Charles University in Prague

Faculty of Science

Department of Botany

Ph.D. Thesis

Influence of climatic fluctuations in the Neogene and Quaternary on evolution of ecologically diverse plant genus: an example of *Hippophae* L. (Elaeagnaceae)

M.Sc. Dongrui Jia

Supervised by Dr. Igor V. Bartish

Prague, 2013

Statement of originality

Hereby I declare that I performed my dissertation work on my own and that I cited all used sources of information and literatures. Neither this work nor any considerable part of it has not been submitted for obtaining any other or the same academic title.

In Prague, October, 2013

Signature: Dongrui Jia

Acknowledgement

With the completion of this thesis, my journey for obtaining a doctorate degree arrived its destination. Along the way, many people provided guidance, inspiration and support to me. I would like to take this opportunity to thank all of them, including those whose names are not mentioned.

First and foremost, I wish to express my deepest appreciation to my supervisors Dr. Jianquan Liu and Dr. Igor V. Bartish. Dr. Jianquan Liu accepted me as his Master student in 2006 and thereby opened for me the door to the academic world of Evolutionary Biology. His ideas and guidance have been of great help throughout the process from the initial stage, all the way through to the end. And, without his support and encouragement, I could not start my PhD study in Prague, where I spent a pleasant and fruitful time with Dr. Igor V. Bartish. I am very thankful to Igor for all his contributions of time, ideas in past three years. He was always willing to transfer his knowledge and share his research experience. His attention to every detail and academic precision provided me the necessary direction and focus in my future research career.

I would like to sincerely acknowledge the China Scholarship Council (CSC), Natural Science Foundation of China and Fund of Alexey Kudrin of Russia for providing financial assistance. I am also grateful to the Institution of Botany of the Academy of Sciences of the Czech Republic (ASCR) for providing excellent research facilities.

I am extremely indebted to Dr. Jan Kirschner, Dr. Miroslav Vosatka and Ms. Magdalena Doksanska of the Institution of Botany ASCR and Ms. Magdalena Curikova

of the Office of Education and Student Affairs at Faculty of Science of Charles University for their continuous help and efforts in facilitating my stay in Prague. My thanks also due to Mr. Zhenkang Zhang and Ms. Yun Tang of the Department of Education of the Chinese Embassy in the Czech Republic for their care and concern.

I owe gratitude to Ms. Katerina Wolfova, Ms. Maria Surinova, Ms. Eva Nicova and Dr. Karol Krak for their help in lab. I would like to expand my thanks to all my friends and colleagues for their regards and companionship.

My heartfelt gratitude goes to my parents, and little brother for their continuous love, support and encouragement. Last but not the least, I would like to express my indebtedness towards my wife, Tengliang Liu, for her love, sacrifice and unconditional support. With her, the journey is no longer lonely. I dedicate this thesis to her.

Abstract

The Neogene geologic processes and climatic changes had tremendous impact on evolution of biota in different regions of Northern Hemisphere (NH). The Qinghai-Tibetan Plateau (QTP) was a central part of these processes. Migrations from the QTP to other temperate regions represent one of the main biogeographic patterns for Northern Hemisphere. However, this 'out-of-QTP' hypothesis has never been tested through a phylogeographic analysis of a widely distributed species and the ages and routs of these migrations are largely not resolved. On the other hand, climate change played an important role in shaping the amount and structure of intraspecific genetic diversity, which provide the main basic substrate for any evolutionary change. Therefore, a detailed understanding of the effects of historic climate alterations on intraspecific genetic diversity can provide valuable insights into the evolutionary consequences of past climate changes and predicting the likely direction of global warming effects on sustainability of extant populations and species. In this thesis, I first studied the phylogeography of *Hippophae rhamnoides* to test the 'out-of-QTP' hypothesis (Chapter II). Then, I performed phylogenetic, dating and biogeographic analyses of the genus *Hippophae* (Chapter III). Finally, I studied the correlation between genetic diversity and changes in climatic niche suitability of *H. rhamnoides* ssp. *sinensis* (Chapter IV). The phylogeographic analyses supported an 'out-of-QTP' hypothesis for *H. rhamnoides* followed by allopatric divergence, hybridization and introgression. The biogeographic analyses of *Hippophae* highlighted the impact of different stages in uplift of the QTP and Eurasian mountains and climatic changes in the Neogene on diversification and range shifts in highland flora on the continent. Patterns of genetic diversity in *H. rhamnoides* ssp. *sinensis* suggested strong impact of historical climatic niche suitability and both extent and direction of changes in niche suitability throughout the Late Quaternary on current within-population genetic diversity in this pioneer plant species.

Content

Chapter I. Introduction

The distribution patterns and ranges of present flora are developed not only by their life histories and ecological traits (Hamrick *et al.* 1992), but also by historical events like climate oscillations (Comes & Kadereit 1998, 2003; Hewitt 2000; Hewitt 2004) as well as geological changes. Around the Northern Hemisphere, temperate plant species occurred discontinuous in eastern Asia, Europe, eastern North America, and/or western North America. Several explanations were put forward for this pattern of disjunction. One standard explanation is that one widespread vegetation type was followed by increased extinction owing to drying climates and/or glaciations (Manchester 1999; Milne & Abbott 2002; Tiffney 1985a, b, 2000; Tiffney & Manchester 2001; Wen 1999; Wolfe 1975). However, several recent botanical studies of Northern Hemisphere attempted to integrate references of the movement, the origination and even the directionality of movement with elements of the traditional view (see Manos & Donoghue 2001; Milne & Abbott 2002). Donoghue & Smith (2004) compiled a large sample of relevant phylogenies for 66 Northern Hemisphere disjunct plant genera consisting of 100 disjunct lineages to infer disjunction patterns, ancestral areas and directions of movement. Their results showed that out of the 39 lineages with directionality resolved, 31 disjunct lineages within these genera originated and diversified within East Asia, followed by movement out of Asia at different times. The importance of East Asia as a center of arcto-Tertiary species diversity and diversification was further supported statistically in the recent work of Harris *et al.* (2013), who reconstructed the stem lineage distributions for 185 endemic lineages from 23 disjunct genera representing

17 vascular plant families using Bayes-DIVA approach that handles phylogenetic uncertainty and uncertainty in ancestral area construction. The generation and retention of particularly broad diversity in East Asia is attributable to its topographic and climatic diversity and history (Guo *et al.* 1998; Harrison *et al.* 2001; Qian 2001, 2002; Qian & Ricklefs 1999, 2000). While within East Asia, the Qinghai-Tibetan Plateau (QTP) comprises major parts of three global biodiversity hotspots, i.e. 'Himalaya', 'Indo-Burma' and 'Mountains of Southwest China' (Mittermeier *et al.* 2005; Myers *et al.* 2000). Its southern and south-eastern fringes, namely the Himalaya-Hengduan Mountains region, were further considered to be the largest 'evolutionary front' of the world's temperate zone (López-Pujol *et al.* 2011a, 2011b).

Uplift of the Qinghai-Tibetan Plateau

The Qinghai-Tibetan Plateau (QTP), situated in southwest China, is the largest and highest plateau in the world. It occupies an area of approximately 2.5×10^6 km^2 at an average elevation extending 4500 m above sea level (Zhang *et al.* 2002; Zheng 1996), embracing Tibet, Qinghai, south Xinjiang, southwest Gansu, west Sichuan and Yunnan provinces in China and extending into parts of Afghanistan, Pakistan, India, Nepal and Bhutan. Thus, it is reputed as "the Roof of the world" and "the Third Pole". Additionally, the QTP is also one of the youngest plateaus on the planet, uplifted in the process of the collision and post-collisional convergence of the Indian subcontinent with the Eurasian plate since about 50 million years ago (Ma) (Harrison *et al.* 1992; Molnar & Tapponnier 1975; Royden *et al.* 2008; Tapponnier *et al.* 2001). Following the slow uplift stage in the Late Cretaceous-Oligocene, further rapid uplift of the QTP occurred at 25–17 Ma in the

Early Miocene in southern Tibet and the central Himalaya (Harrison *et al.* 1992; Molnar *et al.* 1993; Shi *et al.* 1999; Zhang *et al.* 2013), which was termed the Himalayan Movement and characterized by the rise of mean altitude of the QTP to 2 000 m, the westward withdrawal of the Paratethys Sea, the central Asian aridity (Guo *et al.* 2002) and the formation of the 'paleo' Asian monsoons (An *et al.* 2001). Based on thermochronologic, sedimentologic, oceanographic and paleoclimatic studies, a major uplift event at 8 Ma that raised the southern Tibet to its present height was suggested (Clark *et al.* 2010; Harrison *et al.* 1992; Molnar *et al.* 1993; Wang *et al.* 2011c; Zhang *et al.* 2008; Zhang *et al.* 2013), while some other geologists and geographers advocated that close to current altitude of some portions of the southern Tibet has been existing from at least 14–15 Ma (Coleman & Hodges 1995; Spicer *et al.* 2003) or even older (~40 Ma; DeCelles *et al.* 2007; Rowley & Currie 2006; Wang *et al.* 2008b). During the Pliocene to Pleistocene, the plateau entered an intense compressional phase (Zhang *et al.* 2013) and became elevated to its maximum altitude as present (Fang *et al.* 2007; Li *et al.* 2001; Liu *et al.* 1998; Shi *et al.* 1999; Xu *et al.* 1973). Three stages of this most recent intensive uplifting of the QTP were further proposed: the Qingzang (Qinghai and Tibet) Movement at 3.6–1.7 Ma (including three phases commencing at 3.6, 2.5 and 1.7 Ma, respectively), the Kunhuang (Kunlun Mts. and Huanghe River, i.e. Yellow River) Movement at 1.1–0.6 Ma and the Gonghe Movement at 0.15 Ma to present (Li *et al.* 1996; Li *et al.* 2001; Shi *et al.* 1999). The orogenesis of this period changed the geomorphologic, hydrologic, sedimentologic and tectonic configurations on and around the QTP dramatically, and was associated with the formation of present like Asian monsoons, the enhanced aridity in Asian inland, the beginning of Chinese loess, the appearance and deflection of the

Yellow river, and the isolation of the Qinghai Lake (Li *et al.* 2001; Zheng *et al.* 2000).

Biodiversity of the Qinghai-Tibetan Plateau

Vegetational History and ecosystematic diversity of the Qinghai-Tibetan Plateau

Despite controversy remains regarding the exact timing and episodes as described above, the uplift of the QTP and Himalayas, as a fascinating geographic event in the Cenozoic, altered profoundly the geography and topography of the area and modified dramatically the climatic and ecological environment of the globe and especially East Asia. By analyzing sporopollen assemblage data and paleobotanical findings obtained from 75 localities on the QTP, Tang & Shen (1996) revealed drastic vegetation changes during the Late Cenozoic. In the Early Miocene, sclerophyllous evergreen broad-leaf *Quercus* forests were developed along with the newly-formed Mediterranean climate and Indian Ocean Monsoon, which replaced sclerophyllous evergreen broad-leaf *Eucalyptus* forests that dominated the plateau during the Late Cretaceous–Eocene interval. Subsequently, the broad-leaf and needle-leaf forests on the plateau showed a trend of general decrease (Wu *et al.* 2006). For example, temperate/warm-temperate broad-leaf forest of *Quercus*, *Ulmus* and *Betula* that dominated the northeast QTP region (Tianshui, Gansu province) between 17.1 and 14.7 Ma was replaced by forest or forest-steppe of *Ulmus*, *Artemisia* and *Betula* between 14.7 and 11.7 Ma, which in turn returned to a broad-leaf forest of *Betula* and *Quercus* during 11.7–8.5 Ma. After 8.5 Ma, the forest decreased rapidly and was replaced mostly by steppe vegetation mainly composed of *Artemisia*, Chenopodiaceae and Poaceae, corresponding to a permanent drying of the Asian interior

at this time (Hui *et al.* 2011). However, the Late Miocene vegetation in the QTP was still mainly characterized by forest, and can be divided into three zones from north to south: alpine evergreen needle-leaf forest and forest grassland in the northern slope of Kunlun Mts.; warm temperate (subtropical) deciduous broadleaved forest and shrubs in the central plateau and the south flank of the Kunlun Mts. and subtropical sclerophyllous evergreen broad-leaf forest mixed with cedar forest in the areas south of the Gangdise and Nyenchen Tanglha Mts. With a cooler climate in the Early and Middle Pliocene, shrub grassland appeared and occupied large areas of the former north zone; alpine evergreen needle-leaf forest shifted southwards and replaced deciduous broadleaved forest of the former central zone; while sclerophyllous evergreen broad-leaf forest mixed with cedar forest remained dominant in the south plateau. From the Late Pliocene on, decreased temperature and precipitation accelerated the replacement of forest by shrub grassland in the wide area of the north and central QTP. Meanwhile, forest vegetation was developed in the eastern edge of the QTP because of a comparatively temperate humid environment. Climatic oscillations during the late Quaternary led to quasi-periodical fluctuations of the QTP vegetation: alpine needle-leaf forests dominated in the warm interglacial periods, while desert grassland or shrub grassland dominated in the cold glacial periods. In the Late Holocene, grassland/meadow occupied the entire plateau, with desert grassland in the northwest part, alpine grassland in the central and west part, sub-alpine meadow and grassland in the northeast part and sub-alpine forest, shrub and meadow in the southeast part (Tang & Shen 1996). Currently, all of the large ecosystems of the terrestrial ecosystem, i.e. tundra, forest, scrub, desert and grassland, and aquatic formations, which can be fully displayed only on a continental scale, are concentrated on

the plateau (Li 1995). Four major plateau ecosystemic zones with distinct regional features have been proposed from southeast to northwest as follows: humid and semi-humid forest on the eastern and southern edges; plateau semi-humid scrub and meadow; plateau semi-arid, high-cold steppe and plateau arid high-cold desert, and each of these large ecosystems contains many mesoecosystems with lots of special plateau types (Li 1995; Zhang 1978).

Species diversity of the Qinghai-Tibetan Plateau

Diversified ecosystems, complicated and varied interfaces between ecosystems (Li 1995), complex landscapes and heterogeneous habitats made the QTP as one important biodiversity museum and cradle where ancient species conserved while young species evolved (López-Pujol *et al.* 2011a). The QTP harbors more than 13 000 vascular plant species, including 800 species in 124 genera and 43 families of Pteridophyta, 88 species in 18 genera and 7 families of Gymnospermae and more than 12 000 species in 1494 genera and 169 families of Angiospermae, which amount to about 40%, 44% and 44.4% of the relevant total species in China, respectively. The richness of fauna in the plateau is also very high: 206 species of mammals, 678 species of birds, 83 species of reptiles, 80 species of amphibians and 152 species of fishes account for 41.3%, 57.2% 22.1%, 28.7% and 5.4% of the entire species of China, respectively (Feng 1996; Feng & Li 1998; Li 1994; Wang 2000; Wu 1996). However, the distribution of species on the plateau is extremely uneven, with few species occurring on the vast inland plateau because of the harsh environment (Li 1995). It is estimated that the Qiangtang Plateau in northern Tibet,

occupying one-fourth of the total plateau area, possesses only one-tenth of the total species; while the Himalayan and the Hengduan mountain regions, covering less than one-fifth of the plateau, contain more than 80% of the species found on the plateau (Li 1995). On the other hand, the QTP boasts abundant endemic species: about 3500 vascular plant species and at least 29 genera, including about 100 endemic ferns (Miller 2003), 40 endemic mammals (60% of China's total), 28 endemic birds (about 30% of the total), two endemic reptiles and 10 endemic amphibians (Li 1995). Similarly, most endemic species are on the southern and eastern fringes of the plateau.

