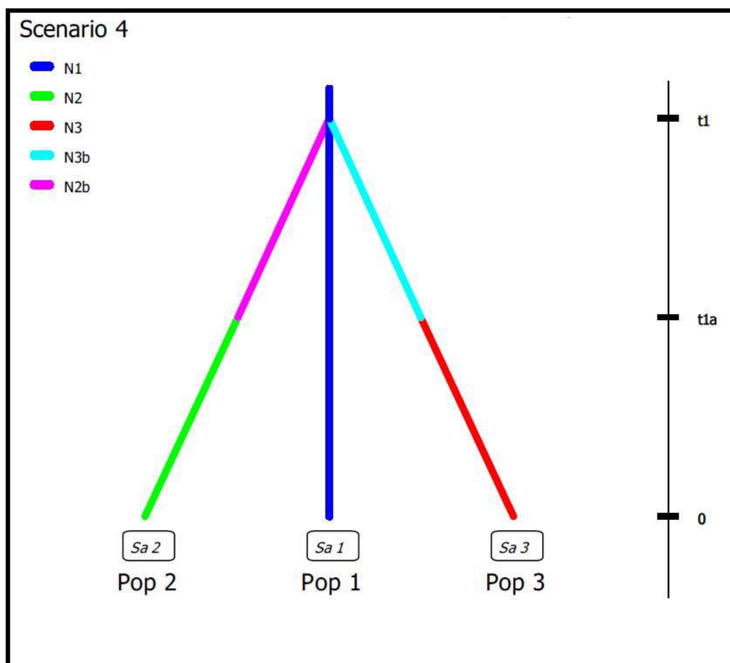
4.2. From CE; EE and CE refugium



4.3. Admixture of CE and EE



4.4. From CE; only CE refugium



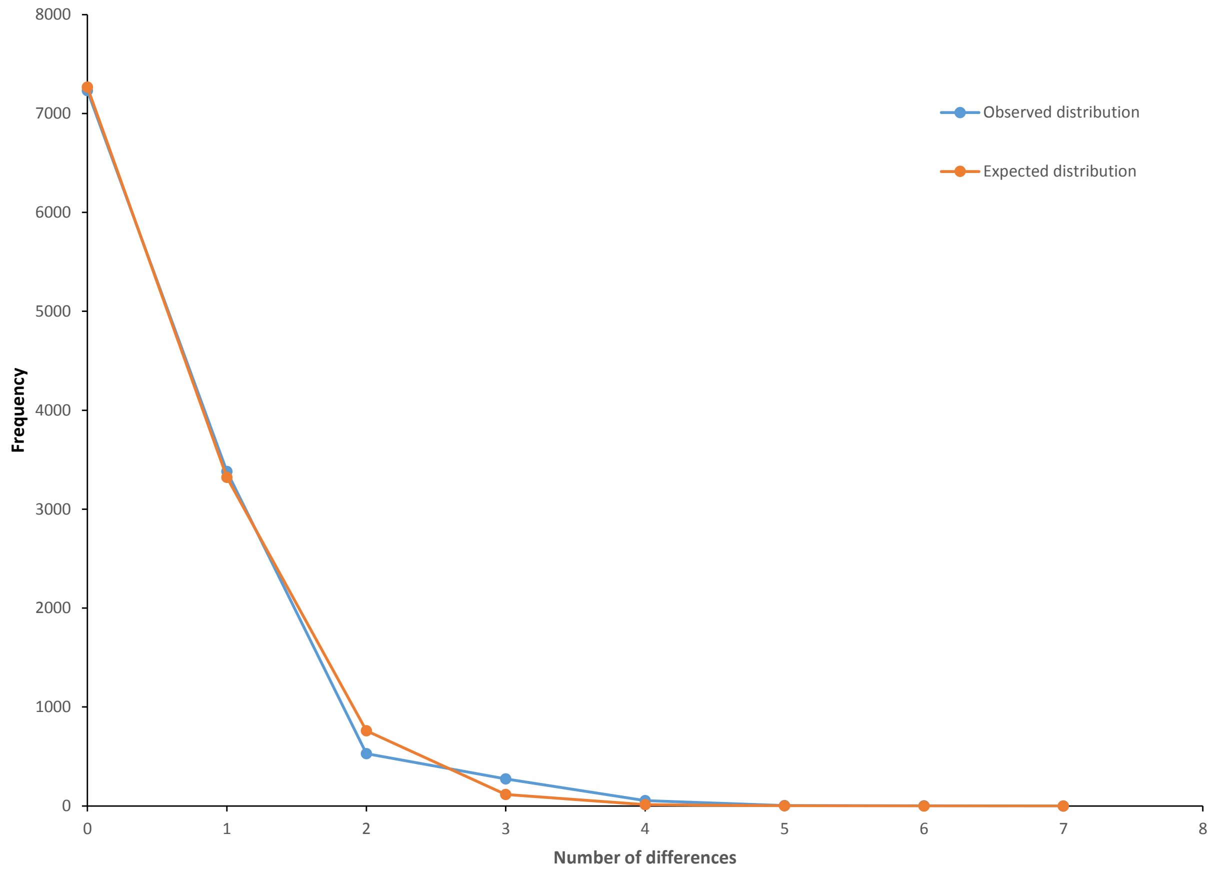


Fig. S2 Distribution of pairwise differences among 14 haplotypes belonging to Lineage 1.

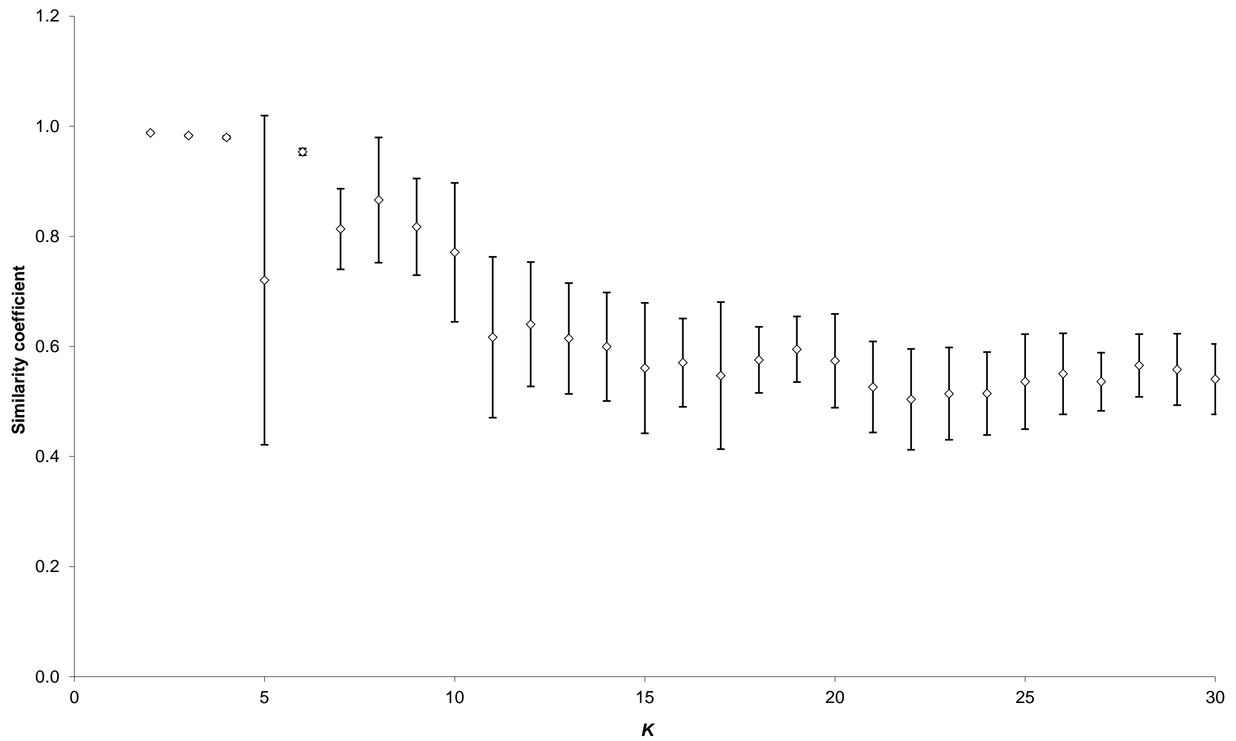


Fig. S3 Estimation of the number of genetic clusters following the method of Nordborg *et al.* (2005). Average similarity coefficients for each K with standard deviations according to Nordborg *et al.* (2005) and Ehrich *et al.* (2007).