Biogeographic studies of plants on and around the Qinghai-Tibetan Plateau

As described above, the continuous uplifts of the QTP in the Cenozoic caused dramatic climatic and ecosystematic shifts, and configured complex landscapes and heterogeneous habitats, which made the QTP the most significant "evolutionary front" of China (López-Pujol *et al.* 2011b). Case studies with biogeographic inferences and/or age estimates from taxa with ranges on or around the QTP were summarized in Table 1-1. In most cases, age estimates were based on secondary calibrations and were therefore very approximate, and thus should be taken as preliminary until carefully selected and properly assigned into molecular phylogenies fossil records can be included into dating analyses. Nevertheless, significant correlations between different uplift stages of the plateau from the Miocene to the Quaternary and endemic species diversifications were revealed. Few studies use

Table 1-1 Case studies with biogeographic inferences and/or age estimates from taxa with ranges on or around the QTP

Taxon	Family	Distribution / Sampling	Growth form	Markers	Stem node age (Ma)	Crown node Age (Ma)	Ancestral Area	Biogeographic inferences	Reference
Cyananthus	Campanulaceae	S and SE QTP	Herb	**nDNA:** nrITS; **cpDNA:** *mat*K, *rbc*L, *psb*A-*trn*H, *trn*G-S, *atp*B, *trn*L-F	*15.12	9.30	S QTP	Migration to SE QTP 9.3 Ma; secondary dispersals from SE to S QTP	Zhou *et al.* 2013
Phyllolobium	Fabaceae	QTP and adjacent areas	Herb	**nDNA:** nrITS	*4.16	3.62	–	–	Zhang *et al.* 2012b
Lilium-Nomocharis complex	Liliaceae	N temperate	Herb	**nDNA:** nrITS; **cpDNA:** *mat*K	*15.96 (cpDNA)	13.70 (cpDNA)	NE QTP	Migrations to E. Asia and Siberia 12.6 and 7.3 Ma; to N America 6.9 Ma; to the Caucasus 7.5 Ma	Gao *et al.* 2013
Androsace	Primulaceae	N temperate; Arctic	Herb	**nDNA:** nrITS; **cpDNA:** *trn*L-F	~37	~35	S and E QTP	Migration from Himalayas (S QTP) to the Caucasus ~14 Ma	Roquet *et al.* 2013
Lepisorus	Polypodiaceae	Asia, Africa	Fern	**cpDNA:** *rbc*L, *rbc*L-*atp*B, *rps*4-*trn*S, *trn*L-F	18.9	17	SE QTP	Migrations to Japan 17.0 and 8.2Ma; to N China 13.8 Ma	Wang *et al.* 2012
Rheum	Polygonaceae	QTP, Xinjiang, C/W Asia, Europe	Herb	**cpDNA:** *rbc*L, *mat*K, *ndh*F, *acc*D, *psa*A, *rbc*L-*acc*D, *trn*K, *trn*L-F	*20.8	12.0	QTP/C Asia/W Asia/combination	Migrations from the QTP to C Asia 9.9 and 4.2 Ma	Sun *et al.* 2012
Gentiana sect. *Cruciata*	Gentianaceae	Eurasia	Herb	**cpDNA:** *atpB-rbcL*, *rpl*20-*rps*12, *trn*L-F, *trn*S-G	5.6–16.86	1.29–3.88	QTP	Dispersals out of QTP to central Asia and Europe	Zhang *et al.* 2009
Saussurea	Asteraceae	Eurasia	Herb	**nDNA:** nrITS; **cpDNA:** *trn*L-F, *psb*A-*trn*H	*8.91 (combined)	7.33 (combined)	QTP	Island-like adaptive radiation with dispersals to C Asia, E Asia and Europe	Wang *et al.* 2009b

Table 1-1 (*continued*)

Ligularia-Cremanthodium-Parasenecio complex	Asteraceae	QTP, C China, Japan	Herb	**nDNA:** nrITS; **cpDNA:** *ndh*F, *trn*L-F	–	19.51 (ITS)	QTP	Radiation with 16 dispersals to C China and Japan	Liu *et al.* 2006
Phymatopteris	Polypodiaceae	SE Asia to Himalayan region	Fern	**cpDNA:** *rbc*L, *trn*L-F, *rps*4, *rps*4-*trn*S	31.90	26.75	Malay archipelago	Migration to Himalayan region 26.75 Ma followed by subsequent diversification 21.22 Ma	Li *et al.* 2012c
Nannoglottis	Asteraceae	Endemic to QTP	Herb	**nDNA:** nrITS; **cpDNA:** *ndh*F, *trn*L-F	22.86–32	2.41–3.37	SE Asia	Migration to QTP and re-diversification in QTP	Liu *et al.* 2002
Solms-laubachia	Brassicaceae	C Asia, QTP	Herb	**nDNA:** *LEAFY*, *G3pdh*; **cpDNA:** *pet*N-*psb*M, *psb*M-*trn*D	*5.27	2.31	C Asia, W Himalayas	Eastward migration to SE QTP	Yue *et al.* 2009
Pseudodraba, *Baimashania*, *Sinoarabis*	Brassicaceae-Arabideae	C Asia, QTP	Herb	**nDNA:** nrITS, *chs*; **cpDNA:** *trn*L-F	–	7.85 (cpDNA)	E Irano-Turanian region	Eastward migration to SE QTP and further speciation during the Pliocene and Early Pleistocene	Koch *et al.* 2012
Doronicum	Asteraceae-Senecioneae	E, C and SW Asia, Europe, N Africa	Herb	**Morphology**; **nDNA:** nrITS1, nrITS2; **cpDNA:** *trn*L-F	–	–	S Europe/ Mediterra-nean basin	Early European diversification followed by eastward migration to C and E Asia	Fernández *et al.* 2001
Klasea	Asteraceae–Cardueae	Old world	Herb	**nDNA:** nrITS, nrETS	–	–	W Asia	Eastward dispersal to Tianshan-Pamir-Altai, followed by dispersals to E Asia, Himalayas, Europe and W Asia	Martins 2006
Angelica	Apiaceae-Selineae	Northern Hemisphere	Herb	**nDNA:** nrITS, nrETS; **cpDNA:** *rps*16, *rps*16-*trn*K, *rpl*32-*trn*L, *trn*L-T	~16	13.6	NE Asia	Dispersals to E/C China, C Asia, Europe, N America and E Himalayas; radiation in E Himalyas since 7.7 Ma	Liao *et al.* 2012

*Fossil calibrations were used for dating analysis. Only mean values for time estimates were shown.

ancestral area reconstructions based on phylogenetic tests and likelihood analyses. Generally, three biogeographic patterns could be illustrated. First, genera endemic to the QTP and adjacent areas diversified during a period from the Miocene to the Quaternary, probably correlated with the different uplift stages of the plateau (e.g. Zhang *et al.* 2012b; Zhou *et al.* 2013). Second, some genera with a wide distribution originated in the QTP and adjacent areas and migrated to East Asia (e.g. Gao *et al.* 2013; Liu *et al.* 2006; Wang *et al.* 2012; Wang *et al.* 2009b), North America (e.g. Gao *et al.* 2013), Central Asia (e.g. Sun *et al.* 2012; Wang *et al.* 2009b; Zhang *et al.* 2009), the Caucasus and Western Europe (e.g. Gao *et al.* 2013; Roquet *et al.* 2013; Wang *et al.* 2009b; Zhang *et al.* 2009). Third, some genera originated in other regions of the world, such as the Northeast Asia, the Southeast Asia, the Central Asia, the Irano-Turanian region and the Mediterranean region, and diversified greatly after their ancestors reached the QTP (e.g. Fernández *et al.* 2001; Koch *et al.* 2012; Li *et al.* 2012c; Liao *et al.* 2012; Liu *et al.* 2002; Martins 2006; Yue *et al.* 2009).

Phylogeographic studies of plants in the Qinghai-Tibetan Plateau and adjacent areas

The most recent violent tectonic uplifts of the QTP during the Kunhuang Movement (1.1–0.6 Ma) initiated the formation of mountain glaciers, ice caps and valley glaciers in high regions (Shi 2002; Zhang *et al.* 2000; Zheng *et al.* 2002; Zheng *et al.* 1998). During the Quaternary, the QTP experienced four major glaciations (Zheng *et al.* 2002) with Naynayxungla Glaciation (reaching its maximum ca. 0.8–0.5 Ma) as the largest, when an ice sheet covered an area five to seven times larger than it does today and glaciers may have been retained in the high mountains of the central QTP regions even during the

interglacial warm stages (Ehlers & Gibbard 2007; Owen *et al.* 2005; Shi & Ren 1990; Wu *et al.* 2001; Zheng *et al.* 2002). Afterwards, the glaciations had become less extensive. During the Last Glacial Maximum (LGM) only limited glacier advances existed in the QTP areas (Lehmkuhl & Owen 2005). It is generally believed that a massive unified ice sheet covering the whole QTP has never existed during the Quaternary (Owen *et al.* 2008; Seong *et al.* 2008; Shi 2002; Shi *et al.* 1998).

Repeated Quaternary climatic oscillations certainly affected the range dynamics of organisms (Avise 2000), and have left genetic signatures in current populations (Hewitt 2000; Hewitt 2004). Since Zhang *et al.* (2005) provided the first study on *Juniperus przewalskii* (Cupressaceae), a growing body of molecular phylogeographic studies has begun to elucidate how plant species, including trees, shrubs, herbs and ferns, that occur in the QTP and adjacent regions shifted their distribution ranges in response to the Quaternary climatic oscillations by using mainly chloroplast DNA (cpDNA) sequence variation (Table 1-2). Remarkable differences in the demographic history of the QTP plant species were indicated. Some species may have retreated to the glacial refugia located in the eastern and southeastern edges of the QTP, and either recolonized the QTP platform (e.g. Chen *et al.* 2008b; Cun & Wang 2010; Duan *et al.* 2011; Li *et al.* 2011b; Meng *et al.* 2007; Wu *et al.* 2010; Yang *et al.* 2008; Zhang *et al.* 2005; Zhang *et al.* 2010; Zhao *et al.* 2013) and other regions outside of the QTP (e.g. Chen *et al.* 2012; Fan *et al.* 2013; Wang *et al.* 2011b; Zhang *et al.* 2010; Zhao *et al.* 2013), or experienced local expansions (e.g. Gao *et al.* 2012; Li *et al.* 2011a; Liu *et al.* 2013; Wang *et al.* 2011a; Zhang *et al.* 2011a; Zou *et al.* 2012) or even remained stable (e.g. Li *et al.* 2012a; Qiu *et*

Table 1-2 Phylogeographic patterns of different plants in the QTP and adjacent areas

Species	Family	Distribution/ Sampling	Markers	Crown node age (Ma)	Refugia	Expansion direction	Expansion age (Kya)	Reference
Herbs								
Bupleurum smithii	Apiaceae	QTP, N China	**cpDNA:** *trn*H-*psb*A, *trn*L-F	–	NE, E QTP edge; micro-refugia on the platform	From edge to the platform; northward to N China	–	Zhao *et al.* 2013
Hippuris vulgaris	Hippuridaceae	QTP, Xinjiang	**cpDNA:** *ycf*6-*psb*M, *trn*T-L, *rps*16, *atp*I-H	0.48	No specific refugia	–	120; 170	Chen *et al.* 2013
Cyananthus delavayi	Campanulaceae	SE QTP	**cpDNA:** *trn*H-*psb*A, *psb*D-*trn*T; **nDNA:** *ncp*GS	0.86	Multiple refugia throughout the range	No	–	Li *et al.* 2012a
Meconopsis integrifolia	Maximowicz	QTP	**cpDNA:** *trn*fM-S, *trn*L-F; **nDNA:** nrITS	5.77 cpDNA); 5.33 (nrITS)	–	–	–	Yang *et al.* 2012a
Ligularia vellerea	Asteraceae	SE QTP	**cpDNA:** *trn*H-*psb*A, *trn*L-*rpl*32	1.34	–	–	710	Yang *et al.* 2012b
Primula ovalifolia	Primulaceae	SE QTP, C China	**cpDNA:** *trn*T-L, *trn*L-F, *rps*l6	2.44	Multiple refugia throughout its range	No	–	Xie *et al.* 2012
Incarvillea sinensis	Bignoniaceae	E/SE QTP, N/NE China	**cpDNA:** *trn*H-*psb*A, *trn*S-fM	4.40	Not specified	North-eastward colonization to NE China	–	Chen *et al.* 2012

Table 1-2 (*continued*)

Sinopodophyllum hexandrum	Berberidaceae	S/SE/E QTP	**AFLPs**; **cpDNA:** *atp*H-I, *rps*18-*clpp*, *rpl*32-*trn*L	0.48–0.37	E Himalayas; SE QTP	Local expansions	–	Li *et al.* 2011a
Psammosilene tunicoides	Caryophyllaceae	SE QTP	**cpDNA:** *rpl*16, *trn*Q-*rps*16; **nDNA:** *GPA*1	–	SE QTP	Local expansions	–	Zhang *et al.* 2011a
Swertia tetraptera	Gentianaceae	E/C QTP	**cpDNA:** *trn*L-F	–	C platform; E plateau edge	Local expansions	–	Yang *et al.* 2011
Stellera chamaejasme	Thymelaeaceae	QTP, C/N China	**cpDNA:** *trn*T-L, *trn*L-F, *rpl*16	*2.11	SE QTP	Onto the platform and northwards to N China	–	Zhang *et al.* 2010
Allium przewalskianum	Alliaceae	QTP	**cpDNA:** *rpl*16, *trn*L-F, *acc*D-*psa*I, *trn*H-*psb*A, *trn*S-G; **nDNA:** nrITS	–	E QTP edge	Westward expansions to the platform	5–150	Wu *et al.* 2010
Aconitum gymnandrum	Ranunculaceae	QTP	**cpDNA:** *rpl*20-*rps*12, *trnV*, *psbA*-*trnH*; **nDNA:** nrITS	1.45– 3.65	E QTP edge; QTP platform	–	–	Wang *et al.* 2009a
Dysosma versipellis	Berberidaceae	SE QTP, C/E China	**cpDNA:** *trn*L-F, *trn*L-*ndh*J, *trn*S-fM	1.48	SE QTP; C China	Stable in SE QTP; expansions in C/E China	327	Qiu *et al.* 2009
Rhodiola alsia	Froderstrom	QTP	**cpDNA:** *rbc*L	–	Multiple refugia on the platform; SE QTP edge	Local expansions	–	Gao *et al.* 2009

Table 1-2 (*continued*)

Metagentiana striata	Gentianaceae	E QTP	**cpDNA:** *trn*S-G	–	SE edge	Onto the platform	–	Chen *et al.* 2008b
Pedicularis longiflora	Orobanchaceae	QTP	**cpDNA:** *trn*T-F	0.67	SE QTP	Onto the platform	120–17	Yang *et al.* 2008
Shrubs								
Sophora davidii	Leguminosae	SE QTP, C/N China	**cpDNA:** *psb*A-*trn*H, *rpl*32-*trn*L; **nDNA:** *ncp*GS	1.28	SE QTP	Eastwards out of QTP	179–22	Fan *et al.* 2013
Reaumuria soongarica	Tamaricaceae	N QTP, NW China	**cpDNA:** *trn*S-G	2.03–6.33	Outside of the QTP	To the N QTP	37.4–112.4	Li *et al.* 2012b
Spiraea alpina	Rosaceae	QTP	**cpDNA:** *trn*L-F	–	Multiple refugia on the platform	Local expansions	–	Zhang *et al.* 2012a
Hippophae tibetana	Elaeagnaceae	QTP	**cpDNA:** *trn*L-F, *trn*S-G; **nDNA:** nrITS	1.08–3.25 (cpDNA); 1.26–3.17 (nrITS)	Multiple refugia	Local expansion in E region	26.4–79.1	Jia *et al.* 2011
Hippophae tibetana	Elaeagnaceae	QTP	**cpDNA:** *trn*T-F	3.15	Multiple micro-refugia throughout the range	Local expansions	–	Wang *et al.* 2010
Sibiraea angustata	Rosaceae	QTP	**cpDNA:** *trn*S-G	–	SE edge	Onto the platform	–	Duan *et al.* 2011

Table 1-2 (*continued*)

Potentilla fruticosa	Rosaceae	QTP	**cpDNA:** *trn*S-G, *rpl*20-*rps*12	–	C QTP	North-eastwards to the edge	52–25; 5.1–2.5	Li *et al.* 2010
Potentilla fruticosa	Rosaceae	QTP	**cpDNA:** *trn*T-L	–	SE, W, C QTP	Radical expansions across the QTP	30–325	Sun *et al.* 2010
Potentilla fruticosa	Rosaceae	QTP	**cpDNA:** *mat*K	–	C QTP	North-eastwards to the edge	–	Shimono *et al.* 2010
Juniperus sabina	Cupressaceae	Xinjiang, NE QTP, N/NE China	**cpDNA:** *trn*L-F, *trn*S-G, *trn*D-T	0.78	N China; Xinjiang	Eastwards to the NE China and westward to the NE QTP	–	Guo *et al.* 2010
Potentilla glabra	Rosaceae	QTP	**cpDNA:** *trn*T-L		SE QTP edge; QTP platform	Northwards to NE QTP	316; 201	Wang *et al.* 2009c
Ostryopsis davidiana	Betulaceae	E QTP, N China	**cpDNA:** *trn*L-F, *psb*A-*trn*H	–	Multiple refugia throughout its range	-	–	Tian *et al.* 2009
Forests								
Picea likiangensis	Pinaceae	E QTP	**mtDNA:** *nad*1, *nad5*; **cpDNA:** *trn*L-F, *trn*S-G, *ndh*K/C	–	Multiple refugia thoughout its range	Local expansion along the distributional edge	–	Zou *et al.* 2012
Pinus densata	Pinaceae	SE QTP	**mtDNA:** *nad*1, *nad*4, *nad5*; **5 cpSSRs**		Multiple refugia throughout its range	Westward range expansions	–	Wang *et al.* 2011a

Table 1-2 (*continued*)

Pinus densata	Pinaceae	SE QTP	**nDNA:** 8 loci	6.6	Multiple refugia throughout its range	Westward range expansions	–	Gao *et al.* 2012
Juniperus przewalskii	Cupressaceae	N/E QTP	**cpDNA:** *trn*T-F, *trn*S-G	–	E plateau edge	Onto QTP platform	–	Zhang *et al.* 2005
Juniperus przewalskii	Cupressaceae	N/E QTP	**nDNA:** 8 loci	–	E plateau edge	Onto QTP platform	–	Li *et al.* 2011b
Juniperus tibetica group	Cupressaceae	QTP	**cpDNA:** *trn*T-L, *trn*L-F, *trn*L	–	Microrefugia throughout the range	-	–	Opgenoorth *et al.* 2010
Tsuga dumosa	Pinaceae	SE/S QTP	**mtDNA:** *nad*5, *cox*1; **cpDNA:** *trn*S-fM, *atp*H-I; **nDNA:** *LEAFY*	–	SE QTP	Westwards to S QTP	818	Cun & Wang 2010
Picea crassifolia	Pinaceae	NE QTP, adjacent Helan/Daqing Mts.	**mtDNA:** *nad*1, *nad*5; **cpDNA:** *trn*C-D	–	E plateau edge; Helan/Daqing Mts.	From E edge onto the platform	–	Meng *et al.* 2007
Taxus wallichiana	Taxaceae	S/SE QTP	**cpDNA:** *trn*L-F; **8 SSRs**	4.2	SE QTP	Local expansions	60.4; 85.8	Liu *et al.* 2013
Fern								
Lepisorus clathratus	Polypodiacae	S/SE QTP, N China, Altai	**cpDNA:** *rps*4-*trn*S, *trn*L-F	1.5	SE, S QTP,	Range expansions into N to C China	–	Wang *et al.* 2011b

*Fossil calibrations were used for dating analysis. Only mean values for time estimates were shown.

al. 2009; Xie *et al.* 2012) during the interglacial phases or at the end of the LGM. Alternatively, some species may probably have survived through the Quaternary in multiple microrefugia at high-altitude and experienced expansions of different extents (e.g. Gao *et al.* 2009; Jia *et al.* 2011; Li *et al.* 2010; Opgenoorth *et al.* 2010; Shimono *et al.* 2010; Sun *et al.* 2010; Wang *et al.* 2010; Wang *et al.* 2009c; Wang *et al.* 2009a; Yang *et al.* 2011; Zhang *et al.* 2012a; Zhao *et al.* 2013). Additionally, there are also cases that plant colonized the northern QTP area from their refugia outside of the plateau (e.g. Guo *et al.* 2010; Li *et al.* 2012b). On the other hand, climatic oscillations together with extensive orogenic processes since the Pliocene may have led to intraspecific divergences within species with long origin history (e.g. Fan *et al.* 2013; Gao *et al.* 2012; Jia *et al.* 2011; Li *et al.* 2012b; Liu *et al.* 2013; Qiu *et al.* 2009; Wang *et al.* 2010; Wang *et al.* 2011b; Wang *et al.* 2009a; Xie *et al.* 2012; Yang *et al.* 2012a; Yang *et al.* 2012b; Zhang *et al.* 2010), followed by hybridizations and introgressions between intraspecific lineages (e.g. Wang *et al.* 2009a; Zou *et al.* 2012) when they expanded to form hybrid zones after the end of the glacial ages or during the interglacial or retreated into a single refugium (Liu *et al.* 2012). Finally, it is worthy to mention that except for the study on *Stellera chamaejasme* that used fossil calibrations (Zhang *et al.* 2010), time estimates in the other studies listed in Table 1-2 were mainly based on mutation rates, and thus should be taken with extreme caution.