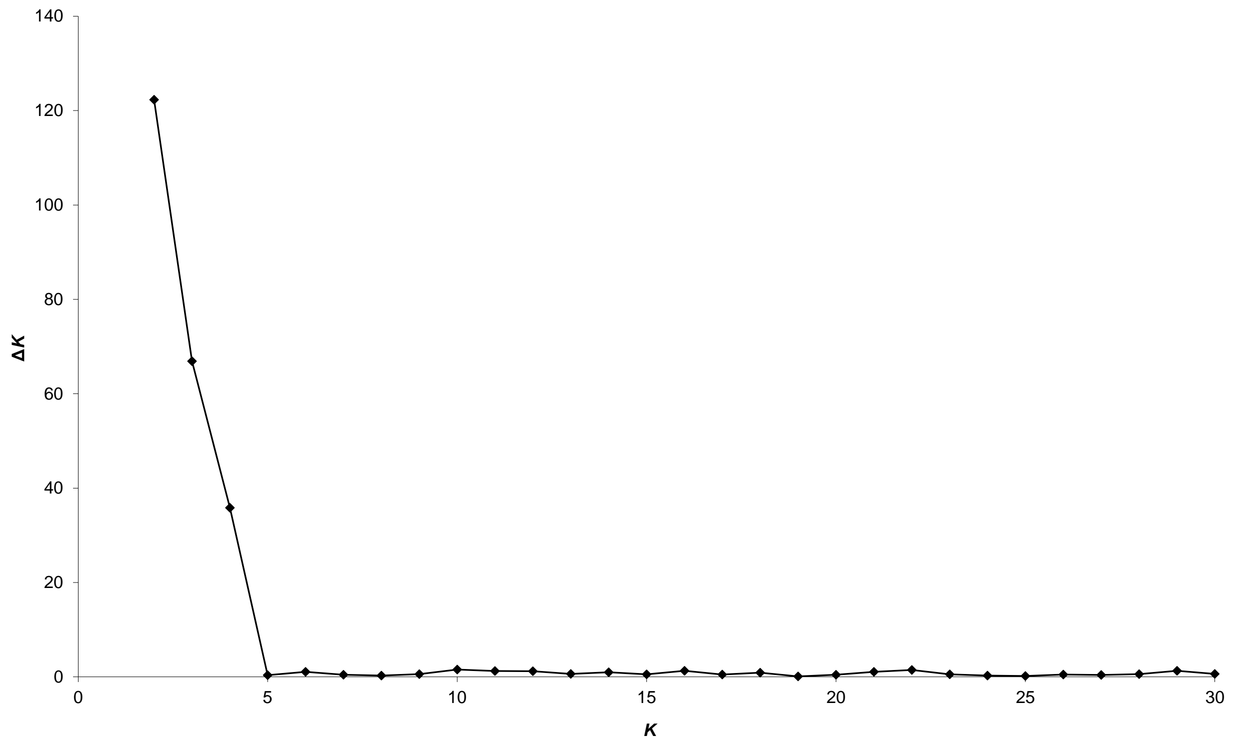


Fig. S4 Estimation of the number of genetic clusters following the method of Evanno, *et al.* (2005). Second-order rate of change in the probability between successive runs (DK) as a function of K (number of clusters).

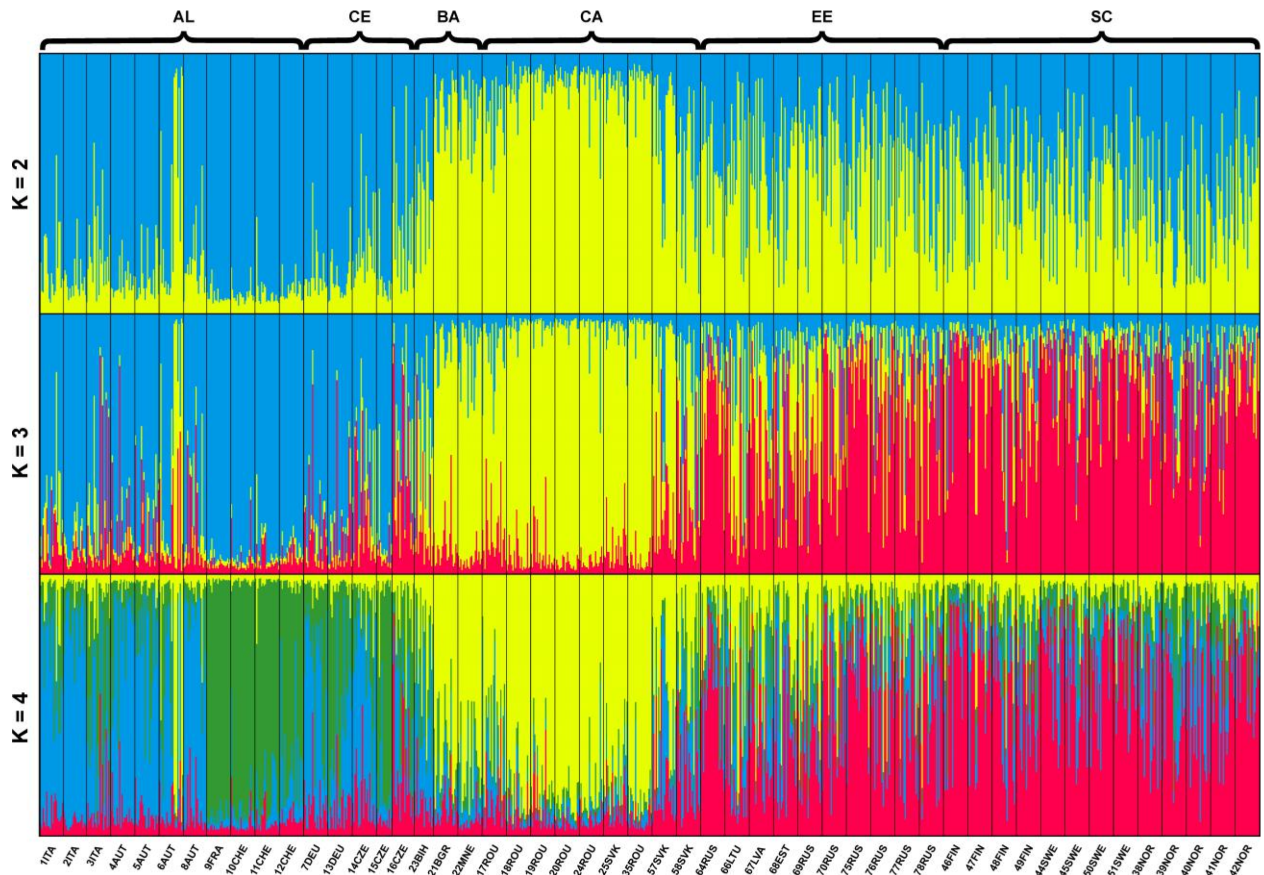


Fig. S5 Population structure of *Alnus incana* estimated by STRUCTURE from 1004 individuals from 51 localities. Each individual is represented by a vertical line, which is proportioned into *K* coloured segments, the length of each colour being proportional to the estimated membership coefficient. Black lines separate different populations labelled at the bottom of the figure with abbreviations corresponding to those in Table 1. Abbreviations at the top of the figure refer to six geographical areas defined in Table 1 and Fig. 1.

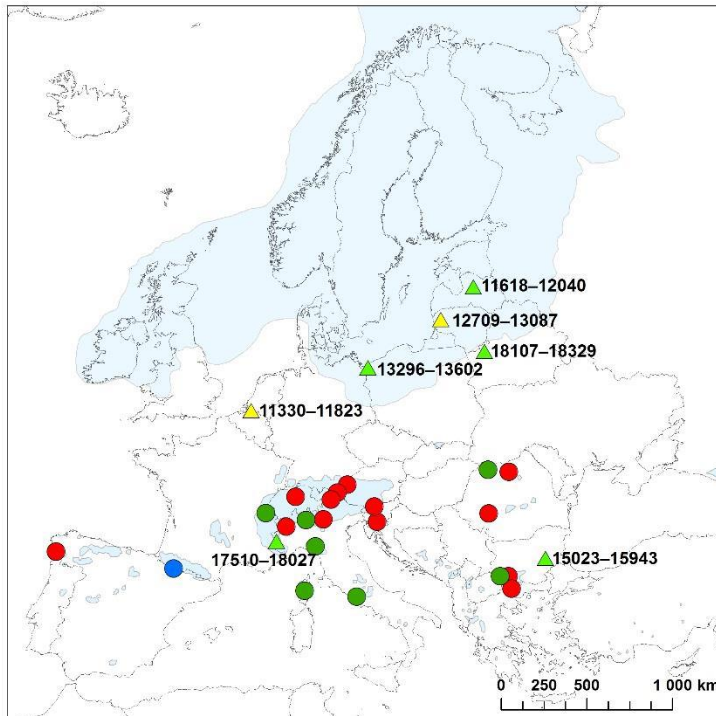


Fig. S6 Pollen records and macrofossils of *Alnus incana* at selected sites since the Late Pleniglacial (LPG) based on the data of Douda *et al.* (2014). The dots indicate pollen records from the LPG (blue), Late Glacial (green) and early Holocene (red). The triangles indicate macrofossil records of *A. incana* (yellow) and *Alnus* sp. (green) accompanied by the approximate age of the record. The bluish area indicates the maximal extent of the ice sheet during the LPG.