General information of Elaeagnaceae Juss.

Elaeagnaceae is a family of shrubs or trees often with thorns and distinctive silvery or brownish peltate scales and/or stellate hairs. The family consists of three genera

(*Elaeagnus* L., *Hippophae* L. and *Shepherdia* Nutt.) and ca. 90 species (Qin & Gilbert 2007). It is mostly deciduous and has a typical Northern Hemisphere distribution, but in Southeast Asia extends to Malesia and Northeast Queensland. *Elaeagnus* occurs throughout the range, *Hippophae* occurs throughout Eurasia, while *Shepherdia* occurs only in North America (Bartish & Swenson 2003). The basic chromosome number in *Hippophae* is x = 12 (Rousi 1965), and x = 14 in *Elaeagnus* (Arohonka & Rousi 1980). Two species of *Shepherdia* have basic chrome numbers x = 11 (*S. canadensis*) and x = 13 (*S. argentea*) (Arohonka & Rousi 1980, and references therein). *Hippophae* is exclusively anemophylous, while *Elaeagnus* and *Shepherdia* often have fragrant flowers and are entomophilous. *Hippophae* and *Shepherdia* are dioecious, whereas *Elaeagnus* is hermaphrodite or androdioecious or gynodioecious. Elaeagnaceae was placed into order Rosales by APG III (2009), and its close relationships with Barbeyaceae, Dirachmaceae and Rhamnaceae were indicated by recent phylogenetic analyses based on sequence data (Wang *et al.* 2009d; Zhang *et al.* 2011b). Within the family, sequence analysis of the cpDNA *trn*L-F and *rbc*L revealed a strongly supported clade consisting of *Shepherdia* and *Elaeagnus*, to which *Hippophae* was placed sister (Richardson *et al.* 2000). However, another phylogenetic analysis of cpDNA sequences indicated a close relationship between *Shepherdia* and *Hippophae*, but without providing support values (Richardson *et al.* 2004). Therefore, the phylogenetic relationships among the three genera should be considered as not resolved.

General information of *Hippophae* L.

The genus *Hippophae* has a wide but fragmented distribution raging from northern

Europe through central Europe and Central Asia to China. All species are diploid (2n = 24; Rousi 1965), dioecious, and wind-pollinated. They reproduce vegetatively with root and sexually with bird-dispersed seeds (Bartish *et al.* 2000). Roots of *Hippophae* form nitrogen-fixing nodules (Bond *et al.* 1956) and posses an efficient dual symbiosis with mycorrhiza and *Frankia* (Tian *et al.* 2002), which enables *Hippophae* to colonize infertile and bare soil after disturbance. Thus, its natural habitats include seashores, river deltas, valley slopes and meadows at altitudes from sea level to 5 000 meters above sea level (Lian *et al.* 1998; Rousi 1971).

Systematics of *Hippophae* L.

The systematic treatment of *Hippophae* L. had been in dispute for many years (Table 1-3). Servettaz (1908) recognized one species *H. rhamnoides* L. with three subspecies, which were raised to species level (i.e. *H. rhamnoides* L., *H. tibetana* Schlecht and *H. salicifolia* D. Don) by Rousi (1971). However, Avdeyev (1983) put forth arguments against the use of strongly variable and mostly continuous characters and inappropriate statistical methods by Rousi (1971), and proposed one single species *H. rhamnoides* on the basis of one brachyblast character (monocarpic or dicarpic) and several quantitative and continuous characters. Based on fruit and seed characteristics, Lian (1988) divided the genus into two sections: sect. *Hippophae* Lian with carpodermis separated from seed coat and seed surface smooth and shining, and sect. *Gyantsensis* Lian with carpodermis adhering to seed coat and seed surface rough and dull. Further, Lian & Chen (1993)

Table 1-3 Overview of systematic treatments of *Hippophae* L. (modified from Bartish *et al.* 2002)

Servettaz (1908)	Rousi (1971)	Avdeyev (1983)	Lian & Chen (1993)	Hyvönen (1996)	Lian *et al.* (1998)	Swenson & Bartish (2002)
H. rhamnoides	*H. rhamnoides*	*H. rhamnoides*	**Sect. *Hippophae***	*H. rhamnoides*	**Sect. *Hippophae***	*H. rhamnoides*
ssp. *rhamnoides*	ssp. *carpatica*	ssp. *rhamnoides*	*H. rhamnoides*	ssp. *carpatica*	*H. rhamnoides*	ssp. *carpatica*
ssp. *salicifolia*	ssp. *caucasica*	ssp. *salicifolia*	ssp. *carpatica*	ssp. *caucasica*	ssp. *carpatica*	ssp. *caucasica*
ssp. *tibetana*	ssp. *fluviatilis*		ssp. *caucasica*	ssp. *fluviatilis*	ssp. *caucasica*	ssp. *fluviatilis*
	ssp. *gyantsensis*		ssp. *fluviatilis*	ssp. gyantsensis	ssp. *fluviatilis*	ssp. *mongolica*
	ssp. *mongolica*		ssp. *mongolica*	ssp. *neurocarpa*	ssp. *mongolica*	ssp. *rhamnoides*
	ssp. *rhamnoides*		ssp. *rhamnoides*	ssp. *mongolica*	ssp. *rhamnoides*	ssp. *sinensis*
	ssp. *sinensis*		ssp. *turkestanica*	ssp. *rhamnoides*	ssp. *sinensis*	ssp. *turkestanica*
	ssp. *turkestanica*		*H. salicifolia*	ssp. *tibetana*	ssp. *turkestanica*	ssp. *yunnanensis*
	ssp. *yunnanensis*		*H. sinensis*	ssp. *turkestanica*	ssp. *yunnanensis*	*H. salicifolia*
	H. salicifolia		ssp. *sinensis*	*H. salicifolia*	*H. salicifolia*	*H. goniocarpa*
	H. tibetana		ssp. *yunnanensis*	ssp. *salicifolia*	**Sect. *Gyantsensis***	*H. litangensis*
			Sect. *Gyantsensis*	ssp. *sinensis*	*H. goniocarpa*	*H. gyantsensis*
			H. goniocarpa	ssp. *yunnanensis*	ssp. *goniocarpa*	*H. neurocarpa*
			H. gyantsensis		ssp. *litangensis*	ssp. *neurocarpa*
			H. neurocarpa		*H. gyantsensis*	ssp. *stellatopilosa*
			H. tibetana		*H. neurocarpa*	*H. tibetana*
					ssp. *neurocarpa*	
					ssp. *stellatopilosa*	
					H. tibetana	

suggested treating *H. rhamnoides* ssp. *sinensis* and *H. rhamnoides* ssp. *yunnanensis* of Rousi (1971) as two subspecies of *H. sinensis*. After conducting a cladistic analysis on morphological characters, Hyvönen (1996) found two distinct lineages within the genus and interpreted them as two species, *H. rhamnoides* and *H. salicifolia*. Lian *et al.* (1998) recognized six species in two sections and described new subspecies. However, based on parsimony analyses of all 15 taxa in *Hippophae* recognized by Lian *et al.* (1998) using cpDNA and a combined data set of morphological characters and cpDNA, Bartish *et al.* (2002) refrained from recognizing any sections within the genus due to weak internal support, and accepted the two subspecies of *H. goniocarpa*, for which a hybrid origin were suggested (Bartish *et al.* 2000; Lian *et al.* 1995), as independent taxa: *H. goniocarpa* and *H. litangensis* (Swenson & Bartish 2002). Almost simultaneously, Tzvelev (2002) suggested taxonomic changes within *H. rhamnoides* that several subspecies should be raised to species level: *H. sinensis* (Rousi) Tzvel. comb. et stat. nov., *H. yunnanensis* (Rousi) Tzvel. comb. et stat. nov., *H. caucasica* (Rousi) Tzvel. comb. et stat. nov., *H. mongolica* (Rousi) Tzvel. comb. et stat. nov. and *H. turkestanica* (Rousi) Tzvel. comb. et stat. nov. However, he suggested no formal changes to spp. *carpatica* and *fluviatilis*, although considered the latter one might be raised to species level.

Molecular studies of *Hippophae* L.

Phylogenetic studies

Yao & Tigerstedt (1993) were probably the first who applied molecular markers to study the genetic diversity and phylogeny in *Hippophae*. They performed electrophoretic

analysis of isozymes in 25 populations representing *H. tibetana*, *H. neurocarpa* and three subspecies of *H. rhamnoides* (ssp. *rhamnoides*, *sinensis* and *turkestanica*). The phylogenetic tree justified the division of subspecies under *H. rhamnoides* and specific status of *H. neurocarpa* and *H. tibetana*. RAPD analysis of wild populations representing all taxa, according to the systematic classification suggested by Lian *et al.* (1998), strongly supported the systematic treatment of *Hippophae* developed by Lian *et al.* (1998) (Bartish *et al.* 2000). Additionally, the results also provided further evidence for the hybridization theory of Lian & Chen (1993) and Lian *et al.* (1995) that *H. goniocarpa* ssp. *goniocarpa* and *litangensis* have originated from interspecific hybridization between *H. neurocarpa* ssp. *neurocarpa* and *H. rhamnoides* ssp. *sinensis*, and between *H. neurocarpa* ssp. *stellatopilosa* and *H. rhamnoides* ssp. *yunnanensis*, respectively (see also Sun *et al.* 2003; Wang *et al.* 2008a). Both parsimony analyses of cpDNA RFLP data and a combined data set of cpDNA and morphological characters strongly supported monophyly of *Hippophae* (Bartish *et al.* 2002). All recognized subspecies of *H. rhamnoides* formed a monophyletic group that was distinguished by a single molecular synapomorphy (Bartish *et al.* 2002). Sun *et al.* (2002) analyzed the phylogenetic relationships among 15 taxa of the genus by comparing sequences of the internal transcribed spacer (ITS) region of nuclear ribosomal DNA. In agreement with Bartish *et al.* (2002), the monophyly of *Hippophae* was supported by high bootstrap value. *H. tibetana* was placed basal to two highly supported clades, of which one consisted of all subspecies of *H. rhamnoides* (Sun *et al.* 2002).

Assessment of genetic diversity

Information on genetic variation is of great significance not only for evolutionary aspects of speciation but also for applied research, for instance, conservation and breeding. Yao & Tigerstedt (1993) found that a large amount of genetic diversity within *H. rhamnoides* (ssp. *rhamnoides*, *sinensis* and *turkestanica*) was allocated within populations and between subspecies while not among the populations of the same subspecies, indicating that there were a relatively restricted population differentiation while significant differences existed among subspecies. RAPD analysis on *Hippophae rhamnoides* ssp. *rhamnoides* indicated that, compared with plants with different breeding systems, genetic diversity of this taxon was considerably lower than the tree species and most of the perennial, outcrossing herbaceous species, but higher than annual or short-lived, selfing species (Bartish *et al.* 1999). Beyond that, several studies have been conducted on groups of populations of *Hippophae* to assess genetic diversity and genetic structure using different markers, for example, on *H. rhamnoides* ssp. *sinensis* using ISSR (Tian *et al.* 2004a; Tian *et al.* 2004b; Wang *et al.* 2011d), cpSSR (Wang *et al.* 2011d), PCR-RFLP (Chen *et al.* 2010b) and RAPD (Sheng *et al.* 2006; Sun *et al.* 2006); on *H. rhamnoides* ssp. *yunnanensis* using RAPD (Chen *et al.* 2010a) and ISSR (Chen *et al.* 2008a; Tian *et al.* 2004a); on *H. rhamnoides* ssp. *turkestanica* using AFLP (Raina *et al.* 2012; Shah *et al.* 2009) and RAPD (Singh *et al.* 2006); on *H. rhamnoides* ssp. *carpatica* using RAPD (Alexandra *et al.* 2012; Bartish *et al.* 2000); on *H. rhamnoides* ssp. *caucasica* and *fluviatilis* using RAPD (Bartish *et al.* 2000); on *H. salicifolia* using AFLP (Raina *et al.* 2012) and on *H. tibetana* using AFLP (Raina *et al.* 2012). However, because of limited

and incomplete sampling across the whole range of each taxon, these studies should be considered as preliminary.

Phylogeographic studies

Bartish *et al.* (2006) performed the first phylogeographic study within the genus *Hippophae*. They analyzed populations of *H. rhamnoides* (ssp. *caucasica*, *carpatica*, *fluviatilis* and *rhamnoides*) from Europe/Asia Minor for both sequence variation at the nuclear chalcone synthase intron (*Chsi*) and cpDNA restriction fragment length polymorphism (PCR-RFLP) variation. This study supported the hypothesis of postglacial recolonization of central and northern Europe from a south-eastern refugium (Gams 1943), and indicated a hybrid suture zone north of Alps due to the secondary contact of previously isolated populations from different refugia, which nearly synchronously expanded from Alps and Carpathians in the Late Quaternary. Genetic structure of populations from Asia Minor revealed demographic trends, which were different from European, indicating variable evolutionary responses to global climatic processes in different regions. Taking an advantage of the detailed fossil record of *H. rhamnoides* in Europe, these authors estimated approximate ages of differentiation between several main genetic lineages of *Chs*i within western part of its range. Meng *et al.* (2008) examined cpDNA *trn*L-F variation in 14 populations of *H. neurocarpa* and recovered eight haplotypes totally with one shared between ssp. *neurocarpa* and *stellatopilosa*, indicating discordance between genealogy and morphology in this fragment. In addition, private haplotypes were found in the high altitude, which suggested the species might have

survived the glacial periods in multiple refugia located on the plateau, from which contiguous range expansions occurred. Cheng *et al.* (2009) investigated the phylogeographic relationships between *H. gyantsensis* and *H. rhamnoides* ssp. *yunnanensis* by sequencing cpDNA *trn*L-F and *trn*S-G fragments for 109 trees of 14 populations from the two taxa. The study identified 11 haplotypes with two shared between taxa. Phylogenetic and Nested Clade analyses further revealed that the divergence of the haplotypes was highly inconsistent with morphological differentiation, suggesting complex maternal lineage sorting between the two taxa. The authors thus argued against the previous taxonomic treatment of *H. gyantsensis*. Multiple refugia were also suggested for both taxa across their distribution ranges during the LGM. Wang *et al.* (2010) applied cpDNA *trn*T-F sequence variation to examine the effects of the Quaternary glaciations and the QTP uplift on the phylogeographic structure of *H. tibetana*. The recovered 50 haplotypes clustered into three lineages, distributed in the western, central and eastern geographical range, respectively. Based on a mutation rate of 4.87×10^{-10} substitutions per site per year for *trn*T-F region, the three lineages was estimated to diverge at 3.15 Ma, corresponding to the most recent rapid and abrupt uplift of the QTP in the late Pliocene. A total of 33 (66%) private haplotypes occurring in single populations spread over the whole range of the species and originated from multiple differentiations in many lineages during a long period of more than 1 million years, indicating that these haplotypes went through the LGM and possibly some other Quaternary glaciations in their respective locations as refugia. Because some private haplotypes located at very high sites, the authors also advocated these microrefugia as the highest of all known refugia so far. Similar results were reported in another

phylogeographic study on *H. tibetana* using both cpDNA (*trn*L-F and *trn*S-G) and nrITS sequence variation, wherein only two distinct lineages were recovered due to missing samples from the central portion of its range (Jia *et al.* 2011).

Biogeographic hypotheses for *Hippophae* L.

Bobrov (1962) was probably the first to suggest east to west migrations along mountain ranges for this species. He also suggested the Late Miocene as the most likely period of this migration (based only on paleofloristic data). Hyvönen (1996) conducted the first phylogenetic analysis in *Hippophae* using morphological characters. Contrary to Bobrov (1962), he suggested western part of the genus' range as its most likely ancestral area, and an ancient fragmentation of a wide Eurasian range in the genus. Based exclusively on palynological data, Gams (1943) suggested re-colonization of Scandinavia by *H. rhamnoides* in the Holocene from a glacial refugium in south-eastern Europe.

References

Alexandra SG, Ecaterina T, Nicoleta C*, et al.* (2012) The assessment of the genetic diversity of sea buckthorn populations from Romania using RAPD markers. *Romanian Biotechnological Letters* **17**, 7749-7756.

An Z, Kutzbach JE, Prell WL, Porter SC (2001) Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature* **411**, 62-66.

APG III (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**, 105-121.

Arohonka T, Rousi A (1980) Karyotypes and C-bands in *Shepherdia* and *Elaeagnus*. *Annales Botanici Fennici* **17**, 258-263.

Avdeyev VI (1983) Novaya taxonomiya roda oblepikha: *Hippophae* L. *Izvestiya Akademii Nauk*

Tadzhikskoy SSR, Otdeleniye Biologicheskih Nauk **93**, 11-17.

Avise JC (2000) *Phylogeography: The History and Formation of Species* Harvard University Press, London.

Bartish, Jeppsson, Nybom (1999) Population genetic structure in the dioecious pioneer plant species *Hippophae rhamnoides* investigated by random amplified polymorphic DNA (RAPD) markers. *Molecular Ecology* **8**, 791-802.

Bartish IV, Jeppsson N, Bartish GI, Lu R, Nybom H (2000) Inter- and intraspecific genetic variation in *Hippophae* (Elaeagnaceae) investigated by RAPD markers. *Plant Systematics and Evolution* **225**, 85-101.

Bartish IV, Jeppsson N, Nybom H, Swenson U (2002) Phylogeny of *Hippophae* (Elaeagnaceae) inferred from parsimony analysis of chloroplast DNA and morphology. *Systematic Botany* **27**, 41-54.

Bartish IV, Kadereit JW, Comes HP (2006) Late Quaternary history of *Hippophae rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (*Chsi*) sequences and chloroplast DNA variation. *Molecular Ecology* **15**, 4065-4083.

Bartish IV, Swenson U (2003) Elaeagnaceae. In: *The Families and Genera of Vascular Plants (Vol. VI)* (ed. Kubitzki K), pp. 131-134. Springer-Verlag, Berlin, Heidelberg, New York.

Bobrov EG (1962) Obzor roda *Myricaria* Desv. i jego istorija. *Botanicheskij Zhurnal* **52**, 924-936.

Bond G, MacConnell J, McCallum A (1956) The nitrogen-nutrition of *Hippophaë rhamnoides* L. *Annals of Botany* **20**, 501-512.

Chen G, Wang Y, Zhao C, Korpelainen H, Li C (2008a) Genetic diversity of *Hippophae rhamnoides* populations at varying altitudes in the Wolong natural reserve of China as revealed by ISSR markers. *Silvae Genetica* **57**, 29-36.

Chen JM, Du ZY, Sun SS, Gituru RW, Wang QF (2013) Chloroplast DNA phylogeography reveals repeated range expansion in a widespread aquatic herb *Hippuris vulgaris* in the Qinghai-Tibetan Plateau and adjacent areas. *Plos One* **8**.

Chen LH, Yu Z, Jin HP (2010b) Comparison of ribosomal DNA ITS regions among *Hippophae rhamnoides* ssp. *sinensis* from different geographic areas in China. *Plant Molecular Biology Reporter* **28**, 635-645.

Chen ST, Xing YW, Su T*, et al.* (2012) Phylogeographic analysis reveals significant spatial genetic structure of *Incarvillea sinensis* as a product of mountain building. *Bmc Plant Biology* **12**, 58.

Chen SY, Wu G, Zhang DJ*, et al.* (2008b) Potential refugium on the Qinghai-Tibet Plateau revealed by the chloroplast DNA phylogeography of the alpine species *Metagentiana striata* (Gentianaceae). *Botanical Journal of the Linnean Society* **157**, 125-140.

Chen W, Su X, Zhang H*, et al.* (2010a) High genetic differentiation of *Hippophae rhamnoides* ssp. *yunnanensis* (Elaeagnaceae), a plant endemic to the Qinghai-Tibet Plateau. *Biochemical*

Genetics **48**, 565-576.

Cheng K, Sun K, Wen HY*, et al.* (2009) Maternal divergence and phylogeographical relationships between *Hippophae gyantsensis* and *H. rhamnoides* subsp. *yunnanensis*. *Journal of Plant Ecology (Chinese Version)* **33**, 1-11.

Clark MK, Farley KA, Zheng D, Wang Z, Duvall AR (2010) Early Cenozoic faulting of the northern Tibetan Plateau margin from apatite (U–Th)/He ages. *Earth and Planetary Science Letters* **296**, 78-88.

Coleman M, Hodges K (1995) Evidence for Tibetan plateau uplift before 14 Myr ago from a new minimum age for east-west extension. *Nature* **374**, 49-52.

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

Comes HP, Kadereit JW (2003) Spatial and temporal patterns in the evolution of the flora of the European Alpine System. *Taxon* **52**, 451-462.

Cun YZ, Wang XQ (2010) Plant recolonization in the Himalaya from the southeastern Qinghai-Tibetan Plateau: Geographical isolation contributed to high population differentiation. *Molecular Phylogenetics and Evolution* **56**, 972-982.

DeCelles PG, Kapp P, Ding L, Gehrels G (2007) Late Cretaceous to middle Tertiary basin evolution in the central Tibetan Plateau: Changing environments in response to tectonic partitioning, aridification, and regional elevation gain. *Geological Society of America Bulletin* **119**, 654-680.

Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1633-1644.

Duan Y, Gao Q, Zhang F*, et al.* (2011) Phylogeographic analysis of the endemic species *Sibiraea angustata* reveals a marginal refugium in the Qinghai-Tibet Plateau. *Nordic Journal of Botany* **29**, 615-624.

Ehlers J, Gibbard PL (2007) The extent and chronology of Cenozoic global glaciation. *Quaternary International* **164**, 6-20.

Fan DM, Yue JP, Nie ZL*, et al.* (2013) Phylogeography of *Sophora davidii* (Leguminosae) across the "Tanaka-Kaiyong Line', an important phytogeographic boundary in Southwest China. *Molecular Ecology* **22**, 4270-4288.

Fang XM, Zhang WL, Meng QQ*, et al.* (2007) High-resolution magnetostratigraphy of the Neogene Huaitoutala section in the eastern Qaidam Basin on the NE Tibetan Plateau, Qinghai Province, China and its implication on tectonic uplift of the NE Tibetan Plateau. *Earth and Planetary Science Letters* **258**, 293-306.

Feng ZJ (1996) The fauna characteristics and formation, evolution of terrestrial vertebrates. In:

Formation and Evolution of Qinghai-Xizang Plateau, pp. 209-222. Shanghai Science & Technology Press, Shanghai.

Feng ZJ, Li BS (1998) The biodiversity of the high-cold region of the Qinghai-Xizang Plateau. In: *Research Report on National Conditions of the Chinese Biodiversity*, pp. 82-89. China Environmental Science Press, Beijing.

Fernández IÁ, Aguilar JF, Panero JL, Feliner GN (2001) A phylogenetic analysis of *Doronicum* (Asteraceae, Senecioneae) based on morphological, nuclear ribosomal (ITS), and chloroplast (*trn*L-F) evidence. *Molecular Phylogenetics and Evolution* **20**, 41-64.

Gams H (1943) Der Sanddorn (*Hippophae rhamnoides* L.) im Alpengebiet. *Beihefte zum Botanischen Centralblatt, Abteilung B* **2**, 68-96.

Gao J, Wang B, Mao I-F*, et al.* (2012) Demography and speciation history of the homoploid hybrid pine *Pinus densata* on the Tibetan Plateau. *Molecular Ecology* **21**, 4811-4827.

Gao QB, Zhang DJ, Chen SY*, et al.* (2009) Chloroplast DNA phylogeography of *Rhodiola alsia* (Crassulaceae) in the Qinghai-Tibet Plateau. *Botany-Botanique* **87**, 1077-1088.

Gao YD, Harris AJ, Zhou SD, He XJ (2013) Evolutionary events in *Lilium* (including *Nomocharis*, Liliaceae) are temporally correlated with orogenies of the Q-T plateau and the Hengduan Mountains. *Molecular Phylogenetics and Evolution* **68**, 443-460.

Guo Q, Ricklefs RE, Cody ML (1998) Vascular plant diversity in eastern Asia and North America: historical and ecological explanations. *Botanical Journal of the Linnean Society* **128**, 123-136.

Guo YP, Zhang R, Chen CY, Zhou DW, Liu JQ (2010) Allopatric divergence and regional range expansion of *Juniperus sabina* in China. *Journal of Systematics and Evolution* **48**, 153-160.

Guo Z, Ruddiman WF, Hao Q*, et al.* (2002) Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature* **416**, 159-163.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**, 95-124.

Harris AJ, Wen J, Xiang Q-Y (2013) Inferring the biogeographic origins of inter-continental disjunct endemics using a Bayes-DIVA approach. *Journal of Systematics and Evolution* **51**, 117-133.

Harrison S, Yu G, Takahara H, Prentice I (2001) Palaeovegetation (Communications arising): Diversity of temperate plants in East Asia. *Nature* **413**, 129-130.

Harrison TM, Copeland P, Kidd W, Yin A (1992) Raising tibet. *science* **255**, 1663-1670.

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

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

Hui ZC, Li JJ, Xu QH*, et al.* (2011) Miocene vegetation and climatic changes reconstructed from a sporopollen record of the Tianshui Basin, NE Tibetan Plateau. *Palaeogeography, Palaeoclimatology, Palaeoecology* **308**, 373-382.

Hyvönen J (1996) On phylogeny of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **16**, 51-62.

Jia DR, Liu TL, Wang LY, Zhou DW, Liu JQ (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* **102**, 37-50.

Koch MA, Karl R, German DA, Al-Shehbaz IA (2012) Systematics, taxonomy and biogeography of three new Asian genera of Brassicaceae tribe Arabideae: An ancient distribution circle around the Asian high mountains. *Taxon* **61**, 955-969.

Lehmkuhl F, Owen LA (2005) Late Quaternary glaciation of Tibet and the bordering mountains: a review. *Boreas* **34**, 87-100.

Li BS (1994) The characteristics and conservation of biodiversity of the Qinghai-Tibetan Plateau. In: *East Asia Covered with Green Color*, pp. 635-661. China Environmental Science Press, Beijing.

Li BS (1995) *Biodiversity of the Qinghai-Tibetan Plateau and its conservation* International Centre for Integrated Mountain Development (ICIMOD), Kathmandu, Nepal.

Li C, Shimono A, Shen H, Tang Y (2010) Phylogeography of *Potentilla fruticosa*, an alpine shrub on the Qinghai-Tibetan Plateau. *Journal of Plant Ecology* **3**, 9-15.

Li CX, Lu SG, Ma JY*, et al.* (2012c) From the Himalayan region or the Malay Archipelago: Molecular dating to trace the origin of a fern genus *Phymatopteris* (Polypodiaceae). *Chinese Science Bulletin* **57**, 4569-4577.

Li GD, Yue LL, Sun H, Qian ZG (2012a) Phylogeography of *Cyananthus delavayi* (Campanulaceae) in Hengduan Mountains inferred from variation in nuclear and chloroplast DNA sequences. *Journal of Systematics and Evolution* **50**, 305-315.

Li JJ, Fang XM, Ma HZ*, et al.* (1996) Geomorphological and environmental evolution in the upper reaches of the Yellow River during the late Cenozoic. *Science in China Series D(Earth Sciences)* **39**, 380-390.

Li JJ, Fang XM, Pan BT, Zhao ZJ, Song YG (2001) Late Cenozoic intensive uplift of Qinghai-Xizang Plateau and its impacts on environments in surrounding area. *Quaternary Sciences* **21**, 381-391.

Li Y, Zhai SN, Qiu YX*, et al.* (2011a) Glacial survival east and west of the 'Mekong-Salween Divide' in the Himalaya-Hengduan Mountains region as revealed by AFLPs and cpDNA sequence variation in *Sinopodophyllum hexandrum* (Berberidaceae). *Molecular Phylogenetics and Evolution* **59**, 412-424.

Li ZH, Chen J, Zhao GF*, et al.* (2012b) Response of a desert shrub to past geological and climatic change: A phylogeographic study of *Reaumuria soongarica* (Tamaricaceae) in western China. *Journal of Systematics and Evolution* **50**, 351-361.

Li ZH, Zhang QA, Liu JQ, Kallman T, Lascoux M (2011b) The Pleistocene demography of an alpine

juniper of the Qinghai-Tibetan Plateau: *tabula rasa*, cryptic refugia or something else? *Journal of Biogeography* **38**, 31-43.

Lian Y (1988) New discoveries of the genus *Hippophae* L.(Elaeagnaceae). *Acta Phytotaxonomica Sinica* **26**, 235-237.

Lian YS, Chen XL (1993) Study on the germplasm resource of the genus *Hippophae* L. In: *International Symposium on Sea Buckthorn (Hippophae rhamnoides L.)*, pp. 157-161, Novosibirsk, Russia.

Lian YS, Chen XL, Lian H (1998) Systematic classification of the genus *Hippophae* L. *Seabuckthorn Research* **1**, 13-23.

Lian YS, Chen XL, Sun K (1995) New discoveries of the genus *Hippophae* L. In: *Proceedings of International Workshop on Seabuckthorn*, pp. 60-66. China Science and Technology Press, Beijing.

Liao C-Y, Downie SR, Yu Y, He X-J (2012) Historical biogeography of the *Angelica* group (Apiaceae tribe Selineae) inferred from analyses of nrDNA and cpDNA sequences. *Journal of Systematics and Evolution* **50**, 206-217.

Liu J-Q, Gao T-G, Chen Z-D, Lu A-M (2002) Molecular phylogeny and biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics and Evolution* **23**, 307-325.

Liu J-Q, Sun Y-S, Ge X-J, Gao L-M, Qiu Y-X (2012) Phylogeographic studies of plants in China: Advances in the past and directions in the future. *Journal of Systematics and Evolution* **50**, 267-275.

Liu J-Q, Wang Y-J, Wang A-L, Hideaki O, Abbott RJ (2006) Radiation and diversification within the *Ligularia–Cremanthodium–Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. *Molecular Phylogenetics and Evolution* **38**, 31-49.

Liu J, Moller M, Provan J*, et al.* (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist* **199**, 1093-1108.

Liu TS, Zheng MP, Guo ZT (1998) Initiation and evolution of the Asian monsoon system timely coupled with the ice-sheet growth and the tectonic movements in Asia. *Quaternary Sciences* **18**, 194-204.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011a) Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**, 1267-1280.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011b) Mountains of Southern China as "Plant Museums" and "Plant Cradles": Evolutionary and Conservation Insights. *Mountain Research and Development* **31**, 261-269.

Manchester SR (1999) Biogeographical relationships of North American tertiary floras. *Annals of the Missouri Botanical Garden* **86**, 472-522.

Manos PS, Donoghue MJ (2001) Progress in northern hemisphere phytogeography: an introduction. *International Journal of Plant Sciences* **162**, S1-S2.

Martins L (2006) Systematics and biogeography of *Klasea* (Asteraceae–Cardueae) and a synopsis of the genus. *Botanical Journal of the Linnean Society* **152**, 435-464.

Meng L-H, Yang H-L, Wu G-L, Wang Y-J (2008) Phylogeography of *Hippophae neurocarpa* (Elaeagnaceae) inferred from the chloroplast DNA *trn*L-F sequence variation. *Journal of Systematics and Evolution* **46**, 32-40.

Meng LH, Yang R, Abbott RJ*, et al.* (2007) Mitochondrial and chloroplast phylogeography of *Picea crassifolia* Kom. (Pinaceae) in the Qinghai-Tibetan Plateau and adjacent highlands. *Molecular Ecology* **16**, 4128-4137.

Miller D (2003) Tibet environmental analysis. In: *Background paper in preparation for USAID's program USAID Bureau for Asia and Near East*. U.S. Agency for International Development (USAID) Website.

Milne RI, Abbott RJ (2002) The origin and evolution of Tertiary relict floras. *Advances in Botanical Research* **38**, 281-314.

Mittermeier RA, Gil PR, Hoffman M*, et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions* Cemex, Conservation International and Agrupación Sierra Madre, Monterrey, Mexico.

Molnar P, England P, Martinod J (1993) Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon. *Reviews of Geophysics* **31**, 357-396.

Molnar P, Tapponnier P (1975) Cenozoic tectonics of Asia: Effects of a continental collision. *science* **189**, 419-426.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**, 853-858.

Opgenoorth L, Vendramin GG, Mao K*, et al.* (2010) Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. *New Phytologist* **186**, 780-780.

Owen LA, Caffee MW, Finkel RC, Seong YB (2008) Quaternary glaciation of the Himalayan–Tibetan orogen. *Journal of Quaternary Science* **23**, 513-531.

Owen LA, Finkel RC, Barnard PL*, et al.* (2005) Climatic and topographic controls on the style and timing of late Quaternary glaciation throughout Tibet and the Himalaya defined by ^{10}Be cosmogenic radionuclide surface exposure dating. *Quaternary Science Reviews* **24**, 1391-1411.

Qian H (2001) A comparison of generic endemism of vascular plants between East Asia and North America. *International Journal of Plant Sciences* **162**, 191-199.

Qian H (2002) A comparison of the taxonomic richness of temperate plants in East Asia and North America. *American Journal of Botany* **89**, 1818-1825.

Qian H, Ricklefs RE (1999) A comparison of the taxonomic richness of vascular plants in China and the United States. *The American Naturalist* **154**, 160-181.

Qian H, Ricklefs RE (2000) Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* **407**, 180-182.

Qin HN, Gilbert MG (2007) Elaeagnaceae. In: *Flora of China (Vol. 13)* (eds. Wu ZY, H. RP, Hong DY), pp. 251-273. Science Press, Beijing, China.

Qiu Y-X, Guan B-C, Fu C-X, Comes HP (2009) Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Molecular Phylogenetics and Evolution* **51**, 281-293.

Raina S, Jain S, Sehgal D*, et al.* (2012) Diversity and relationships of multipurpose seabuckthorn (*Hippophae* L.) germplasm from the Indian Himalayas as assessed by AFLP and SAMPL markers. *Genetic Resources and Crop Evolution* **59**, 1033-1053.

Richardson J, Chatrou L, Mols J, Erkens R, Pirie M (2004) Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1495-1508.

Richardson JE, Fay MF, Cronk QC, Bowman D, Chase MW (2000) A phylogenetic analysis of Rhamnaceae using *rbc*L and *trn*L-F plastid DNA sequences. *American Journal of Botany* **87**, 1309-1324.

Roquet C, Boucher FC, Thuiller W, Lavergne S (2013) Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *Journal of Biogeography* **40**, 1874-1886.