Table S1 Model checking by deviations of summary statistics for the observed data form the posterior predictive distribution of the most likely scenario of population origin. NAL (mean number of alleles), VAR (mean allele size variance), MGW (M index), N2P (mean number of alleles, two samples), V2P (mean allele size variance, two samples), FST (Wright's F_{ST}), LIK (mean index of classification, two samples) and DAS (shared allele distance, two samples). The Significance column indicates the position of the observed summary statistic in the 5% (*), 1% (**), or 0.1% (***) tails of the posterior predictive distribution.

a) Fennoscandia (Scenario 4.4)

Summary	Observed	Probability	Significance
NAL_1_1	6.1667	0.3115	
NAL_1_2	7.9444	0.34	
NAL_1_3	6.9444	0.153	
VAR_1_1	7.1304	0.722	
VAR_1_2	5.0766	0.4925	
VAR_1_3	5.2653	0.5275	
MGW_1_1	0.5401	0	***
MGW_1_2	0.7901	0.158	
MGW_1_3	0.7246	0.0505	
N2P_1_1&2	8.5556	0.344	
N2P_1_1&3	7.6111	0.1575	
N2P_1_2&3	8.5	0.1775	
V2P_1_1&2	5.8472	0.5845	
V2P_1_1&3	5.8936	0.593	
V2P_1_2&3	5.1924	0.507	
FST_1_1&2	0.0271	0.4885	
FST_1_1&3	0.0293	0.425	
FST_1_2&3	0.007	0.014	*
LIK_1_1&2	0.8249	0.0195	*
LIK_1_1&3	0.824	0.015	*
LIK_1_2&1	0.9177	0.027	*
LIK_1_2&3	0.8223	0.018	*
LIK_1_3&1	0.8391	0.009	**
LIK_1_3&2	0.7531	0.006	**
DAS_1_1&2	0.5022	0.987	*
DAS_1_1&3	0.5111	0.9905	**
DAS_1_2&3	0.5105	0.9885	*

b) Eastern Europe (Scenario 3.6)

Summary	Observed	Probability	Significance
NAL_1_1	8.4444	0.636	
NAL_1_2	6.1667	0.4545	
NAL_1_3	7.9444	0.577	
VAR_1_1	4.376	0.7425	
VAR_1_2	7.1304	0.927	
VAR_1_3	5.0766	0.8295	
MGW_1_1	0.729	0.01	**
MGW_1_2	0.5401	0	***
MGW_1_3	0.7901	0.0815	
N2P_1_1&2	9.1111	0.632	
N2P_1_1&3	9.5	0.599	
N2P_1_2&3	8.5556	0.5275	
V2P_1_1&2	5.5077	0.845	
V2P_1_1&3	4.785	0.797	
V2P_1_2&3	5.8472	0.872	
FST_1_1&2	0.0538	0.8775	
FST_1_1&3	0.0325	0.9475	
FST_1_2&3	0.0271	0.19	
LIK_1_1&2	1.1927	0.756	
LIK_1_1&3	0.993	0.5205	
LIK_1_2&1	0.9276	0.428	
LIK_1_2&3	0.8254	0.1605	
LIK_1_3&1	0.8904	0.3595	
LIK_1_3&2	0.9195	0.2	
DAS_1_1&2	0.4581	0.6625	
DAS_1_1&3	0.459	0.68	
DAS_1_2&3	0.5022	0.8345	

c) Central Europe (Scenario 2.5)

Summary	Observed	Probability	Significance
NAL_1_1	7.1667	0.284	
NAL_1_2	8.4444	0.5915	
NAL_1_3	6.1667	0.1465	
VAR_1_1	6.2913	0.494	
VAR_1_2	4.376	0.283	
VAR_1_3	7.1304	0.579	
MGW_1_1	0.7187	0.1135	
MGW_1_2	0.729	0.104	
MGW_1_3	0.5401	0.0025	**
N2P_1_1&2	9.4444	0.474	

N2P_1_1&3	7.9444	0.205	
N2P_1_2&3	9.1111	0.454	
V2P_1_1&2	5.7308	0.423	
V2P_1_1&3	6.5387	0.508	
V2P_1_2&3	5.5077	0.4025	
FST_1_1&2	0.0656	0.8275	
FST_1_1&3	0.0039	0	***
FST_1_2&3	0.0538	0.659	
LIK_1_1&2	0.9606	0.023	*
LIK_1_1&3	0.7472	0	***
LIK_1_2&1	1.1824	0.3165	
LIK_1_2&3	1.1918	0.337	
LIK_1_3&1	0.7478	0	***
LIK_1_3&2	0.9277	0.013	*
DAS_1_1&2	0.4557	0.99	**
DAS_1_1&3	0.5287	1	***
DAS_1_2&3	0.4581	0.989	*

d) Alps and Carpathians (Scenario 1.2)

Summary	Observed	Probability	Significance
NAL_1_1	7.1667	0.2325	
NAL_1_2	8.4444	0.3905	
NAL_1_3	6.3333	0.2895	
VAR_1_1	6.2913	0.505	
VAR_1_2	4.376	0.245	
VAR_1_3	4.7452	0.3045	
MGW_1_1	0.7187	0.092	
MGW_1_2	0.729	0.0615	
MGW_1_3	0.7015	0.1955	
N2P_1_1&2	9.4444	0.352	
N2P_1_1&3	8.0556	0.2175	
N2P_1_2&3	8.8333	0.3185	
V2P_1_1&2	5.7308	0.409	
V2P_1_1&3	6.0888	0.4685	
V2P_1_2&3	4.5013	0.256	
FST_1_1&2	0.0656	0.868	
FST_1_1&3	0.0402	0.4645	
FST_1_2&3	0.0247	0.31	
LIK_1_1&2	0.9601	0.036	*
LIK_1_1&3	0.9016	0.0045	**
LIK_1_2&1	1.1819	0.206	
LIK_1_2&3	1.0222	0.043	*

LIK_1_3&1	1.109	0.0945	
LIK_1_3&2	0.9487	0.0275	*
DAS_1_1&2	0.4557	0.988	*
DAS_1_1&3	0.4919	0.9985	**
DAS_1_2&3	0.4553	0.987	*

Table S2 Parameter estimates of the most likely demographic scenarios of the origin of populations of *Alnus incana* in different regions. Median estimates and 95% credible interval are shown. NCE, NCA, NEE, NSC, NBA, NAL are estimates of effective population size, FoundNEE, FoundNSC, FoundNCE, FoundNEE, FoundNCA indicate founding effective population size in the region. The number of generations is the time unit (t1 and t1a).

	Parametr	Median	Q_{05}	Q_{95}
Alps and Carpathians (Scenario 1.2)	NAL	2.75E+04	6.58E+03	8.67E+04
	NCA	7.79E+04	3.89E+04	9.85E+04
	NBA	8.76E+03	5.37E+03	1.33E+04
	FoundNAL	3.02E+03	6.61E+02	8.60E+03
	FoundNCA	5.29E+03	1.53E+03	9.40E+03
	t1a	3.74E+02	8.74E+01	8.05E+02
	t1	6.24E+02	5.10E+02	9.23E+02
Central Europe (Scenario 2.5)	NAL	8.52E+03	5.34E+03	1.27E+04
	NCA	7.75E+04	3.85E+04	9.85E+04
	NCE	1.87E+04	3.49E+03	8.22E+04
	FoundNCA	2.84E+03	6.38E+02	8.47E+03
	FoundNCE	5.94E+03	1.94E+03	9.54E+03
	t1a	3.05E+02	6.73E+01	7.37E+02
	t1	5.94E+02	5.06E+02	9.06E+02
Eastern Europe (Scenario 3.6)	NCA	3.46E+04	1.43E+04	8.41E+04
	NCE	2.94E+04	5.71E+03	9.00E+04
	NEE	4.98E+04	1.51E+04	9.52E+04
	FoundNCE	2.65E+03	1.06E+03	5.56E+03
	FoundNEE	6.29E+03	2.47E+03	9.56E+03
	t1a	3.35E+02	9.89E+01	7.57E+02
	t1	7.21E+02	5.21E+02	9.63E+02
Fennoscandia (Scenario 4.4)	NCE	7.64E+03	4.67E+03	1.18E+04
	NEE	7.51E+04	3.49E+04	9.82E+04
	NSC	2.71E+04	6.54E+03	8.64E+04
	FoundNEE	7.34E+03	3.03E+03	9.76E+03
	FoundNSC	6.96E+03	2.74E+03	9.71E+03
	t1a	3.71E+02	8.61E+01	8.03E+02
	t1	5.96E+02	5.07E+02	8.98E+02

5. Supporting information to the paper Havrdová A, Douda J, Krak K, Vít P, Hadincová V, Zákravský P, Mandák B (in press) **Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture.** *Molecular Ecology*.

Supporting information S1

Estimation of DNA ploidy level, genome size a chromosome counts

DNA ploidy levels (Suda *et al.* 2006) and absolute genome sizes (C-values; Greilhuber *et al.* 2005) of *Alnus* species were estimated using flow cytometry. Both fresh leaves and leaves stored in silica gel were used for analyses of different populations depending on the time needed to transport the material from the field to the laboratory. Young, intact leaf tissue of the analyzed plants and an appropriate amount of leaf tissue of the internal reference standard [*Bellis perennis*; 2C-value set to 3.38 pg following Schönswetter *et al.* (2007)] were chopped together using a sharp razor blade in a plastic Petri-dish containing 0.5 ml of ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20) (Otto 1990; Doležel *et al.* 2007). The crude suspension was filtered through a 0.42 μm nylon mesh to remove tissue debris and then incubated for at least 30 min at room temperature. Isolated nuclei were stained with 1 ml of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12\ H_2O$) supplemented with the AT-selective fluorochrome 4',6-diamidino-2-phenylindole (DAPI) and ß-mercaptoethanol at final concentrations of 4 μg/ml and 2 μg/ml, respectively. Immediately after staining, the relative fluorescence intensity of at least 3,000 particles was recorded on a CyFlow Space flow cytometer (Partec GmbH, Münster, Germany) equipped with a diode UV chip as an excitation light source.