Rousi A (1965) Observation on the cytology and variation of European and Asiatic populations of *Hippophaë ramnoides*. *Annales Botanici Fennici* **2**, 1-18.

Rousi A (1971) The genus *Hippophae* L. A taxonomic study. *Annales Botanici Fennici* **8**, 177-227.

Rowley DB, Currie BS (2006) Palaeo-altimetry of the late Eocene to Miocene Lunpola basin, central Tibet. *Nature* **439**, 677-681.

Royden LH, Burchfiel BC, van der Hilst RD (2008) The geological evolution of the Tibetan Plateau. *science* **321**, 1054-1058.

Seong YB, Owen LA, Bishop MP*, et al.* (2008) Reply to comments by Matthias Kuhle on "Quaternary glacial history of the central Karakoram". *Quaternary Science Reviews* **27**, 1656-1658.

Servettaz C (1908) Monographie der Elaeagnaceae. *Beihefte zum Botanischen Centralblatt* **25**, 18.

Shah AH, Ahmad SD, Khaliq I*, et al.* (2009) Evaluation of phylogenetic relationship among seabuckthorn (*Hippophae rhamnoides* L. spp. *turkestanica*) wild ecotypes from Pakistan using amplified fragment length polymorphism (AFLP). *Pak J Bot* **41**, 2419-2426.

Sheng H, An L, Chen T, *et al.* (2006) Analysis of the genetic diversity and relationships among and within species of *Hippophae* (Elaeagnaceae) based on RAPD markers. *Plant Systematics and Evolution* **260**, 25-37.

Shi YF (2002) Characteristics of late Quaternary monsoonal glaciation on the Tibetan Plateau and in East Asia. *Quaternary International* **97**, 79-91.

Shi YF, Li JJ, Li BY (1998) *Uplift and Environmental Changes of Qinghai-Tibetan Plateau in the Late Cenozoic* Guangdong Science and Technology Press, Guangzhou.

Shi YF, Li JJ, Li BY, *et al.* (1999) Uplift of the Qinghai-Xizang (Tibetan) plateau and east Asia environmental change during late Cenozoic. *Acta Geographica Sinica* **54**, 10-21.

Shi YF, Ren JW (1990) Glacier recession and lake shrinkage indicating the climatic warming and drying trend in central Asia. *Acta Geographica Sinica* **45**, 1-13.

Shimono A, Ueno S, Gu S, *et al.* (2010) Range shifts of *Potentilla fruticosa* on the Qinghai-Tibetan Plateau during glacial and interglacial periods revealed by chloroplast DNA sequence variation. *Heredity* **104**, 534-542.

Singh R, Mishra SN, Dwivedi SK, Ahmed Z (2006) Genetic variation in seabuckthorn (*Hippophae rhamnoides* L.) populations of cold arid Ladakh (India) using RAPD markers. *Current Science* **91**, 1321-1322.

Spicer RA, Harris NB, Widdowson M, *et al.* (2003) Constant elevation of southern Tibet over the past 15 million years. *Nature* **421**, 622-624.

Sun K, Chen W, Ma R, *et al.* (2006) Genetic variation in *Hippophae rhamnoides* ssp. *sinensis* (Elaeagnaceae) revealed by RAPD markers. *Biochemical Genetics* **44**, 186-197.

Sun K, Chen X, Ma R, *et al.* (2002) Molecular phylogenetics of *Hippophae* L. (Elaeagnaceae) based on the internal transcribed spacer (ITS) sequences of nrDNA. *Plant Systematics and Evolution* **235**, 121-134.

Sun K, Ma R, Chen X, Li C, Ge S (2003) Hybrid origin of the diploid species *Hippophae goniocarpa* evidenced by the internal transcribed spacers (ITS) of nuclear rDNA. *Belgian Journal of Botany* **136**, 91-96.

Sun YS, Ikeda H, Wang YJ, Liu JQ (2010) Phylogeography of *Potentilla fruticosa* (Rosaceae) in the Qinghai-Tibetan Plateau revisited: a reappraisal and new insights. *Plant Ecology & Diversity* **3**, 249-257.

Sun YS, Wang AL, Wan DS, Wang Q, Liu JQ (2012) Rapid radiation of *Rheum* (Polygonaceae) and parallel evolution of morphological traits. *Molecular Phylogenetics and Evolution* **63**, 150-158.

Swenson U, Bartish IV (2002) Taxonomic synopsis of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **22**, 369-374.

Tang L, Shen C (1996) Late Cenozoic vegetational history and climatic characteristics of Qinghai-

Xizang Plateau. *Acta Micropalaeontologica Sinica* **13**, 321-337.

Tapponnier P, Zhiqin X, Roger F*, et al.* (2001) Oblique stepwise rise and growth of the Tibet Plateau. *science* **294**, 1671-1677.

Tian B, Liu RR, Wang LY*, et al.* (2009) Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. *Journal of Biogeography* **36**, 2148-2155.

Tian C, He X, Zhong Y, Chen J (2002) Effects of VA mycorrhizae and *Frankia* dual inoculation on growth and nitrogen fixation of *Hippophae tibetana*. *Forest ecology and management* **170**, 307-312.

Tian C, Nan P, Shi S, Chen J, Zhong Y (2004a) Molecular genetic variation in Chinese populations of three subspecies of *Hippophae rhamnoides*. *Biochemical Genetics* **42**, 259-267.

Tian CJ, Lei YD, Shi SH*, et al.* (2004b) Genetic diversity of sea buckthorn (*Hippophae rhamnoides*) populations in northeastern and northwestern China as revealed by ISSR markers. *New Forests* **27**, 229-237.

Tiffney BH (1985a) The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* **66**, 243-273.

Tiffney BH (1985b) Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. *Journal of the Arnold Arboretum* **66**, 73-94.

Tiffney BH (2000) Geographic and climatic influences on the Cretaceous and Tertiary history of Euramerican floristic similarity. *Acta- Universitatis Carolinae Geologica* **44**, 5-16.

Tiffney BH, Manchester SR (2001) The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* **162**, S3-S17.

Tzvelev NN (2002) On the genera *Elaeagnus* and *Hippophae* (Elaeagnaceae) in Russia and adjacent states. *Botanical journal (Ботанический журнал)* **87**, 74-86.

Wang A, Schluetz F, Liu J (2008a) Molecular evidence for double maternal origins of the diploid hybrid *Hippophae goniocarpa* (Elaeagnaceae). *Botanical Journal of the Linnean Society* **156**, 111-118.

Wang BS, Mao JF, Gao J, Zhao W, Wang XR (2011a) Colonization of the Tibetan Plateau by the homoploid hybrid pine *Pinus densata*. *Molecular Ecology* **20**, 3796-3811.

Wang C, Zhao X, Liu Z*, et al.* (2008b) Constraints on the early uplift history of the Tibetan Plateau. *Proceedings of the National Academy of Sciences* **105**, 4987-4992.

Wang G, Cao K, Zhang K*, et al.* (2011c) Spatio-temporal framework of tectonic uplift stages of the Tibetan Plateau in Cenozoic. *Science China Earth Sciences* **54**, 29-44.

Wang H, Moore MJ, Soltis PS*, et al.* (2009d) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences* **106**, 3853-3858.

Wang H, Qiong L, Sun K, *et al.* (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**, 2964-2979.

Wang JT (2000) Biodiversity: Biota and Biocoenose. In: *Mountain Geoecology and Sustainable Development of the Tibetan Plateau* (eds. Zheng D, Zhang QS, Wu SH), pp. 139-157. Springer Netherlands.

Wang L-Y, Ikeda H, Liu T-L, Wang Y-J, Liu J-Q (2009c) Repeated range expansion and glacial endurance of *Potentilla glabra* (Rosaceae) in the Qinghai-Tibetan Plateau. *Journal of Integrative Plant Biology* **51**, 698-706.

Wang L, Schneider H, Zhang XC, Xiang QP (2012) The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *Bmc Plant Biology* **12**, 210.

Wang L, Wu ZQ, Bystriakova N, *et al.* (2011b) Phylogeography of the Sino-Himalayan Fern *Lepisorus clathratus* on "The Roof of the World". *Plos One* **6**, e25896.

Wang LY, Abbott RJ, Zheng W, *et al.* (2009a) History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnandrum* (Ranunculaceae). *Molecular Ecology* **18**, 709-721.

Wang Y-J, Susanna A, Von Raab-Straube E, Milne R, Liu J-Q (2009b) Island-like radiation of *Saussurea* (Asteraceae: Cardueae) triggered by uplifts of the Qinghai–Tibetan Plateau. *Biological Journal of the Linnean Society* **97**, 893-903.

Wang Y, Jiang H, Peng S, Korpelainen H (2011d) Genetic structure in fragmented populations of *Hippophae rhamnoides* ssp. *sinensis* in China investigated by ISSR and cpSSR markers. *Plant Systematics and Evolution* **295**, 97-107.

Wen J (1999) Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology and Systematics* **30**, 421-455.

Wolfe JA (1975) Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* **62**, 264-279.

Wu L-L, Cui X-K, Milne RI, Sun Y-S, Liu J-Q (2010) Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**, 1691-1704.

Wu SG (1996) The characteristics, formation, and evolution of the flora. In: *Formation and Evolution of Qinghai-Xizang Plateau*, pp. 194-209. Shanghai Science & Technology Press, Shanghai.

Wu Y, Cui Z, Liu G, *et al.* (2001) Quaternary geomorphological evolution of the Kunlun Pass area and uplift of the Qinghai-Xizang (Tibet) Plateau. *Geomorphology* **36**, 203-216.

Wu ZH, Wu Z-H, Ye PS, Hu DG, Peng H (2006) Late Cenozoic environmental evolution of the Qinghai-Tibet Plateau as indicated by the evolution of sporopollen assemblages. *Geology in*

china **33**, 973-974.

Xie XF, Yan HF, Wang FY*, et al.* (2012) Chloroplast DNA phylogeography of *Primula ovalifolia* in central and adjacent southwestern China: Past gradual expansion and geographical isolation. *Journal of Systematics and Evolution* **50**, 284-294.

Xu J, Tao JR, Sun XJ (1973) On the discovery of a Quercus semicarpifolia bed in Mount Shisha Pangma and its significance in botany and geology. *Acta Botanica Sinica* **15**, 103-119.

Yang FS, Li YF, Ding X, Wang XQ (2008) Extensive population expansion of *Pedicularis longiflora* (Orobanchaceae) on the Qinghai-Tibetan Plateau and its correlation with the Quaternary climate change. *Molecular Ecology* **17**, 5135-5145.

Yang FS, Qin AL, Li YF, Wang XQ (2012a) Great genetic differentiation among populations of *Meconopsis integrifolia* and its implication for plant speciation in the Qinghai-Tibetan Plateau. *Plos One* **7**.

Yang LC, Zhou GY, Li CL, Song WZ, Chen GC (2011) Potential refugia in Qinghai-Tibetan Plateau revealed by the phylogeographical study of *Swertia tetraptera* (Gentianaceae). *Polish Journal of Ecology* **59**, 753-764.

Yang Z-Y, Yi T-S, Pan Y-Z, Gong X (2012b) Phylogeography of an alpine plant *Ligularia vellerea* (Asteraceae) in the Hengduan Mountains. *Journal of Systematics and Evolution* **50**, 316-324.

Yao Y, Tigerstedt PA (1993) Isozyme studies of genetic diversity and evolution in *Hippophae*. *Genetic Resources and Crop Evolution* **40**, 153-164.

Yue J-P, Sun H, Baum DA*, et al.* (2009) Molecular phylogeny of *Solms-laubachia* (Brassicaceae) s.l., based on multiple nuclear and plastid DNA sequences, and its biogeographic implications. *Journal of Systematics and Evolution* **47**, 402-415.

Zhang D, Fengquan L, Jianmin B (2000) Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. *Environmental Geology* **39**, 1352-1358.

Zhang FQ, Gao QB, Zhang DJ*, et al.* (2012a) Phylogeography of *Spiraea alpina* (Rosaceae) in the Qinghai-Tibetan Plateau inferred from chloroplast DNA sequence variations. *Journal of Systematics and Evolution* **50**, 276-283.

Zhang K, Wang G, Cao K*, et al.* (2008) Cenozoic sedimentary records and geochronological constraints of differential uplift of the Qinghai-Tibet Plateau. *Science in China Series D: Earth Sciences* **51**, 1658-1672.

Zhang K, Wang G, Xu Y*, et al.* (2013) Sedimentary Evolution of the Qinghai–Tibet Plateau in Cenozoic and its Response to the Uplift of the Plateau. *Acta Geologica Sinica - English Edition* **87**, 555-575.

Zhang ML, Kang Y, Zhong Y, Sanderson SC (2012b) Intense uplift of the Qinghai-Tibetan Plateau triggered rapid diversification of *Phyllolobium* (Leguminosae) in the Late Cenozoic. *Plant Ecology & Diversity* **5**, 491-499.

Zhang Q, Chiang TY, George M, Liu JQ, Abbott RJ (2005) Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. *Molecular Ecology* **14**, 3513-3524.

Zhang QY, Zhao YJ, Gong X (2011a) Genetic variation and phylogeography of *Psammosilene tunicoides* (Caryophyllaceae), a narrowly distributed and endemic species in south-western China. *Australian Journal of Botany* **59**, 450-459.

Zhang S-D, Soltis DE, Yang Y, Li D-Z, Yi T-S (2011b) Multi-gene analysis provides a well-supported phylogeny of Rosales. *Molecular Phylogenetics and Evolution* **60**, 21-28.

Zhang XL, Wang YJ, Ge XJ*, et al.* (2009) Molecular phylogeny and biogeography of *Gentiana* sect. *Cruciata* (Gentianaceae) based on four chloroplast DNA datasets. *Taxon* **58**, 862-870.

Zhang XS (1978) The plateau zonality of Tibetan vegetation. *Journal of Integrative Plant Biology* **02**, 140-149.

Zhang Y-H, Volis S, Sun H (2010) Chloroplast phylogeny and phylogeography of *Stellera chamaejasme* on the Qinghai-Tibet Plateau and in adjacent regions. *Molecular Phylogenetics and Evolution* **57**, 1162-1172.

Zhang Y, Li B, Zheng D (2002) A discussion on the boundary and area of the Tibetan Plateau in China. *Geographical Research* **21**, 1-8.

Zhao C, Ma XG, Liang QL, Wang CB, He XJ (2013) Phylogeography of an alpine plant (*Bupleurum smithii*, Apiaceae) endemic to the Qinghai-Tibetan Plateau and adjacent regions inferred from chloroplast DNA sequence variation. *Journal of Systematics and Evolution* **51**, 382-395.

Zheng BX, Xu QQ, Shen YP (2002) The relationship between climate change and Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and speculation. *Quaternary International* **97-98**, 93-101.

Zheng D (1996) The system of physico-geographical regions of the Qinghai-Xizang (Tibet) Plateau. *Science in China (series D)* **39**, 410-417.

Zheng H, Powell CM, An Z, Zhou J, Dong G (2000) Pliocene uplift of the northern Tibetan Plateau. *Geology* **28**, 715-718.

Zheng Z, Yuan B, Nicole P-M (1998) Paleoenvironments in China during the Last Glacial Maximum and the Holocene optimum. *Episodes* **21**, 152-158.

Zhou Z, Hong DY, Niu Y*, et al.* (2013) Phylogenetic and biogeographic analyses of the Sino-Himalayan endemic genus *Cyananthus* (Campanulaceae) and implications for the evolution of its sexual system. *Molecular Phylogenetics and Evolution* **68**, 482-497.

Zou JB, Peng XL, Li L*, et al.* (2012) Molecular phylogeography and evolutionary history of *Picea likiangensis* in the Qinghai-Tibetan Plateau inferred from mitochondrial and chloroplast DNA sequence variation. *Journal of Systematics and Evolution* **50**, 341-350.

Chapter II. Out of the Qinghai-Tibetan Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophae rhamnoides* (Elaeagnaceae)

Abstract

Numerous temperate plants now distributed across Eurasia are hypothesized to have originated and migrated from the Qinghai-Tibetan Plateau (QTP) and adjacent regions. However, this hypothesis has never been tested through a phylogeographic analysis of a widely distributed species. Here we use *Hippophae rhamnoides* L. as a model to test this hypothesis. We collected 635 individuals from 63 populations of the nine subspecies of *H. rhamnoides*. We sequenced two maternally inherited chloroplast (cp) DNA fragments and also the bi-paternally inherited nuclear ribosomal ITS. We recovered five major clades in phylogenetic trees constructed from cpDNA and ITS sequence variation. Most sampled individuals of six subspecies that are distributed in northern China, central Asia and Asia Minor/Europe, respectively, comprised monophyletic clades (or subclades) nested within those found in the QTP. Two subspecies in the QTP were paraphyletic while the placement of another subspecies from the Mongolian Plateau differed between the ITS and cpDNA phylogenetic trees. Our phylogeographic analyses supported an 'out-of-QTP' hypothesis for *H. rhamnoides* followed by allopatric divergence, hybridization and introgression. These findings highlight the complexity of intraspecific evolutions and the importance of the QTP as a center of origin for many temperate plants.