A different staining procedure was used for absolute genome size estimation. The suspension of isolated nuclei was stained with a solution containing 1 ml of Otto II buffer (0.4M Na2HPO4·12 H2O), β-mercaptoethanol (final concentration of 2 μl/ml), propidium iodide (PI) and RNase IIA (both at final concentrations of 50 μg/ml). Samples were stained for 5 min at room temperature before being run through the flow cytometer CyFlow SL (Partec GmbH, Münster, Germany). Isolated stained nuclei were excited with a laser beam of 532 nm (solid-state laser Samba, Cobolt AB, Solna, Sweden), and the fluorescence intensity of 5,000 particles was recorded. Resulting histograms were evaluated using the application FloMax (PartecGmbH, Münster, Germany); DNA ploidy levels and absolute genome sizes were determined on the basis of the sample/standard ratio. Each plant was analyzed separately. Our previous pilot

study confirmed the lack of variation in the sample/standard ratio between fresh and silica-dried samples analyzed in the same way. The reliability of FCM measurements (i.e. between-plant differences) was repeatedly confirmed in simultaneous runs of *Alnus* accessions yielding distinct fluorescence intensities (i.e. resulting in furcate double peaks in FCM histograms (Greilhuber 2005).

To confirm the reliability of the ploidy estimates, FCM results were supplemented by conventional chromosome counts. Chromosome counts were obtained from somatic mitotic cells in root-tips of pot cultivated plants. The root tips were pre-treated in a saturated water solution of p-dichlorbenzene for approximately 2 hours, then fixed in a 3:1 mixture of 96% ethanol and acetic acid, macerated in a 1:1 mixture of ethanol and hydrochloric acid for 15 s, washed in water, and stained with lacto-propionic orcein. The number of chromosomes was determined under an NU Zeiss microsope with an Olympus E 510 camera attached. Two diploids (locality No. 1) and two tetraploids (locality No. 48) were analyzed.

References

Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, **2**, 2233–2244.

Greilhuber J, Doležel J, Lysák MA, Bennett MD (2005) The origin, evolution and proposed stabilization of the terms “genome size” and “C-value” to describe nuclear DNA contents. *Annals of Botany*, **95**, 255–260.

Greilhuber J (2005) Intraspecific Variation in Genome Size in Angiosperms: identifying its existence. *Annals of Botany*, **95**, 91–98.

Otto F (1990) DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman H, Darzynkiewicz AZ, editors. Methods in cell biology. *New York: Academic Press*, **33**, 105–110.

Schönswetter P, Suda J, Popp M, Weiss-Schneeweiss H, Brochmann C (2007) Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and Evolution*, **42**, 92–103.

Supporting Tables S1–S3

Table S1: Locations, haplotype and microsatellite details. Pop no and **Pop name** are the codes and names of populations in our database, populations with the prefix L and KF are related to samples previously published by Lepais et al. 2013 (L) and King and Ferris 1998 (KF); **Country** – country abbreviation; **Latitude** and **Longitude** – GPS coordinates of sampling sites in WGS84; **Ploidy** – ploidy levels according to (Mandák et al., unpublished data, for methods see Appendix S1); **Hapl** – haplotypes occurring in each populations with the number of individuals of the specific haplotype in brackets; haplotype labels correspond to those given in Fig. 3; **N SSRs** – number of individuals from each population used for microsatellite analysis; **Group** – geographic groups delimited for Bayesian analysis of European diploid populations, two groups in some cases mean different groups for each analysis, i.e. Structure/ABC; N_A – number of alleles; A_r – rarefied allelic richness; hA_r – rarefied allelic richness for haploid genome; H_e – gene diversity corrected for sample size; hH_e – gene diversity corrected for sample size for haploid genome; F_{IS} – inbreeding coefficient for diploid populations and its significance tested by 999 permutations, * P < 0.05, ** P < 0.01, *** P < 0.001; **NA** – not available.

Pop no	Pop name	Country	Latitude	Longitude	Ploidy	Hapl	N SSRs	Group	N_A	A_r	hA_r	H_e	hH_e	F_{IS}	
1	České Budějovice	CZE	49.009135	14.436995	2	1 (3)	20	HER	5.74	4.87	3.27	0.6057	0.4925	-0.010	
2	San Daniele del Friuli	ITA	46.135526	12.949137	2	1 (3)	20	ALP	5.37	4.41	3.06	0.5809	0.4861	-0.040	
3	Busche	ITA	46.035783	11.987871	2	1 (3)	20	ALP	5.63	4.82	3.45	0.6042	0.5217	0.067	*
4	Denno	ITA	46.26612	11.06297	2	1 (3)	20	ALP	5.74	4.86	3.42	0.6208	0.5263	0.040	
5	CasolaValsenio	ITA	44.197562	11.602491	2	5 (2)	20	APE	5.26	4.50	3.13	0.5945	0.5342	0.069	*
6	Montalto	ITA	43.52565	11.685982	2	5 (3)	20	APE	5.68	4.78	3.53	0.5964	0.5438	0.117	***
7	Gorreto	ITA	44.606015	9.296819	2	1 (3)	20	APE	6.16	4.98	3.54	0.6191	0.5548	-0.010	
8	Aspet	FRA	43.029703	0.787468	2	1 (3)	20	WES/IBE	5.53	4.66	3.52	0.5962	0.5551	0.088	**
9	Eugi	ESP	43.014991	-1.507978	2	1 (3)	20	WES/IBE	5.37	4.72	3.37	0.6069	0.5528	0.020	
10	Vargas	ESP	43.33294	-3.970437	4	10 (3)	20	NA	7.74	5.84	4.24	0.6755	0.6155	NA	
11	Chaves	PRT	41.753617	-7.458716	4	10 (2)	20	NA	8.11	5.84	4.17	0.6717	0.6296	NA	
12	Vilares	PRT	41.881106	-8.723562	4	10 (3)	20	NA	8.11	5.98	4.11	0.6745	0.6273	NA	
13	Fernitz bei Graz	AUT	46.974117	15.487067	2	1 (2)	19	ALP	6.00	4.96	3.52	0.6374	0.5632	0.040	
14	Gibina	SVN	46.52324	16.299652	2	1 (3)	20	ALP	5.84	4.90	3.54	0.6201	0.5524	0.040	
15	Klagenfurt am Wörthersee	AUT	46.620793	14.179016	2	1 (2)	20	ALP	6.26	5.00	3.38	0.6197	0.5230	0.050	
16	Breitbrunn am Chimsee	DEU	47.88315	12.385533	2	1 (3)	20	ALP	6.11	5.07	3.44	0.6298	0.5327	0.030	

17	Bruckberg	DEU	48.496548	11.995551	2	1 (3)	20	HER	6.95	5.63	3.85	0.6444	0.5912	0.065	*
18	Cazzago Brabbia	ITA	45.794781	8.733642	2	1 (3)	20	ALP	6.32	5.12	3.53	0.6159	0.5549	0.030	
19	Laffrey	FRA	44.998195	5.777764	2	1 (3)	20	ALP	6.21	5.16	3.70	0.6168	0.5666	0.060	*
20	Kisslegg	DEU	47.773332	9.921796	2	1 (3)	20	ALP	5.89	4.99	3.52	0.6143	0.5574	0.040	
21	Abtsgmünd	DEU	48.908048	9.955931	2	1 (3)	20	HER	6.42	5.33	3.71	0.6298	0.5598	0.040	
22	Chomutov	CZE	50.500603	13.329811	2	1 (1)	20	HER	6.16	5.06	3.49	0.5952	0.5324	-0.010	
23	Orlické Záhoří	CZE	50.237385	16.537688	2	1 (3)	20	HER	6.16	5.10	3.62	0.6248	0.5727	0.062	*
24	Negreni	ROU	46.950346	22.724281	2	26 (2), 27 (1)	20	CAR	6.42	5.24	3.76	0.6135	0.5784	0.040	
25	Praid	ROU	46.568674	25.167343	2	1 (3)	20	CAR	6.16	5.07	3.7	0.6132	0.5701	0.058	*
26	Comăneşti	ROU	46.436926	26.370439	2	1 (3)	20	CAR	6.63	5.47	3.65	0.6403	0.5654	0.069	*
27	Tatarani	ROU	44.994135	25.284871	2	27 (2)	20	CAR	6.74	5.37	3.82	0.6193	0.5798	0.071	*
28	Vaglevtsi	BGR	42.931827	25.635955	2	32 (3)	20	BAL	6.53	5.23	3.98	0.6179	0.6056	0.058	*
29	Simitli	BGR	41.882727	23.145251	2	34 (3)	20	BAL	6.32	5.08	3.5	0.5968	0.5374	0.010	
30	Lakatnik	BGR	43.087282	23.380836	2	32 (3)	20	BAL	6.26	5.20	3.62	0.6277	0.5837	0.020	
31	Rudozem	BGR	41.514264	24.886083	2	33 (3)	20	BAL	5.63	4.80	3.53	0.6016	0.5537	0.030	
32	Marina	GRC	40.859622	21.487761	2+3+4	35 (1), 36 (2)	20	NA	8.95	6.51	4.52	0.7222	0.6717	NA	
33	Librazhd	ALB	41.159971	20.192281	4	36 (1), 37 (2)	20	NA	9.26	6.62	4.59	0.7440	0.6970	NA	
34	Žabljak	MNE	43.147406	19.296672	4	39 (1), 40 (1)	20	NA	7.95	6.12	4.44	0.7174	0.6787	NA	
35	Vitez	BIH	44.188385	17.737766	2+3+4	1 (3)	20	NA	8.05	6.19	4.45	0.7212	0.6812	NA	
36	Pakrac	HRV	45.444461	17.243923	2	1 (3)	20	ALP	5.79	4.90	3.65	0.6162	0.5675	0.095	***
38	Beckum	DEU	51.76021	8.009907	2	1 (2)									
39	Eichhorst	DEU	52.896181	13.642833	NA	1 (2)									
41	Tübingen	DEU	48.557697	9.045914	NA	1 (3)									
42	Międziwodzie	POL	54.005942	14.701656	2	1 (1)									
45	Oxford	GBR	51.746714	-1.248544	2	1 (1)									
48	Alcalá de los Gazules	ESP	36.520361	-5.619806	4	9 (3)	20	NA	8.00	5.94	4.07	0.6847	0.6469	NA	
49	Collias	FRA	43.951389	4.477194	2	1 (2)									
52	Dumbrăviţa	ROU	46.1015	22.102028	NA	27 (3)									
53	Cârțișoara	ROU	45.693583	24.572972	NA	27 (1), 29 (1)									
56	Ersizlerdere	TUR	41.84117	33.73594	2	23 (3)									
57	Maçka	TUR	40.68294	39.66249	2	21 (3)									
59	Giosla	GBR	58.101531	-6.859649	NA	1 (1)									

62	SanktGalen	CHE	47.416745	9.412158	2	1 (1)									
63	Bjeloperica	SRB	43.955735	19.982071	4	41 (1)									
69	Rangsdorf	DEU	52.283306	13.453375	2	1 (3)	20	NEA	6.42	5.32	3.81	0.6337	0.5801	0.063	*
70	Salzwedel	DEU	52.885725	11.124123	2	1 (3)	20	NEA	6.42	5.22	3.58	0.6221	0.5540	0.050	
71	Vierde	DEU	52.882186	9.743381	2	1 (3)	20	WES	6.16	5.09	3.47	0.6235	0.5540	0.060	*
72	Holdorf	DEU	52.544087	8.179687	2	1 (3)	20	WES	6.16	5.20	3.76	0.6495	0.5983	0.050	
73	Naarden	NLD	52.285302	5.142885	2	1 (3)	20	WES	6.42	5.29	3.52	0.6290	0.5396	0.000	
74	Kintbury	GBR	51.402706	-1.43685	2	1 (3)	20	BRI	6.37	5.22	3.58	0.6260	0.5548	0.050	
75	Soulme	BEL	50.184217	4.742358	2	1 (3)	20	WES	5.79	4.84	3.28	0.6027	0.5127	0.067	*
76	Fourges	FRA	49.120367	1.64608	2	1 (3)	20	WES	7.00	5.58	3.82	0.6441	0.5856	0.020	
77	Langeais	FRA	47.338355	0.436995	2	1 (3)	20	WES	6.21	5.02	3.33	0.5951	0.5169	0.040	
78	Les Graulges	FRA	45.49031	0.47829	2	1 (2)	20	WES	6.68	5.63	3.74	0.6540	0.5886	0.059	*
79	Sainte-Marguerite	FRA	45.214136	3.602951	2	1 (2)	20	HER	5.89	4.90	3.44	0.6032	0.5423	0.087	**
80	La Charité-sur-Loire	FRA	47.177084	3.007033	2	1 (2)	16	HER	6.16	5.39	3.61	0.6312	0.5790	0.030	
81	Saulxures-sur-Moselotte	FRA	47.948524	6.755554	2	1 (3)	20	HER	6.11	5.08	3.57	0.6161	0.5639	0.076	**
82	Erlensee	DEU	50.150774	8.97237	2	1 (3)	20	HER	6.05	5.03	3.43	0.6089	0.5402	0.030	
83	Zorge	DEU	51.641773	10.631134	2	1 (3)	20	HER	6.42	5.25	3.73	0.6241	0.5596	0.020	
84	Tullstorp	SWE	56.330425	13.003779	2	1 (3)	20	SCA	6.58	5.34	3.76	0.6486	0.5859	0.064	**
85	Ljungskile	SWE	58.253694	11.971803	2	1 (3)	20	SCA	6.37	5.28	3.58	0.6156	0.5494	0.020	
86	Notodden	NOR	59.565325	9.191127	2	1 (3)	20	SCA	6.11	5.26	3.74	0.6550	0.5938	0.072	**
87	Etne	NOR	59.635095	5.904407	2	1 (3)	20	SCA	7.05	5.55	3.81	0.6409	0.5784	0.057	*
88	Muruvik	NOR	63.438946	10.830525	2	1 (1)									
89	Kuopio	FIN	62.901204	27.603156	2	1 (3)	20	SCA	6.16	4.96	3.47	0.6040	0.5299	0.069	*
90	Pälkäne	FIN	61.334245	24.268583	2	1 (3)	20	SCA	6.63	5.24	3.73	0.6154	0.5682	0.010	
91	Turku	FIN	60.461435	22.38869	2	1 (3)	20	SCA	6.11	5.01	3.54	0.6211	0.5496	0.040	
92	Örbyhus	SWE	60.188525	17.710743	2	1 (3)	20	SCA	6.16	5.00	3.55	0.6115	0.5560	0.010	
93	Friggesund	SWE	61.853668	16.62217	2	1 (3)									
94	Väderstad	SWE	58.331634	14.82418	2	1 (3)	20	SCA	6.68	5.49	4.04	0.6568	0.6362	0.068	*
95	Nykøbing Falster	DNK	54.743651	11.926331	2	1 (3)									
96	Norra Kättbo	SWE	60.817195	14.10483	2	1 (3)									
97	Sandanski	BGR	41.667906	23.38705	2	33 (1)									
98	Šaštín-Stráže	SVK	48.621684	17.159782	2	1 (3)	20	CAR	6.21	5.10	3.49	0.6211	0.5668	0.030	

99	Tisovec	SVK	48.704411	19.922088	2	1 (3)	20	CAR	5.74	4.80	3.56	0.6025	0.5359	0.030	
100	Kráľovský Chlmec	SVK	48.434422	21.955853	2	1 (2), 27 (1)	20	CAR	6.37	5.26	3.76	0.6314	0.5803	0.040	
101	Klubina	SVK	49.356967	18.904203	2	1 (3)	11	CAR	5.74	5.56	4.08	0.6693	0.6310	0.094	**
102	Caxarias	PRT	39.7239	-8.50795	4	10 (3)	19	NA	8.00	5.99	4.16	0.6828	0.6633	NA	
103	Vila Soeiro	PRT	40.53775	-7.33749	4	10 (3)									
104	Allariz	ESP	42.18572	-7.81806	4	10 (3)									
105	Lugo	ESP	43.00423	-7.57427	4	11 (3)									
106	Laurca	ESP	43.52975	-6.53463	4	10 (3)	20	NA	7.95	5.91	4.12	0.6736	0.6446	NA	
110	Macejl	HRV	46.265492	15.867676	2	1 (3)									
111	Lornís	ESP	42.42055	-7.6733	4	10 (3)									
113	Armadale	GBR	57.065272	-5.898843	2	1 (2)									
114	Sovana	ITA	42.65825	11.635333	2	5 (2), 6 (1)	17	APE	5.26	4.56	3.32	0.6082	0.5515	0.060	
115	Cozla	ROU	44.621917	22.0237	2	27 (2), 28 (1)									
116	Olonets	RUS	60.76	32.81	2	1 (3)	22	NEA	6.21	4.95	3.58	0.6064	0.5644	0.040	
118	Ax-les-Thermes	FRA	42.718833	1.91	2	1 (2)									
125	Chakvistskali	GEO	41.690639	41.829861	2	21 (1), 24 (1)									
127	Tulliemet	GBR	56.66	-3.62	NA	1 (3)	20	BRI	5.74	4.89	3.45	0.6240	0.5490	0.030	
128	Orsomarso	ITA	39.797806	15.890444	2	14 (2)									
129	Martirano	ITA	39.106806	16.253944	2	15 (1)	20	NA	5.89	4.73	3.48	0.5675	0.4989	0.070	*
130	Vladičin Han	SRB	42.732607	22.062353	2	25 (1)									
132	Brzeće	SRB	43.306288	20.899779	4	7 (3)									
133	Jošanička Banja	SRB	43.399503	20.790895	2	7 (2), 30 (1)									
134	Prijepolje	SRB	43.365756	19.724729	4	43 (2)									
135	Milići	BIH	44.188049	19.071714	4	41 (1), 42 (1)									
139	Mosina	POL	52.263695	16.800589	2	1 (1)	20	NEA	6.58	5.31	3.69	0.6279	0.5737	0.030	
140	Leszno	POL	52.357538	20.626312	2	1 (2)	20	NEA	6.16	5.19	3.60	0.6329	0.5596	0.050	
141	Białowieża	POL	52.702965	23.831147	2	1 (2)	20	NEA	7.00	5.62	3.89	0.646	0.5831	0.069	**
142	Strenči	LVA	57.621585	25.677251	2	1 (2)	20	NEA	5.95	4.92	3.43	0.6003	0.5421	-0.010	
143	Pushkino	RUS	56.009539	37.857896	2	1 (3)									
144	Rostov	RUS	57.238451	39.473761	2	1 (2)	20	NEA	6.53	5.28	3.61	0.6259	0.5557	0.090	**
145	Rumšiškės	LTU	54.876544	24.202453	2	1 (3)									
146	Ardu	EST	59.117782	25.372891	2	1 (1)									