Key words: Qinghai-Tibetan Plateau, migration, phylogeography, *Hippophae rhamnoides*, intraspecific evolutions

Introduction

A central objective of biogeography is to identify the underlying causes of how organisms are geographically distributed (Cox & Moore 2000). The origins and migration patterns of Northern Hemisphere (NH) temperate plants, which exhibit the greatest diversification in the highlands of Asia, have attracted the attention of numerous botanists for many years (e.g. Donoghue *et al.* 2001; Donoghue & Smith 2004; Ozenda 1988; Wulff 1943). Several have suggested that the majority of these temperate groups originated in the Qinghai-Tibetan Plateau (QTP) and adjacent highlands, when global temperature decreased and the QTP uplifted extensively during the Neogene (Axelrod *et al.* 1996; Kadereit *et al.* 2008; Ozenda 1988; Wu 1988; Wu & Wu 1996; Wulff 1943; Zhang *et al.* 2001; Zhang *et al.* 2004). Evidence in support of this hypothesis has been sought from phylogenetic analyses of genera containing species that are disjunctly and widely distributed in the NH. Three different biogeographic patterns have emerged from such analyses. The first indicates that most of such genera originated in the QTP and adjacent regions, and then migrated to other NH regions where they gave rise to daughter species (e.g. Xu *et al.* 2010; Zhang *et al.* 2009; Zhang *et al.* 2007). This is especially true for numerous forest genera containing some lineages that migrated to North America and other temperate regions at different times during the past 30 million years (Donoghue & Smith 2004). The second pattern indicates that the disjunct distribution of some genera in the NH originated from local relics of the once continuous Arcto-Tertiary, Tethyan or boreal floras (e.g. Mao *et al.* 2010; Sun *et al.* 2001). The third pattern suggests that a limited number of genera originated in other regions of the world and diversified greatly after their ancestors reached the QTP (e.g. Liu *et al.* 2002; Tu *et al.* 2010). These

different patterns, therefore, suggest that biogeographic connections between the QTP and other NH regions are more complex than previously thought.

Previous phylogenetic studies have focused on species distributions within genera and not below the species level. However, phylogenetic analyses of populations from a single species across its entire range can provide important insights into the origins and migration of organism especially over relatively recent historical time (Abbott *et al.* 2000; Avise 2000; Qiu *et al.* 2011). Here we report a phylogenetic analysis of *Hippophae rhamnoides* L. (Sea buckthorn) to test the out-of-QTP hypothesis for a temperate plant species. *Hippophae* is a small genus of Elaeagnaceae, comprising seven species (Bartish *et al.* 2002; Rousi 1971; Swenson & Bartish 2002). Two of the species are putative homoploid hybrids (*H. goniocarpa* and *H. litangensis*), and their taxonomic status is uncertain (Swenson & Bartish 2002). All species of *Hippophae* are restricted to the QTP region and adjacent areas, except for *H. rhamnoides*, which contains nine subspecies (Lian *et al.* 2003; Swenson & Bartish 2002) that are distributed widely but fragmentally in eastern Asia (ssp. *yunnanensis*, *wolongensis*, *sinensis* and *mongolica*), central Asia (ssp. *turkestanica*), Asia Minor (ssp. *caucasica*) and Europe (ssp. *carpatica*, *fluviatilis* and *rhamnoides*; see Fig. 2-1). Its Eurasian range makes *H. rhamnoides* a good candidate to examine biogeographic connections between the QTP and Europe through intervening mountain ranges in central Asia and Asia Minor (Kadereit *et al.* 2008). Although several phylogenetic studies have indicated east-to-west directional dispersal in the genus (Bartish *et al.* 2000; Bartish *et al.* 2002; Sun *et al.* 2002), this hypothesis has not been tested at the species level.

Fruits of *H. rhamnoides* are juicy favoring dispersal by birds and the possibility of

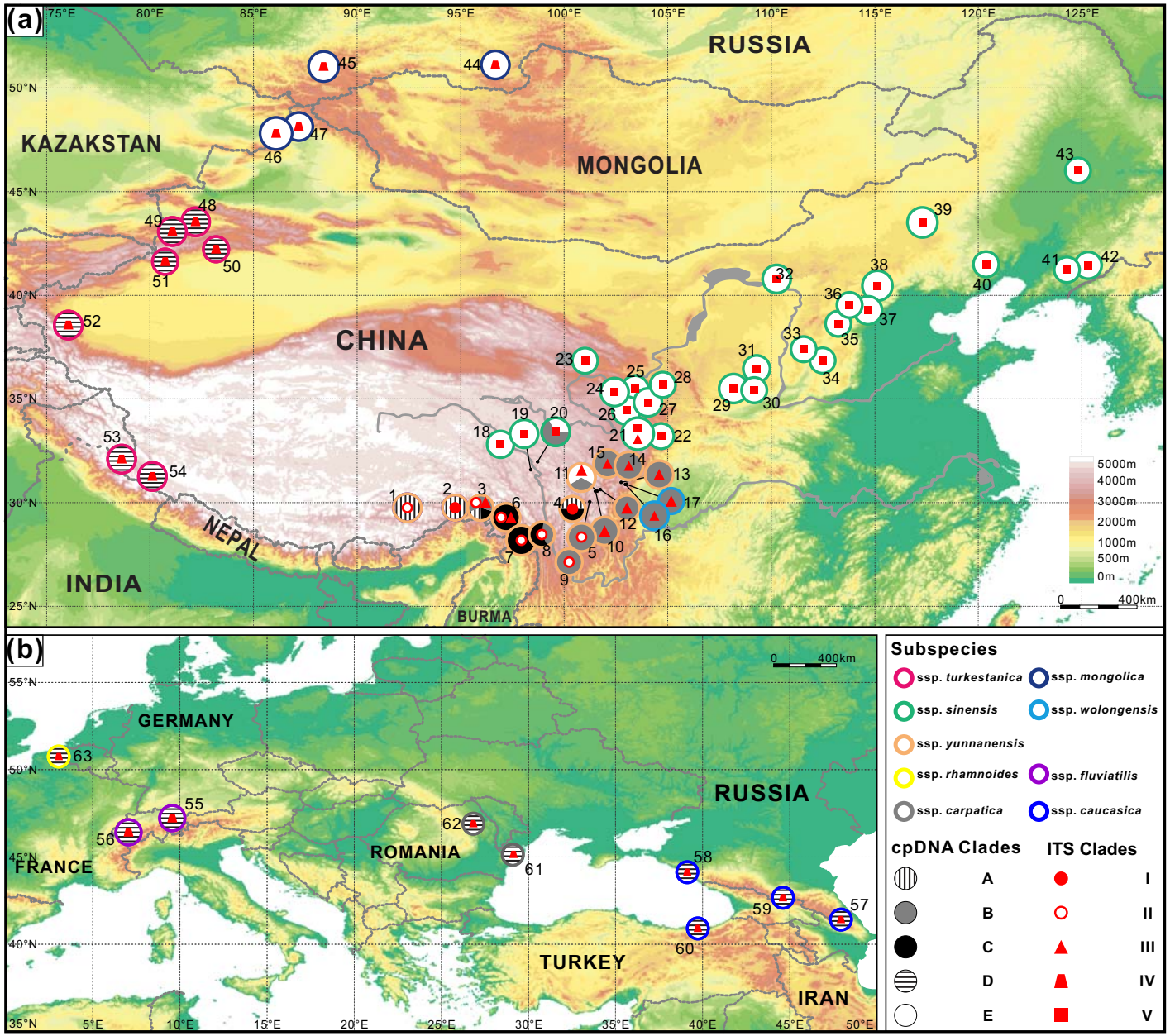


Fig. 2-1 Map of sampling sites (details in Supporting information Table S2-1) and geographic distribution of *Hippophae rhamnoides* cpDNA and internal transcribed spacer (ITS) clades (Fig. 2-2) in Asia (a) and Europe/Minor Asia (b). Pie charts show the relative proportion of each cpDNA clade within each population and presence/absence of ITS clades. The color of the outer ring of each pie corresponds to the subspecies present at a location according to morphological delimitation.

long-distance dispersal leading potentially to inter-regional hybridization (Rousi 1971). Such long-distance dispersal and inter regional hybridization is likely to be further promoted by strong pollen-mediated gene flow because all subspecies of *H. rhamnoides* are dioecious and wind-pollinated (Bartish *et al.* 2000; Hyvönen 1996; Lian *et al.* 1998; Rousi 1971). Two previous phylogenetic analyses of the genus *Hippophae* based on chloroplast DNA restriction fragment length polymorphisms (Bartish *et al.* 2002), and nuclear ribosomal internal transcribed spacer (ITS) sequence variation (Sun *et al.* 2002), included only single representatives of each subspecies of *H. rhamnoides* and suggested that they clustered as a monophyletic lineage, in spite of being genetically distinct. These subspecies occupy different habitats from high mountains (around 3800 m) to low-altitude regions (300 m) and also deserts or seashores (0 m) (Lian *et al.* 1998). Bartish *et al.* (2006) previously conducted a phylogeographic study of *H. rhamnoides*, throughout its range in Europe and Asia Minor. Although this study was highly informative of the demographic history of the species in the western part of its range, it was not, of course, informative of the species' biogeographic history across Eurasia, especially in the eastern Asian mountains.

In the present study, we used maternally inherited chloroplast DNA (cpDNA) (Bartish *et al.* 2002) and bi-parentally inherited ITS data in an attempt to reconstruct historical migration routes and divergence patterns of *H. rhamnoides*. These two sets of molecular markers have been shown in studies of other *Hippophae* species to be highly effective in

tracing intraspecific divergence and detecting possible hybridization and introgression events (Jia *et al.* 2011; Wang *et al.* 2008). From phylogenies constructed from sequence variation of both markers we aimed to answer the following specific questions: *(i)* How many intraspecific clades can be recovered in central and eastern parts of the range of *H. rhamnoides*? Are these clades consistent with the morphological delimitation and geographical distribution of recognized subspecies? *(ii)* Do intraspecific cpDNA divergence patterns reflect dispersal and migration routes of the species from the QTP to other regions? *(iii)* Are the ITS and cpDNA phylogenies consistent with each other? If not, how much hybridization and introgression might be indicated by non-congruent topology?

Materials and Methods

Material sampling

We collected leaves of 635 individuals of *H. rhamnoides* from 63 localities (Table S2-1). Our samples include all nine subspecies acknowledged within this species. Because previous studies by Bartish *et al.* (2002, 2006) indicated that four subspecies (ssp. *carpatica*, *caucasica, rhamnoides* and *fluviatilis*) in Europe and Asia Minor comprise a monophylogenetic lineage, we used 27 individuals collected from nine localities to

represent these subspecies (Table S2-1). We focused our sampling on the remaining five subspecies (54 populations) occurring in the QTP, central Asia and northern China (Fig. 2-1, Table S2-1). Based on morphology, each sampled individual from these populations was assigned to one of the subspecies ssp. *yunnanensis*, *wolongensis*, *sinensis*, *mongolica* and *turkestanica*. We sampled 10–16 trees per population with sampled trees located > 100 m apart in a population. Fresh leaves were dried immediately in the field with silica gel and voucher specimens were deposited in Lanzhou University.

DNA extraction, amplification and sequencing

We used DNeasy™ Tissue Kit (Qiagen GmbH, Hilden, Germany) to isolate total genomic DNA from dried leaves and followed Taberlet *et al.* (1991) and Hamilton (1999) to amplify two cpDNA fragments (*trn*L-F and *trn*S-G, respectively), and White *et al.* (1990) to amplify the nuclear ITS region. Polymerase Chain Reactions (PCR) were performed in 25 μL reaction mixture volumes using reagents and manufacture's instruction for Taq polymerase (Takara, Dalian, China) in the GeneAmp® PCR System 9700 (Applied Biosystems, Hayward, USA). PCR cycling programmes followed Jia *et al.* (2011). PCR products were purified using TIANquick Midi Purification Kit (Tiangen, Beijing, China) and sequencing reactions were conducted with the same PCR primers described above and ABI Prism Bigdye™ Terminator v 3.1 Cycle Sequencing Kit

(Applied Biosystems, Foster City, CA, USA). Sequences were obtained using an ABI 3730XL DNA Analyzer.

CLUSTAL X version 1.81 (Thompson *et al.* 1997) was used to align generated sequences, and MEGA version 4 (Tamura *et al.* 2007) was used to adjust them manually. Newly identified sequences were deposited in GenBank under the accession numbers (JQ289173–JQ289289 and JQ663569–JQ663597).

Phylogeny, divergence and biogeography

Elaeagnus umbellata Thunb. was chosen to be outgroup in the phylogenetic analyses, based on previous molecular and morphological analyses on *Hippophae* of Elaeagnaceae (Bartish *et al.* 2002). Three congeneric taxa (*H. salicifolia*, *H. neurocarpa* ssp. *neurocarpa* and *H. neurocarpa* ssp. *stellatopilosa*) were also included. The sequences of these taxa were taken from GenBank. Sequences from two cpDNA regions (*trn*L-F and *trn*S-G) were concatenated into a matrix.

We analyzed phylogenetic relationships of the different cpDNA and nuclear ITS sequences through maximum parsimony (MP) and Bayesian analyses using PAUP* version 4.0b10 (Swofford 2002) and MRBAYES version 3.0 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003), respectively. GAPCODER (Young & Healy 2003) was used to edit indels as separate characters for inclusion in MP and Bayesian analyses.

MP analyses were performed with characters equally weighted through a heuristic search: 10 replicates of random addition of sequences with ACCTRAN character optimization, MULPARS + TBR branch swapping and STEEPEST DESCENT options on. We calculated bootstrap percentages (BP) using 1000 replicates (Felsenstein 1985). Bayesian analyses were performed over 5 x 10^6 generations with one cold and three incrementally heated Monte Carlo Markov chains (MCMCs). Best-fit models of nucleotide substitution for cpDNA and ITS matrices were determined to be TPM1uf + G and SYM + G, respectively, by the AIC using JMODELTEST version 0.1.1 (Guindon & Gascuel 2003; Posada 2008). A standard discrete model (Lewis 2001) was applied to the indel matrix. Model parameters were unlinked across partitions. We used defaults for the heating scheme as well as for priors on the rate matrix, gamma shape parameter, the proportion of invariable sites and branch lengths. Dirichlet distributions were used to model base frequency parameters and uninformative priors placed on tree topology. We sampled one tree per 500 generations. After discarding the first 2,500 trees out of the 10,001 trees as burn-ins, the remaining trees were used to estimate posterior probability (PP).

To detect genealogical relationships among sequences with shallow genetic divergences, we also constructed cpDNA and ITS haplotype networks using a statistical parsimony algorithm described by Templeton *et al.* (1992) as implemented in TCS version 1.21 (Clement *et al.* 2000). As the TCS program collapses sequences that differ only by additive sites, we grouped the 78 different ITS sequences recorded into 28 types

and selected manually one representative sequence possessing the lowest number of additive sites for each type (Table S2-2). These 28 general ribotypes were used in further analyses to reduce bias introduced by sequence ambiguities (Koch & Matschinger 2007). We ran TCS with a default parsimony connection limit of 95%. In the case of ITS data, after starting initially with the default limit (11 internal steps allowed), we added ribotypes R68, R76 and R77 with the 90% confidence interval (five steps) and finally R55 and R56 when fixing the connection limit at 33 steps.

The datasets obtained were examined to see if they fitted the hypothesis of a molecular clock using a likelihood-ratio test (LRT; Huelsenbeck & Rannala 1997) implemented in PAUP* version 4.0b10 (Swofford 2002) and comparing the log likelihood (ln L) of the ML trees with and without assuming a molecular clock. A molecular clock could not be rejected for both cpDNA and ITS data (cpDNA: TPM1uf + G, 2lnLR = 67.889, d.f. = 51, $P = 0.057$; ITS: SYM + G, 2lnLR = 77.053, d.f. = 80, $P = 0.573$). We therefore used BEAST version 1.5.4 (Drummond *et al.* 2002; Drummond & Rambaut 2007) to estimate genetic divergence. We used GTR + G substitution model, a fixed molecular clock, a constant population size coalescent tree prior and an UPGMA starting tree for both data sets. We sampled all parameters once every 2,000 steps from 20×10^6 MCMC steps after a burn-in of 5×10^6 steps. The program TRACER (Rambaut & Drummond 2007) was used to examine convergence of chains to the stationary distribution and the analysis was repeated before combining the two independent runs together. The cpDNA substitution

rates for most angiosperm species have been estimated to vary between 1 and 3×10^{-9} substitutions per site per year (s s^{-1} yr^{-1}) (Wolfe *et al.* 1987) while those for nrITS in shrubs and herbal plants vary between 3.46 and 8.69×10^{-9} s s^{-1} yr^{-1} (Richardson *et al.* 2001). Given the uncertainties in these rate values, we used normal distribution priors with a mean of 2×10^{-9} and a SD of 6.080×10^{-10} for cpDNA, and a mean of 6.075×10^{-9} and a SD of 1.590×10^{-9} for ITS to cover these rate ranges within the 95% range of the distribution for our estimation of divergence times of major clades.

We ran Bayes-DIVA analyses using RASP version 2.0b (Reconstruct Ancestral State in Phylogenies; Yu *et al.* 2010, 2011) to infer the biogeographic history of *H. rhamnoides* based on the phylogeny constructed from cpDNA and nrITS. In this analysis we defined five distribution regions for the sampled populations: (Q) the QTP region (including Tibet, Qinghai, northwest Yunnan, southwest Sichuan, south Gansu; populations 1–28); (N) northern China (including east Gansu, Inner Mongolia, Shaanxi, Shanxi, Hebei, Liaoning; populations 29–43); (M) the Mongolian Plateau (including north Xinjiang, Mongolia, Siberia; populations 44–47); (C) central Asia (including west Xinjiang and west Tibet; populations 48–54); (E) Europe and Asia Minor (populations 55–63) (Figs 1, 3). We loaded 10,001 trees previously produced in MRBAYES version 3.0 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) and chose the F81 model for the Bayesian MCMC analyses, allowing for different rates of change among ancestral areas.

Population genetic analyses

Population gene diversity (H_S, H_T) and between-population divergence (G_{ST}, N_{ST}) was estimated within each region using the program PERMUT with 1000 permutations tests (Pons & Petit 1996; available at http://www.pierroton.inra.fr/genetics/labo/Software/PermutCpSSR). A significant difference between N_{ST} and G_{ST} may suggest a significant phylogeographic structure (Pons & Petit 1996). We assessed genetic differentiation between all pairs of geographical regions using Φ_{ST} as estimators following analyses of molecular variance (AMOVA; Excoffier *et al.* 1992) in ARLEQUIN version 3.1 (Excoffier *et al.* 2005). We also conducted two types of hierarchical (3-level) AMOVAs, i.e. with the highest level partitioned by either subspecies or regions. Finally, we calculated haplotype diversity (h) and nucleotide diversity (π) for each population and each subspecies using ARLEQUIN.