148	Sovinjak	HRV	45.38689	13.925666	2	1 (2), 8 (1)									
150	Żukowo	POL	54.306444	18.308472	2	1 (3)									
152	Rossleithen	AUT	47.689385	14.293535	2	1 (1)									
153	Novi Yarylovychi	UKR	51.999694	31.015861	2	1 (2)	20	NEA	5.79	4.74	3.50	0.5974	0.5582	0.086	**
154	Komarivka	UKR	50.494778	29.535556	2	1 (3)									
156	Malyns'k	UKR	51.086556	26.578861	2	1 (2)									
157	Lelekhivka	UKR	49.936389	23.7035	2	1 (2)	11	CAR	5.11	4.98	3.47	0.6034	0.5385	0.050	
158	Shenkursk	RUS	62.194511	42.776253	2	1 (3)									
160	Severodvinsk	RUS	64.610619	39.818228	2	1 (1)									
161	Kubenskoye	RUS	59.432248	39.711738	2	1 (2)									
162	Potter Heigham	GBR	52.730932	1.571506	2	1 (3)	20	BRI	6.26	5.17	3.57	0.6362	0.5694	0.050	
163	Sheffield	GBR	53.401544	-1.782336	2	1 (2)	20	BRI	6.26	5.19	3.70	0.6329	0.5801	0.079	*
166	Glengarriff	IRL	51.75465	-9.56896	2	1 (2)									
167	Letterfrack	IRL	53.55962	-9.87507	2	1 (1)									
L_BMO	Ben Mehidi	DZA	36.843333	7.980833	2	16 (3)	20	NA	4.21	3.65	2.67	0.5153	0.4418	-0.050	
L_BOU	Bab Taza	MAR	35.013611	-5.190833	4	13 (2)	20	NA	7.95	6.06	4.13	0.7093	0.6396	NA	
L_DAT	Guerbes	DZA	36.931389	7.2475	2	18 (2)	20	NA	5.00	4.25	3.23	0.5102	0.5105	0.030	
L_DER	El Kala	DZA	36.867222	8.388889	2	16 (3)	20	NA	4.74	4.05	2.92	0.4962	0.4640	-0.010	
L_DHA	Ain Draham	TUN	36.800833	8.659722	2	16 (3)	20	NA	4.21	3.73	2.72	0.5076	0.4446	-0.073	*
L_LAO	El Kala	DZA	36.8825	8.580833	2	16 (3)	20	NA	4.58	3.89	3.01	0.5011	0.4803	0.011	
L_TAY	Tayenza	MAR	35.272778	-5.448333	4	12 (3)	20	NA	8.37	5.85	4.01	0.6882	0.6665	NA	
L_TIT	Ouchtata	TUN	36.964167	8.962778	2	20 (2), 16 (1)	20	NA	4.37	3.74	2.84	0.5457	0.4942	0.030	
L_ZIA	Cap Serrat	TUN	37.189167	9.213333	2	19 (2), 16 (1)	20	NA	4.74	4.11	3.02	0.5368	0.5050	-0.040	
L_ZLE	Ain Draham	TUN	36.823889	8.838056	2	16 (2)	20	NA	4.95	4.21	2.98	0.5604	0.5111	0.010	
KF1	Bodensdorf	AUT	46.681498	13.968251	2	1 (3)									
KF2	Drosendorfan der Thaya	AUT	48.867808	15.623271	2	1 (2)									
KF3	Gloggnitz	AUT	47.676915	15.936001	2	1 (3)									
KF4	Millstatt	AUT	46.804967	13.567049	2	1 (1), 2 (1)									
KF5	Retz	AUT	48.758053	15.946232	2	1 (1)									
KF7	Florennes	BEL	50.251579	4.659573	2	1 (3)									
KF8	Anheé	BEL	50.315208	4.839036	2	1 (2)									
KF9	Plachkovtsi	BGR	42.827547	25.484146	2	32 (2)									

KF11	Voneshta Voda	BGR	42.876279	25.642505	2	32 (2), 31 (1)
KF13	Carazzi	FRA	42.015759	8.885536	2	16 (1)
KF14	Cardo-Torgia	FRA	41.869914	8.978614	2	16 (1)
KF17	Sagone	FRA	42.113824	8.702188	2	16 (1)
KF18	Macinaggio	FRA	42.984558	9.453735	2	17 (1)
KF19	Đurđevac	HRV	46.039742	17.052929	2	1 (2)
KF24	Poděbrady	CZE	50.142781	15.112743	2	1 (2)
KF34	Lezay	FRA	46.267254	0.009378	2	1 (2)
KF36	LeNouvion-en-Thiérache	FRA	50.018108	3.785419	2	1 (1)
KF39	Angoulême	FRA	45.670489	0.107054	2	1 (2)
KF40	Clemecy	FRA	47.472018	3.528215	2	1 (1)
KF44	Dettingen	DEU	48.428847	8.593329	2	1 (2)
KF47	Wasserburg	DEU	47.566881	9.637479	2	1 (2)
KF48	Leianokladi	GRC	38.915613	22.27272	4	38 (1)
KF49	Homokszentgyörgy	HUN	46.119946	17.598059	2	1 (1)
KF50	Mike	HUN	46.235339	17.535228	2	1 (2)
KF53	Somogysárd	HUN	46.41575	17.597299	2	1 (3)
KF57	Viterbo	ITA	42.424217	12.108994	2	1 (1)
KF64	Kuinre	NLD	52.801181	5.915308	2	1 (1)
KF65	Eidesåsen	NOR	60.051616	6.538543	2	1 (2)
KF68	Jarlsberg	NOR	59.301238	10.367812	2	1 (2), H4 (1)
KF72	Biała Podlaska	POL	52.016609	23.326928	2	1 (3)
KF75	Wichrowo	POL	54.026528	20.410194	2	1 (2)
KF76	Wyszków	POL	52.551567	21.507783	2	1 (2)
KF83	Billesholm	SWE	unknown	unknown	2	1 (3)
KF84	Lyckeby	SWE	56.230184	15.629082	2	1 (2), H3 (1)
KF86	Vanneberga	SWE	56.083467	13.90667	2	1 (2)
KF90	Göktaş Köyü	TUR	39.998279	43.560745	2	21 (2)
KF91	Hopa	TUR	41.39694	41.42395	2	21 (2)
KF97	Tirebolu	TUR	41.01203	38.84903	2	21 (3)
KF98	Ulubey	TUR	40.874194	37.762756	2	21 (1), H22 (1)
KF102	unknown	BGR	unknown	unknown	2	32 (2)

KF104	unknown	DEU	unknown	unknown	2	1 (2)
KF106	unknown	FRA	unknown	unknown	2	1 (1)
KF107	unknown	FRA	unknown	unknown	2	1 (2), H3 (1)
KF108	unknown	FRA	unknown	unknown	2	1 (1)
KF109	unknown	FRA	unknown	unknown	2	1 (3)
KF110	unknown	FRA	unknown	unknown	2	1 (1)
KF111	unknown	FRA	unknown	unknown	2	1 (2)
KF112	unknown	ITA	unknown	unknown	2	1 (3)
KF114	unknown	NOR	unknown	unknown	2	1 (2)
KF115	Sulechów	POL	52.106505	15.611572	2	1 (2)
KF116	unknown	NLD	unknown	unknown	2	1 (1)
KF117	Veenendaal	NLD	52.007295	5.500183	2	1 (3)
KF118	Tilburg	NLD	51.652495	5.122242	2	1 (3)
KF122	unknown	GBR	unknown	unknown	2	1 (2)

References

King A, Ferris C (1998). Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Lepais O, Muller SD, Ben Saad-Limam S, Benslama M, Rhazi L, Belouahem-Abed D, Daoud-Bouattour A, Gammar AM, Ghrabi-Gammar Z, Bacles CFE (2013) High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PLoS One*, **8**, e75029.