A mismatch distribution analysis (MDA; Rogers & Harpending 1992; Schneider & Excoffier 1999; Slatkin & Maddison 1989) was conducted to examine the demographic expansions of the five major cpDNA clades identified in the phylogenetic analyses. As population structure has a limited effect on the mismatch distribution (Bernatchez 2001; Rogers 1995), we pooled all haplotypes of each clade and did not consider their frequencies. We used 1000 parametric bootstrap replicates to generate an expected distribution using a model of sudden demographic expansion (Excoffier *et al.* 2005), to

calculate the sum of squared deviations (SSD) and raggedness index (*HRag*) of Harpending (1994) between observed and expected mismatch distributions and to obtain 95% confidence intervals (CIs) around τ. We also calculated Tajima's *D* (Tajima 1989) and Fu's F_S (Fu 1997) to assess possible expansions. The *D* and *F*s statistic should have large negative values within a clade under the expansion hypothesis due to an excess of rare new mutations. We calculated significance of the tests with 10,000 replicates. All of these demographic tests were done using ARLEQUIN version 3.1 (Excoffier *et al.* 2005). When sudden expansions were detected, we used the equation $\tau = 2ut$ (Rogers & Harpending 1992; Rogers 1995) to estimate expansion times: *t*, the expansion time in number of generations, τ the mode of the mismatch distribution, and *u* the mutation rate per generation for the whole analyzed sequence (i.e. chlorotype). We calculated *u* according to $u = \mu kg$, where μ is the substitution rate per nucleotide site per year, *k* is the average sequence length of the DNA region under study, and g is the generation time in years. We assumed the generation time of *H. rhamnoides* to be five years (Bartish *et al.* 2006). We again adopted the substitution rate of 2.0×10^{-9} s s^{-1} yr^{-1} for expansion estimations.

Results

cpDNA variation

From the 635 individuals sampled, we recovered 27 and 29 unique sequences for the *trn*L-F and *trn*S-G fragments, respectively. After combining both sequences we identified a total of 49 chlorotypes (C01–C49). Sequence lengths of these haplotypes varied from 1357 bp to 1514 bp with 59 nucleotide substitutions and 11 indels (1–163 bp) recorded. Thirty-four haplotypes (69.39%) occurred in not more than one population while the remaining 15 (30.61%) were shared between populations (Table S2-1). However, all chlorotypes were subspecies-specific and therefore none was shared by any two subspecies.

Phylogenetic trees constructed using MP and Bayesian methods were largely consistent in topology (Fig. 2-2a). All haplotypes from *H. rhamnoides* comprised a monophyletic lineage with five clades (A–E) showing no exact 'one to one' match with subspecies. All haplotypes of ssp. *yunnanensis* clustered into three major clades (A, B and C) apart from C32 (found in population 11) which was placed in clade E. Haplotypes placed in clades A, B or C were exclusively distributed in the QTP region (Fig. 2-1). Clade E comprised most haplotypes of ssp. *sinensis*, plus one from ssp. *mongolica* (C26) in addition to the one of ssp. *yunnanensis* (C32) mentioned above. However, C08 of ssp. *sinensis* nested within clade B. A monophyletic subclade comprising four haplotypes of ssp. *wolongensis* also nested within clade B. Clade D comprised two sub-clades: one consisted of haplotypes of ssp. *turkestanica* distributed in central Asia, and the other

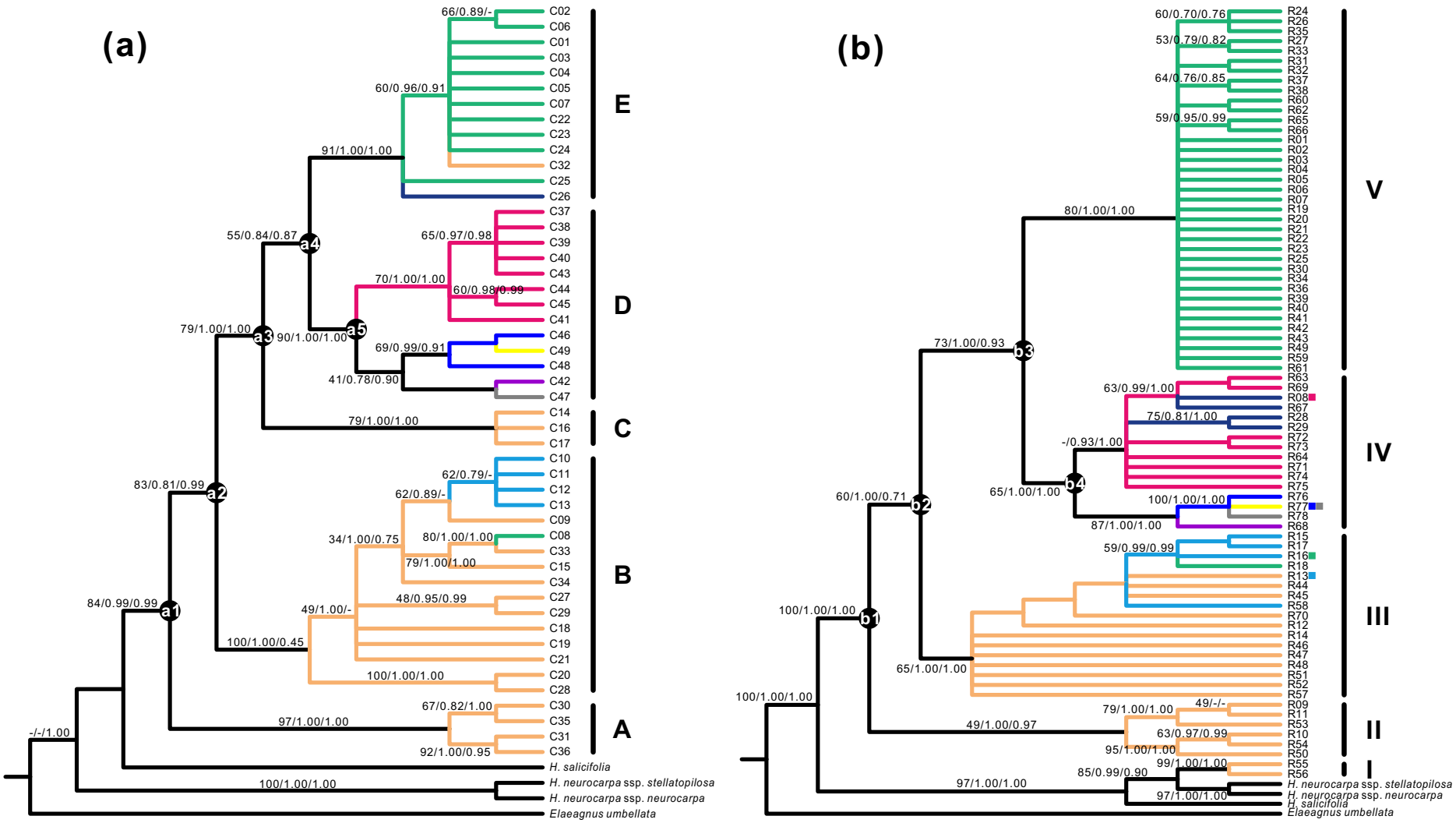


Fig. 2-2 Maximum parsimony 50% consensus trees based on phylogenetic analyses of (a) cpDNA (length = 297, CI = 0.808, RI = 0.883) and (b) ITS (length = 230, CI = 0.848, RI = 0.955). Support values (MP bootstrap/Bayesian posterior probability/BEAST posterior possibilities) are shown at nodes. Colors correspond to the different subspecies examined in Fig. 2-1.

included ssp. *caucasica*, *carpatica*, *fluviatilis* and *rhamnoides* distributed in Asia Minor and Europe (Fig. 2-1 and Fig. 2-2a). The haplotype network (Fig. S2-1a) was largely congruent with the phylogenetic trees (Fig. 2-2a), although depicted relationships between haplotypes in more detail. The dating analysis estimated that the five major clades diversified from the late Miocene to the late Pliocene (Fig. 2-2a), with the first (node a1) and last major events (node a4) occurring 5.46 (95% HPD: 2.23–10.10) Ma and 2.05 (95% HPD: 0.72–3.86) Ma (Table S2-3), respectively.

ITS variation

We recovered 78 different ITS sequences (ribotypes, R01–R78), of which 50 (64.10%) were unique to single populations. A total of 108 base substitutions and five indels (1–2 bp) were found and the two phylogenetic analyses generated topologically similar trees (Fig. 2-2b). Two ribotypes (R55 and R56) of ssp. *yunnanensis* clustered with *H. neurocarpa* ribotypes into a clade also containing *H. salicifolia* ribotype with high support, indicating possible past hybridization and introgression between these species. The other 76 ribotypes comprised a monophyletic *H. rhamnoides* lineage with high support. However, phylogenetic relationships between subspecies differed somewhat from those inferred from cpDNA sequences.

Ribotypes from ssp. *yunnanensis* clustered into three basal clades (I, II and III) with those of ssp. *wolongensis* also placed within clade III. Clade V consisted of only ribotypes from ssp. *sinensis*, while clade IV included all the ribotypes from the other

six subspecies. Four ribotypes were shared between/among subspecies: R13 between ssp. *yunnanensis* (occurred in populations 6, 10, 11, 13, 14 and 15) and ssp. *wolongensis* (population 17), R16 between ssp. *wolongensis* (population 16) and ssp. *sinensis* (population 21), R08 between ssp. *turkestanica* (populations 48, 49, 50, 51 and 52) and ssp. *mongolica* (populations 45, 46 and 47) and R77 among ssp. *caucasica* (populations 57, 58 and 59), ssp. *carpatica* (population 61) and ssp. *rhamnoides* (population 63). R18, the only ribotype of ssp. *sinensis* placed in clade III occurred in population 21. The first divergence within *H. rhamnoides* (node b1; Fig. 2-2b) resulting in the origin of clade II and the ancestor of clades III, IV and V in the ITS tree was estimated to have occurred 3.02 (95% HPD: 1.42–5.18) Ma. The second divergence (node b2) occurred 2.56 (95% HPD: 1.22–4.33) Ma, while the third divergence (node b3) occurred 2.44 (95% HPD: 1.07–4.19) Ma (Table S2-3).

Biogeographic analyses and inter-subspecies differentiation

Bayes-DIVA analyses of cpDNA data supported the QTP region (area Q) as the ancestral area for *H. rhamnoides* (Fig. 2-3b). This species most likely diversified early in the QTP and subsequently multiple dispersals (for each node only the most likely reconstruction was considered) (Figs 3a, 3b) occurred from this area to central Asia (C), Asia Minor/Europe (E), northern China (N) and Mongolian Plateau (M), respectively. Bayes-DIVA analysis based on nrITS data resulted in similar reconstructions, but required additional dispersals events from central Asia to the

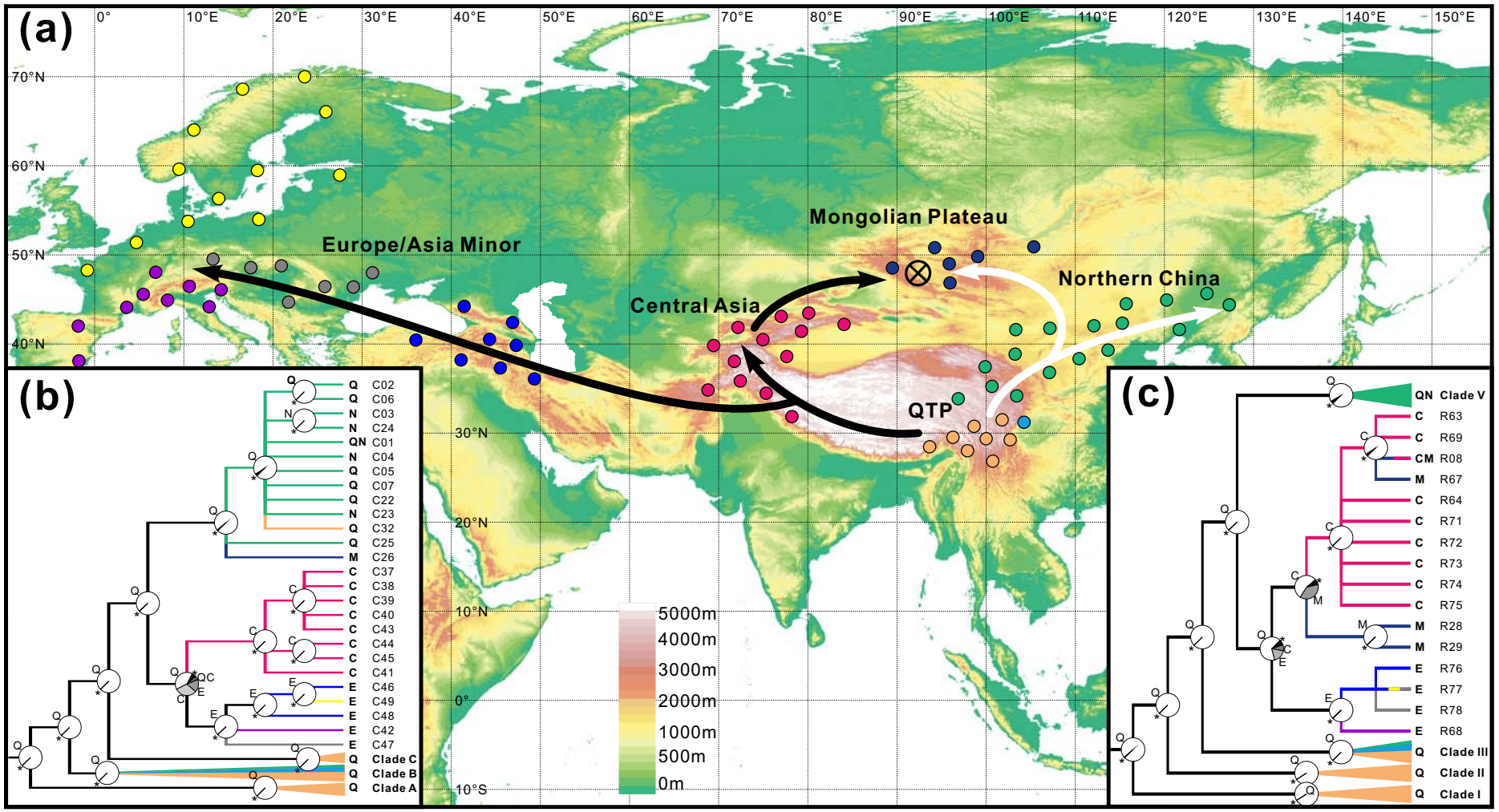


Fig. 2-3 (a) Migration routes of *H. rhamnoides* out of the Qinghai-Tibetan Plateau (QTP) and statistical reconstructions (Pie charts) of ancestral areas based on the Bayes-DIVA analyses of cpDNA data (b) and ITS data (c). Color coding follows Figs 1 and 2. Five major distributions were defined as Qinghai-Tibetan Plateau (Q), northern China (N), Mongolian Plateau (M), central Asia (C) and Europe/Asia Minor (E). Sequences of the same subspecies or distribution were compressed. Ancestral areas with probability < 0.05 were represented by asterisks.

Mongolian Plateau (Figs 3a, 3c). However, both Bayes-DIVA analyses based on cpDNA and nrITS did not detect any distinct vicariance signal, although vicariance between C and E or between N and M seem to be likely.

Table 2-1 Estimates of average gene diversity within populations (H_S) of *Hippophae rhamnoides*, total gene diversity (H_T), interpopulation differentiation (G_{ST}), and number of substitution types (N_{ST}) for chlorotypes and ribotypes across regions

			cpDNA					ITS				
Region	#Pop	*N*	*n*	H_S	H_T	G_{ST}	N_{ST}	*n*	H_S	H_T	G_{ST}	N_{ST}
QTP	28	321	31	0.314	0.955	0.672	0.896*	13	0.117	0.811	0.856	0.947*
Northern China	15	167	5	0.112	0.355	0.683	0.671	6	0.223	0.667	0.666	0.678
Mongolian Plateau	4	54	1	–	–	–	–	3	0.110	0.738	0.851	0.990
Central Asia	7	76	8	0.173	0.856	0.798	0.896*	5	0.287	0.594	0.517	0.596
Total	54	618	44	0.216	0.905	0.761	0.870*	25	0.168	0.854	0.804	0.968*

#Pop, No. of populations; *N*, No. of plants; *n*, No. of chlorotypes/ribotypes; *, $P < 0.005$.

For both cpDNA and ITS data sets, a significantly larger N_{ST} than G_{ST} value estimated across all populations indicated that genetic variation of the species was geographically structured across its distribution (Table 2-1). Highest genetic diversity occurred in the QTP (H_T = 0.955 and 0.811 for cpDNA and ITS data, respectively), corresponding to the occurrence of the most divergent haplotypes (Fig. 2-1). Pairwise Φ_{ST} showed high levels of divergence among regions for both cpDNA and ITS sequences (Table 2-2) and further analysis by hierarchical AMOVA showed that for cpDNA 54% of the total genetic variation was partitioned by subspecies, and 33% partitioned by regions (Table 2-3), while for ITS data, 77% of the total genetic variation was attributed to the differences among subspecies and 59% to the differences among regions.

Table 2-2 Pairwise genetic differentiation (Φ_{ST}) among regions estimated from ITS sequences (upper part) and cpDNA sequences (lower part)

	QTP	Northern China	Mongolian Plateau	Central Asia
QTP		0.960	0.976	0.979
Northern China	0.726		0.990	0.990
Mongolian Plateau	0.703	0.963		0.786
Central Asia	0.815	0.998	0.998	

All values are significant at the 0.01 level in a permutation tests (1000 permutations).