Table S2: Parameter estimates of the most likely scenarios and the bias and precision of the estimations. Bias – mean relative bias; **RMSE** – relative mean square error; **50%** and **95% coverage** – proportion of simulations in which the true value lies within the 50% and 95% credible interval around the estimate; **Factor 2** – proportion of estimated values falling within the interval between 50% and 200% of the true values. Times are given in generations.

	Mean	Median	Mode	Q 0.05	Q 0.95	Bias	RMSE	50% coverage	95% coverage	Factor 2
A. Postglacial expansion from APE to ALP										
N_{APE}	1.29E+04	1.26E+04	1.14E+04	7.87E+03	1.92E+04	0.028	0.546	0.510	0.962	0.966
N_{ALP}	6.51E+04	6.70E+04	9.60E+04	2.65E+04	9.75E+04	-0.143	6.019	0.520	0.950	0.580
NF_{ALP}	6.91E+03	7.26E+03	8.99E+03	2.81E+03	9.72E+03	0.391	1.737	0.514	0.962	0.800
t2	6.34E+02	5.93E+02	5.03E+02	5.06E+02	8.99E+02	0.080	0.267	0.504	0.956	1.000
t2a	4.78E+02	4.72E+02	5.23E+02	1.05E+02	8.61E+02	0.045	3.448	0.484	0.962	0.650
B. Postglacial expansion from BAL to CAR										
N_{BAL}	9.83E+03	9.61E+03	9.66E+03	3.76E+03	1.68E+04	0.026	0.659	0.502	0.962	0.940
N_{CAR}	5.61E+04	5.52E+04	3.98E+04	1.69E+04	9.56E+04	-0.120	2.614	0.522	0.938	0.632
NF_{CAR}	6.23E+03	6.54E+03	8.45E+03	1.80E+03	9.67E+03	0.583	11.873	0.490	0.934	0.670
t2	8.67E+02	6.97E+02	5.74E+02	5.31E+02	1.67E+03	0.111	0.620	0.510	0.946	0.860
t2a	5.94E+02	3.77E+02	1.37E+02	4.62E+01	1.90E+03	-0.181	6.014	0.508	0.960	0.458
C. Postglacial expansion from WES to BRI										
N_{BRI}	5.83E+04	5.94E+04	9.88E+04	1.68E+04	9.68E+04	-0.118	8.296	0.498	0.952	0.494
N_{WES}	2.01E+04	1.97E+04	1.89E+04	1.27E+04	2.85E+04	0.030	0.547	0.506	0.960	0.964
NF_{BRI}	7.22E+03	7.61E+03	9.81E+03	3.13E+03	9.78E+03	0.424	2.425	0.492	0.956	0.772
t2	6.19E+02	5.80E+02	5.00E+02	5.05E+02	8.77E+02	0.093	0.262	0.534	0.962	1.000
t2a	4.37E+02	4.24E+02	3.82E+02	8.91E+01	8.22E+02	0.099	3.387	0.528	0.960	0.642
D. Postglacial expansion from IBE to WES and HER										
N_{HER}	6.80E+04	6.98E+04	9.85E+04	3.06E+04	9.76E+04	-0.057	3.332	0.506	0.948	0.594
N_{IBE}	5.23E+03	5.03E+03	4.30E+03	2.99E+03	8.18E+03	0.053	0.657	0.502	0.950	0.962
N_{WES}	5.56E+04	5.36E+04	3.58E+04	1.79E+04	9.57E+04	-0.084	4.153	0.524	0.958	0.546
NF_{HER}	6.38E+03	6.56E+03	6.79E+03	2.36E+03	9.63E+03	0.283	2.780	0.490	0.950	0.838
NF_{WES}	6.57E+03	6.82E+03	6.85E+03	2.59E+03	9.66E+03	0.227	2.418	0.504	0.960	0.866

t2	6.56E+02	6.18E+02	5.10E+02	5.09E+02	9.22E+02	0.091	0.270	0.496	0.946	1.000
t2a	4.62E+02	4.47E+02	4.40E+02	1.22E+02	8.43E+02	0.118	2.507	0.506	0.952	0.778
E. Admixture from CAR and WES in HER										
NA	2.33E+03	1.50E+03	2.84E+02	1.53E+02	7.11E+03	0.733	14.033	0.486	0.938	0.610
N_{CAR}	5.65E+04	5.50E+04	4.64E+04	2.65E+04	9.08E+04	0.046	0.459	0.524	0.966	0.996
N_{WES}	1.50E+04	1.42E+04	1.28E+04	7.31E+03	2.53E+04	0.029	0.480	0.558	0.980	0.990
N_{HER}	4.99E+04	4.62E+04	3.14E+04	2.10E+04	9.03E+04	0.038	0.683	0.530	0.962	0.982
t1	2.18E+03	1.74E+03	1.12E+03	1.08E+03	4.96E+03	-0.053	0.558	0.516	0.978	0.906
t2	7.20E+02	7.06E+02	5.58E+02	5.20E+02	9.60E+02	-0.100	0.256	0.566	0.938	1.000
r1	4.22E-01	4.11E-01	4.38E-01	1.14E-01	7.84E-01	0.147	2.113	0.540	0.962	0.950
F. Postglacial expansion from CAR to NEA and HER										
N_{HER}	6.47E+04	6.58E+04	9.82E+04	2.70E+04	9.74E+04	-0.043	3.816	0.542	0.946	0.580
N_{CAR}	1.78E+04	1.76E+04	1.74E+04	1.15E+04	2.51E+04	0.054	0.509	0.496	0.948	0.974
N_{NEA}	6.44E+04	6.55E+04	9.94E+04	2.67E+04	9.73E+04	-0.119	3.509	0.516	0.956	0.612
NF_{HER}	6.49E+03	6.69E+03	7.23E+03	2.50E+03	9.64E+03	0.242	2.896	0.518	0.948	0.862
NF_{NEA}	6.75E+03	7.05E+03	7.93E+03	2.73E+03	9.69E+03	0.301	1.964	0.480	0.946	0.870
t2	6.12E+02	5.71E+02	5.00E+02	5.04E+02	8.66E+02	0.039	0.261	0.478	0.944	1.000
t2a	5.67E+02	5.76E+02	5.25E+02	1.95E+02	8.95E+02	0.099	2.450	0.484	0.942	0.720
G. Admixture from WES and NEA in SCA										
NA	4.71E+03	3.84E+03	2.70E+03	4.31E+02	1.18E+04	0.157	1.663	0.492	0.958	0.836
N_{SCA}	6.27E+04	6.21E+04	5.46E+04	3.20E+04	9.52E+04	0.008	1.869	0.492	0.950	0.928
N_{NEA}	6.49E+04	6.55E+04	7.07E+04	3.15E+04	9.58E+04	-0.003	2.095	0.514	0.946	0.822
N_{WES}	4.63E+04	4.17E+04	2.99E+04	1.47E+04	9.10E+04	-0.036	9.629	0.506	0.962	0.816
NF_{NEA}	6.20E+03	6.38E+03	7.19E+03	2.25E+03	9.55E+03	0.122	1.763	0.528	0.948	0.876
NF_{WES}	6.89E+03	7.19E+03	7.90E+03	2.95E+03	9.70E+03	0.116	1.413	0.510	0.960	0.898
t1	1.36E+03	1.21E+03	1.05E+03	1.03E+03	2.08E+03	-0.031	0.562	0.492	0.940	0.900
t1a	8.58E+02	8.83E+02	9.93E+02	6.45E+02	9.91E+02	0.178	0.217	0.490	0.966	1.000
t2	6.41E+02	6.16E+02	5.05E+02	5.10E+02	8.57E+02	-0.200	0.239	0.508	0.944	1.000
r1	4.89E-01	4.86E-01	4.95E-01	1.60E-01	8.26E-01	0.026	0.604	0.614	0.982	0.992

Table S3: Model checking of the most likely scenarios. Deviations of summary statistics for the observed data from the posterior predictive distribution of the most probable scenarios. One sample summary statistics: NAL: mean number of alleles; VAR: mean allele size variance; MGW: mean ratio of the number of alleles to the range of allele sizes. Two sample summary statistics: N2P: mean number of alleles; V2P: mean allele size variance; FST: Wright's F_{ST}; LIK: mean individual assignment likelihoods, DAS: mean index of classification; DM2: mean shared allele distance. Significance denotes the location of the observed summary statistic in the 5%, 1% or 0.1% tails of the posterior predictive distribution (*, **, ***, respectively).

A. Postglacial expansion from APE to ALP			
summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	8.000	0.343	
NAL_1_2	9.684	0.362	
VAR_1_1	5.982	0.264	
VAR_1_2	6.547	0.328	
MGW_1_1	0.772	0.256	
MGW_1_2	0.827	0.191	
N2P_1_1&2	10.053	0.295	
V2P_1_1&2	6.402	0.295	
FST_1_1&2	0.013	0.107	
LIK_1_1&2	1.106	0.035	(*)
LIK_1_2&1	1.154	0.045	(*)
DAS_1_1&2	0.371	0.990	(**)
DM2_1_1&2	0.060	0.006	(**)
B. Postglacial expansion from BAL to CAR			
summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	8.421	0.558	
NAL_1_2	9.579	0.498	
VAR_1_1	6.821	0.446	
VAR_1_2	6.561	0.419	
MGW_1_1	0.800	0.365	
MGW_1_2	0.839	0.270	
N2P_1_1&2	10.263	0.439	
V2P_1_1&2	6.647	0.404	
FST_1_1&2	0.005	0.000	(***)
LIK_1_1&2	1.106	0.026	(*)
LIK_1_2&1	1.168	0.038	(*)
DAS_1_1&2	0.375	0.984	(*)
C. Postglacial expansion from WES to BRI			
summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	8.105	0.098	
NAL_1_2	9.053	0.192	
VAR_1_1	6.353	0.173	
VAR_1_2	6.764	0.208	
MGW_1_1	0.798	0.288	
MGW_1_2	0.823	0.294	
N2P_1_1&2	9.211	0.073	
V2P_1_1&2	6.611	0.187	
FST_1_1&2	0.002	0.000	(***)
LIK_1_1&2	1.083	0.001	(***)
LIK_1_2&1	1.135	0.003	(**)

DAS_1_1&2	0.366	1.000	(***)

D. Postglacial expansion from IBE to WES and HER

summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	5.368	0.400	
NAL_1_2	9.632	0.578	
NAL_1_3	9.053	0.569	
VAR_1_1	5.359	0.598	
VAR_1_2	6.342	0.674	
VAR_1_3	6.764	0.708	
MGW_1_1	0.644	0.067	
MGW_1_2	0.838	0.119	
MGW_1_3	0.823	0.120	
N2P_1_1&2	9.790	0.562	
N2P_1_1&3	9.158	0.508	
N2P_1_2&3	10.000	0.406	
V2P_1_1&2	6.270	0.664	
V2P_1_1&3	6.606	0.694	
V2P_1_2&3	6.518	0.686	
FST_1_1&2	0.019	0.058	
FST_1_1&3	0.011	0.071	
FST_1_2&3	0.000	0.000	(***)
LIK_1_1&2	1.125	0.341	
LIK_1_1&3	1.112	0.303	
LIK_1_2&1	1.318	0.384	
LIK_1_2&3	1.101	0.256	
LIK_1_3&1	1.348	0.437	
LIK_1_3&2	1.114	0.279	
DAS_1_1&2	0.370	0.837	
DAS_1_1&3	0.366	0.816	
DAS_1_2&3	0.369	0.835	

E. Admixture from CAR and WES in HER

summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	9.632	0.094	
NAL_1_2	9.579	0.075	
NAL_1_3	9.053	0.310	
VAR_1_1	6.342	0.161	
VAR_1_2	6.561	0.181	
VAR_1_3	6.764	0.236	
MGW_1_1	0.838	0.095	
MGW_1_2	0.839	0.094	
MGW_1_3	0.823	0.379	
N2P_1_1&2	10.368	0.048	(*)
N2P_1_1&3	10.000	0.073	
N2P_1_2&3	10.105	0.066	
V2P_1_1&2	6.472	0.164	
V2P_1_1&3	6.518	0.184	
V2P_1_2&3	6.698	0.189	
FST_1_1&2	0.003	0.007	(**)
FST_1_1&3	0.000	0.000	(***)
FST_1_2&3	0.005	0.000	(***)
LIK_1_1&2	1.114	0.006	(**)
LIK_1_1&3	1.102	0.000	(***)
LIK_1_2&1	1.127	0.003	(**)
LIK_1_2&3	1.139	0.000	(***)

LIK_1_3&1	1.114	0.015	(*)
LIK_1_3&2	1.143	0.008	(**)
DAS_1_1&2	0.370	1.000	(***)
DAS_1_1&3	0.369	0.998	(**)
DAS_1_2&3	0.366	0.999	(***)
DM2_1_1&2	0.159	0.172	
DM2_1_1&3	0.040	0.000	(***)
DM2_1_2&3	0.217	0.044	(*)

F. Postglacial expansion from CAR to NEA and HER

summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	9.632	0.169	
NAL_1_2	9.579	0.309	
NAL_1_3	9.684	0.200	
VAR_1_1	6.342	0.142	
VAR_1_2	6.561	0.173	
VAR_1_3	6.778	0.178	
MGW_1_1	0.838	0.248	
MGW_1_2	0.839	0.442	
MGW_1_3	0.823	0.174	
N2P_1_1&2	10.368	0.138	
N2P_1_1&3	10.526	0.127	
N2P_1_2&3	10.632	0.191	
V2P_1_1&2	6.472	0.146	
V2P_1_1&3	6.571	0.151	
V2P_1_2&3	6.688	0.173	
FST_1_1&2	0.003	0.000	(***)
FST_1_1&3	0.004	0.026	(*)
FST_1_2&3	0.003	0.000	(***)
LIK_1_1&2	1.114	0.002	(**)
LIK_1_1&3	1.118	0.003	(**)
LIK_1_2&1	1.127	0.006	(**)
LIK_1_2&3	1.131	0.006	(**)
LIK_1_3&1	1.124	0.005	(**)
LIK_1_3&2	1.131	0.004	(**)
DAS_1_1&2	0.370	1.000	(***)
DAS_1_1&3	0.370	1.000	(***)
DAS_1_2&3	0.370	1.000	(***)

G. Admixture from WES and NEA in SCA

summary statistics	observed value	proportion (simulated<observed)	
NAL_1_1	9.895	0.426	
NAL_1_2	9.684	0.411	
NAL_1_3	9.053	0.410	
VAR_1_1	8.426	0.759	
VAR_1_2	6.778	0.670	
VAR_1_3	6.764	0.668	
MGW_1_1	0.540	0.000	(***)
MGW_1_2	0.823	0.080	
MGW_1_3	0.823	0.135	
N2P_1_1&2	11.000	0.399	
N2P_1_1&3	10.421	0.336	
N2P_1_2&3	10.368	0.318	
V2P_1_1&2	7.625	0.712	
V2P_1_1&3	7.716	0.724	
V2P_1_2&3	6.802	0.660	

FST_1_1&2	0.002	0.003	(**)
FST_1_1&3	0.007	0.173	
FST_1_2&3	0.004	0.001	(***)
LIK_1_1&2	1.122	0.320	
LIK_1_1&3	1.141	0.317	
LIK_1_2&1	1.128	0.358	
LIK_1_2&3	1.139	0.243	
LIK_1_3&1	1.151	0.414	
LIK_1_3&2	1.143	0.276	
DAS_1_1&2	0.368	0.725	
DAS_1_1&3	0.361	0.705	
DAS_1_2&3	0.366	0.750	

7. LIST OF PUBLICATIONS

Papers in scientific journals with impact factors

Havrdová A, Douda J, Doudová J (submitted, *Flora*) Local topography affects seed bank successional patterns in alluvial meadows.

Mandák B, **Havrdová A**, Krak K, Hadincová V, Vít P, Zákravský P, Douda J (submitted, *New Phytologist*) Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation.

Mandák B, Vít P, Krak K, Trávníček P, **Havrdová A**, Hadincová V, Zákravský P, Jarolímová V, Bacles CFE, Douda J (in press) Putative glacial refugia inferred from the geographic distribution of Alnus glutinosa cytotypes in Europe. *Annals of Botany*.

Douda J, Boublík K, Slezák M, Biurrun I, Nociar J, **Havrdová A**, Doudová J, Aćić S, Brisse H, Brunet J, Chytry M, Claessens H, Csiky J, Didukh Y, Dimopoulos P, Dullinger S, FitzPatrick Ú, Guisan A, Horchler P, Hrivnák R, Jandt U, Kącki Z, Kevey B, Landucci F, Lecomte H, Lenoir J, Paal J, Paternoster D, Pauli H, Pielech R, Rodwell J, Roelandt B, Svenning J-Ch, Šibík J, Šilc U, Škvorc Ž, Tsiripidis I, Tzonev R, Wohlgemuth T, Zimmermann N (in press) Vegetation classification and biogeography of European floodplain forests and alder carrs. *Applied Vegetation Science*.

Havrdová A, Douda J, Krak K, Vít P, Hadincová V, Zákravský P, Mandák B (in press) Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture. *Molecular Ecology*.

Drašnarová A, Krak K, Vít P, Doudová J, Douda J, Hadincová V, Zákravský P., Mandák B (2014) Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genetics & Genomes*, **10**, 865–873.

Douda J, Doudová J, **Drašnarová A**, Kuneš P, Hadincová V, Krak K, Zákravský P, Mandák B (2014) Migration Patterns of Subgenus *Alnus* in Europe since the Last Glacial Maximum: A Systematic Review. *PloS one*, **9**, e88709.

Douda J, Doudová-Kochánková J, Boublík K, **Drašnarová A** (2012) Plant species coexistence at local scale in temperate swamp forest: test of habitat heterogeneity hypothesis. *Oecologia*, **169**, 523–534.

Papers in other scientific journals

Douda J, **Havrdová A**, Mandák B (in press) Co nám říkají molekulární data o glaciálních refugiích středoevropských dřevin? Zprávy české botanické společnosti.