Table 2-3 Analysis of molecular variance (AMOVA) of chlorotypes and ITS ribotypes for *Hippophae rhamnoides* populations, partitioned by subspecies and regions, respectively

			cpDNA			ITS		
Partitioning	Source of variation	d. f.	SS	VC	PV (%)	SS	VC	PV (%)
By subspecies								
	Among subspecies	4	8986.633	20.409	54.05	3215.827	7.515	76.78
	Among populations	49	6134.421	10.357	27.43	1151.816	2.039	20.84
	Within populations	564	3945.402	6.995	18.53	131.424	0.233	2.38
	Total	617	19066.456	37.762		4499.068	9.787	
By region								
	Among regions	3	5284.371	11.899	33.49	2310.731	5.563	59.30
	Among populations	50	9836.683	16.632	46.82	2056.912	3.586	38.22
	Within populations	564	3945.402	6.995	19.69	131.424	0.233	2.48
	Total	617	19066.456	35.527		4499.068	9.382	

d.f., degrees of freedom; SS, sum of squares; VC, variance components; PV, percentage of variation.

Demographic analyses

Under a model of population expansion, only clade E identified in the cpDNA phylogeny, and containing haplotypes mainly of ssp. *sinensis*, showed a strongly unimodal mismatch distribution (Fig. S2-2), indicating that it underwent an expansion in the recent past. Non-significant SSD and raggedness index values (P_{SSD} = 0.12 and P_{HRag} = 0.10), as well as a

significantly large negative F_S value (-15.557, $P < 0.001$; Table 2-4), also suggested a historical demographic expansion within clade E. Assuming cpDNA mutation rates of 2×10^{-9} s s^{-1} yr^{-1}, the expansion of clade E was estimated to have occurred 90.4 (95% CI: 40.1–139.5) thousand years ago.

Table 2-4 The results of Tajima's *D* and Fu's F_S tests, and mismatch distribution analyses (MDA) for the five lineages (A–E) of *Hippophae rhamnoides* chlorotypes

Clades	Tajima's *D* test		Fu's F_S test		Mismatch distribution					
	D	*P*	F_S	*P*	τ	95% CI	*Hrag*	*P*	SSD	*P*
A	-0.605	0.373	0.314	0.363	8.602	4.879–13.453	0.222	0.780	0.073	0.730
B	-1.577	0.048	-6.576	0.006	8.756	4.930–29.469	0.022	0.630	0.022	0.360
C	0	0.822	0.134	0.272	6.289	0.885–91.289	0.444	0.810	0.186	0.280
D	-0.993	0.162	-8.838	< 0.001	6.602	3.486–9.455	0.030	0.680	0.014	0.535
E	-1.335	0.088	-15.557	< 0.001	2.248	0.998–3.467	0.163	0.100	0.039	0.120

Discussion

Previous studies of the historical biogeography of northern hemisphere, widespread, temperate plant groups having greatest diversification in the highlands of Asia have tended to focus on the phylogenetic analysis of genera showing such distribution (e.g. Mao *et al.* 2010; Tu *et al.* 2010; Xu *et al.* 2010; Zhang *et al.* 2009; Zhang *et al.* 2007). Here we have focused on the historical biogeography of a widespread, temperate species in Eurasia, *H. rhamnoides*, rather than a genus, and shown that a phylogeographic analysis of cpDNA and ITS sequence variation in this species strongly supports an 'out of

QTP' origin and migration pattern as suggested by phylogenetic analyses at the genus level (e.g. Xu *et al.* 2010; Zhang *et al.* 2009; Zhang *et al.* 2007). Our results further suggest that following dispersal from the QTP region, allopatric divergence took place, sometimes followed by secondary contact, hybridization and introgression between divergent intraspecific taxa.

Out of the Qinghai-Tibetan Plateau

"Dispersal is often inferred if the taxa in one area are phylogenetically derived from lineages that have more 'primitive' (phylogenetically basal) members in another area, which is inferred to be the source area" (Futuyma 1998; p. 208). Our phylogenetic trees of both ITS ribotype and cpDNA haplotype variation showed that the phylogenetically basal clades of *H. rhamnoides* (II and III in the ITS tree, and A, B and C in the cpDNA tree) were exclusively distributed in the QTP, thus supporting an out-of-QTP hypothesis for the origin and dispersal of the species (Figs 1, 2). The phylogeographic analysis of cpDNA variation showed that the two other clades identified (D and E) contained haplotypes found only in individuals distributed in northern China, the Mongolian Plateau, central Asia, Asia Minor and Europe. Because clades D and E were nested within two QTP clades (A and B) and sister to QTP clade C, our results further indicated an origin of *H. rhamnoides* in the QTP region and subsequent colonization by dispersal of

other parts of Eurasia. Further support for this hypothesis came from Bayes-DIVA biogeographic analyses of the cpDNA and nrITS trees which suggested that multiple dispersals from the QTP accounted for most of the species' range expansion in Asia (Fig. 2-3). The inferred migration route to Europe of *H. rhamnoides* through intervening mountain ranges in central Asia and Asia Minor is one suggested previously for many plants of the European Alps and neighbouring mountain ranges that originated in the QTP region (Kadereit *et al.* 2008). It was estimated that dispersal to central Asia (represented by clade D of cpDNA and clade IV of nrITS) preceded that to northern China (represented by clade E of cpDNA and clade V of nrITS) with both dispersal events having their source in the QTP. Because cpDNA is maternally inherited in *Hippophae* (Bartish *et al.* 2002), the phylogeographic patterns observed are likely to reflect past migration of the species through long-distance seed dispersal mediated by birds. The fleshy fruits of the species are preferred by migrating birds that stay at high-altitude in the central QTP in summer (to avoid high temperature) and return to northern China or central Asia in autumn (e.g. Blanford's snow finch; Cramp & Perrins 1994; Qu *et al.* 2010).

Allopatric divergence, glacial refugia and local expansion

Clade divergence within the QTP (i.e. clades A–C for cpDNA, and clades II and III for

ITS) and also outside the QTP (clades D and E for cpDNA, and IV and V for ITS) indicates that *H. rhamnoides* underwent allopatric divergence in different parts of its distribution during its evolutionary history. Clades A, B and C comprise cpDNA haplotypes found mainly in the western, eastern and central parts of the species' distribution range in the QTP, respectively, while clades D and E comprise haplotypes that occur mainly in Europe/central Asia and northern China/Mongolia, respectively (Figs 1, 2a). Similarly, ITS clades (I, II and III) are distributed mainly in western-central, southeastern, and northeastern parts of its range in the QTP, while ribotypes in clades IV and V occur in Europe/central Asia/Mongolia and northern China, respectively (Figs 1, 2b). Although there is no 'one to one' match between the five cpDNA clades and the nine subspecies sampled within *H. rhamnoides*, haplotypes contained within ssp. *wolongensis* form a subclade of clade B, while those present in ssp. *turkestanica, caucasica*, *carpatica* and *fluviatilis* comprise clade D, and those in ssp. *mongolica* and *sinensis* mainly comprise clade E. Thus there is evidence of an association between phylogenetic and taxonomic divergence within the species. The same is also apparent from an examination of the ITS phylogeny (Fig. 2-2b).

The major cpDNA and ITS clades were estimated to have diverged from the starting Pliocene to the middle Quaternary (Table S2-3). The available fossil record of *Hippophae* in Europe is too recent (the middle Pleistocene; Krupinski 1992; Lang 1994) for dating analyses based on ITS or cpDNA, while the fossil record of this genus in eastern Asia is

very poor. Therefore, divergence dates were estimated from conventional substitution rates of 1–3 × 10^{-9} s s^{-1} yr^{-1} for cpDNA (Wolfe *et al.* 1987) and 3.46–8.69 × 10^{-9} s s^{-1} yr^{-1} for ITS sequences (Richardson *et al.* 2001). There was a reasonable match in divergence dates for the main cpDNA and ITS clades, i.e. spanning the Pliocene and middle Quaternary periods. It is thought that the QTP underwent an extensive uplift during the Pliocene and Quaternary periods (Li & Li 1991; Shi *et al.* 1998) with climates oscillating greatly during this period (Shi 2002). It is feasible, therefore, that these marked geological and climatic changes could have fragmented the distribution of *H. rhamnoides* within the QTP and triggered the phylogenetic divergence recorded. Clearly, acceptance of this possibility needs to be treated with caution given the uncertainty of the clade divergence dates; however, multiple glacial refugia in the QTP during the Quaternary (Li *et al.* 2011; Tang *et al.* 2010; Wang *et al.* 2009; Wu *et al.* 2010) are thought to have caused similar deep phylogenetic divergence within several other plant species found in this region (e.g. Tang *et al.* 2010; Wang *et al.* 2009), including a congener *H. tibetana* (Jia *et al.* 2011; Wang *et al.* 2010).

The latest divergence within *H. rhamnoides* clades (i.e., between central Asia and Asia Minor/Europe; node a5 in Fig. 2-2a and node b4 in Fig. 2-2b) was estimated to have occurred between 1.45 (95% HPD: 0.53–2.82) and 1.87 (95% HPD: 0.82–3.25) Ma (Table S2-3), indicating that populations of the species in the QTP, northern China, the Mongolian Plateau, central Asia and Asia Minor/Europe have survived in these different

regions since at least this time. The divergences within the Asia Minor/Europe lineage should be later, around the middle Pleistocene, which is congruent with the fossil records in central Europe (Lang 1994). In our analysis, we were only able to detect a significant range expansion in clade E, which is distributed in northern China/Mongolian Plateau. This expansion was dated to 90.4 (95% CI: 40.1–139.5) thousand years ago, which mostly corresponds to the Last Interglacial Period (Zheng *et al.* 2002) in the late Pleistocene and earlier than the Last Glacial Maximum (LGM; 18–21 thousand years ago). Interestingly, this age is highly consistent with the estimate for range expansion of the species in Asia Minor/Europe (Bartish *et al.* 2006). Failure to detect demographic expansions in the other clades might stem from the limited number of samples and haplotypes recorded or for other reasons currently unknown.

Hybridization and/or introgression

Several examples of phylogenetic incongruence between the cpDNA and ITS trees indicated that historical hybridization and introgression had occurred between *H. rhamnoides* and a close relative, and also between some subspecies of *H. rhamnoides*. Interspecific hybridization between *H. rhamnoides* and *H. neurocarpa* was suggested by the finding that two ribotypes (R55 and R56) found in a few individuals of ssp. *yunnanensis* clustered with *H. neurocarpa* in the ITS tree, whereas their cpDNA

haplotypes were placed firmly within *H. rhamnoides* in the cpDNA tree. Bartish *et al.* (2002) previously proposed that hybridization between these two species resulted in the origin of the two homoploid hybrid species, *H. goniocarpa* and *H. litangensis* (Lian *et al.* 1998; Sun *et al.* 2003). Hybridization between ssp. *yunnanensis* and/or ssp. *wolongensis* and ssp. *sinensis* was also indicated (i) by some individuals of ssp. *sinensis* in population 20 (Table S2-1; Fig. 2-1) possessing cpDNA haplotypes characteristic of ssp. *yunnanensis* and *wolongensis* (i.e. placed in clade B), but ribotypes characteristic of ssp. *sinensis* (i.e. placed in clade V of the ITS tree); and (ii) by certain individuals in population 21 of ssp. *sinensis* possessing a cpDNA haplotype characteristic of this subspecies (i.e. placed in clade E), but a ribotype that clustered with those of ssp. *yunnanensis* and *wolongensis* in clade III of the ITS tree. Finally, there was an indication that ssp. *mongolica* originated from hybridization between ssp. *sinensis* and ssp. *turkestanica* in that its cpDNA haplotype (C26) was placed within those of ssp. *sinensis* in the cpDNA tree (clade E; Fig. 2-2a), while its ribotypes were placed in a subclade with those of ssp. *turkestanica* (clade IV; Fig. 2-2b).

Conclusions

Hippophae rhamnoides is one of seven species recognized in the genus (Swenson & Bartish 2002). Whereas the six other species, *H. goniocarpa*, *H. litangensis*, *H.*

gyantsensis, *H. neurocarpa*, *H. salicifolia* and *H. tibetana*, are restricted in their distribution to the QTP and the adjacent Himalaya region, *H. rhamnoides* is disjunctly distributed across Eurasia occurring in the QTP, northern China, Mongolia, Siberia, central Asia, Asia Minor and Europe. Our phylogeographic analysis of *H. rhamnoides* indicates that the species originated in the QTP and then dispersed out of the QTP to other parts of Eurasia. Thus, our results support the hypothesis that many Northern Hemisphere, widespread, temperate plant groups having the greatest diversification in the highlands of Asia originated in the QTP and adjacent highlands, before migrating to other northern hemisphere regions and undergoing divergence. Both within and outside the QTP *H. rhamnoides* is divergent at the molecular level with several distinct clades resolved in both cpDNA and ITS phylogenetic trees. It was not possible to date the origin of these clades precisely; however, it is feasible that they originated during the Pliocene and Quaternary periods when the QTP and other parts of Eurasia underwent considerable geological and/or climatic oscillations. Such oscillations are likely to have fragmented the distribution of *H. rhamnoides* and triggered allopatric divergence and the formation of deep clades. For only one of the cpDNA clades were we able to detect a significant expansion, which using a mean substitution rate was estimated to have occurred before the Last Glacial Maximum. Although there was no 'one to one' match between the molecular clades and the nine subspecies we examined in *H. rhamnoides*, there was an association between phylogenetic and taxonomic divergence. A comparison of the

topologies of the cpDNA and ITS trees revealed several examples of phylogenetic incongruence indicating that historical hybridization and introgression had occurred between *H. rhamnoides* and a closely related species, and also between some subspecies of *H. rhamnoides*.

References

Abbott RJ, Smith LC, Milne RI*, et al.* (2000) Molecular analysis of plant migration and refugia in the Arctic. *Science* **289**, 1343-1346.

Avise JC (2000) *Phylogeography: The History and Formation of Species* Harvard University Press, London.

Axelrod DI, Al SI, Raven PH (1996) History of the modern flora of China. In: *Floristic characteristics and diversity of East Asian plants* (eds. Zhang AL, Wu SG), pp. 43-55. Springer, Hongkong.

Bartish IV, Jeppsson N, Bartish GI, Lu R, Nybom H (2000) Inter- and intraspecific genetic variation in *Hippophae* (Elaeagnaceae) investigated by RAPD markers. *Plant Systematics and Evolution* **225**, 85-101.

Bartish IV, Jeppsson N, Nybom H, Swenson U (2002) Phylogeny of *Hippophae* (Elaeagnaceae) inferred from parsimony analysis of chloroplast DNA and morphology. *Systematic Botany* **27**, 41-54.

Bartish IV, Kadereit JW, Comes HP (2006) Late Quaternary history of *Hippophae rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (*Chsi*) sequences and chloroplast DNA variation. *Molecular Ecology* **15**, 4065-4083.

Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation.

Evolution **55**, 351-379.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Cox CB, Moore PD (2000) *Biogeography: An Ecological and Evolutionary Approach*, 6th edn. Blackwell Science, Oxford.

Cramp S, Perrins CM (1994) *The Birds of the Western Palearctic* Oxford University Press, New York.

Donoghue MJ, Bell CD, Li J (2001) Phylogenetic patterns in Northern Hemisphere plant geography. *International Journal of Plant Sciences* **162**, 41-52.

Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1633.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**, 1307-1320.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology* **7**, 214.

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.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.

Futuyma DJ (1998) *Evolutionary Biology*, 3rd edn. Sinauer Associates, Sunderland.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by

maximum likelihood. *Systematic Biology* **52**, 696-704.

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**, 521-523.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**, 591-600.

Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**, 227-232.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.

Hyvönen J (1996) On phylogeny of Hippophae (Elaeagnaceae). *Nordic Journal of Botany* **16**, 51-62.

Jia DR, Liu TL, Wang LY, Zhou DW, Liu JQ (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* **102**, 37-50.

Kadereit JW, Licht W, Uhink CH (2008) Asian relationships of the flora of the European Alps. *Plant Ecology & Diversity* **1**, 171-179.

Koch MA, Matschinger M (2007) Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 6272-6277.

Krupinski K (1992) Significance of *Hippophae rhamnoides* L. in evolution of the Eemian Interglacial flora in Warsaw area. *Acta Societatia Botanicorum Poloniae* **61**, 131-144.

Lang G (1994) *Quartäre Vegetationsgeschichte Europas: Methoden und Ergebnisse* Spektrum Akademischer Verlag, Heidelberg, Berlin.

Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**, 913.

Li BY, Li JJ (1991) *Quaternary Glacial Distribution Map of Qinghai-Xizang (Tibet) Plateau* Science Press, Beijing.

Li ZH, Zhang QA, Liu JQ, Kallman T, Lascoux M (2011) The Pleistocene demography of an alpine

juniper of the Qinghai-Tibetan Plateau: *tabula rasa*, cryptic refugia or something else? *Journal of Biogeography* **38**, 31-43.

Lian YS, Chen XL, Lian H (1998) Systematic classification of the genus *Hippophae* L. *Seabuckthorn Research* **1**, 13-23.

Lian YS, Chen XL, Sun K, Ma RJ (2003) A new subspecies of *Hippophae* (Elaeagnaceae) from China. *Novon* **13**, 200-202.

Liu JQ, Gao TG, Chen ZD, Lu AM (2002) Molecular phylogeny and biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics and Evolution* **23**, 307-325.

Mao KS, Hao G, Liu JQ, Adams RP, Milne RI (2010) Diversification and biogeography of *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental dispersals. *New Phytologist* **188**, 254-272.

Ozenda P (1988) *Die Vegetation der Alpen im europäischen Gebirgsraum* Fischer, Stuttgart.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**, 1237-1245.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Qiu YX, Fu CX, Comes HP (2011) Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* **59**, 225-244.

Qu Y, Lei F, Zhang R, Lu X (2010) Comparative phylogeography of five avian species: implications for Pleistocene evolutionary history in the Qinghai-Tibetan plateau. *Molecular Ecology* **19**, 338-351.

Rambaut A, Drummond AJ (2007) Tracer v 1.4. Available from http://beast.bio.ed.ac.uk/Tracer.

Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* **293**, 2242-2245.

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise

genetic differences. *Molecular Biology and Evolution* **9**, 552-569.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution* **49**, 608-615.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Rousi A (1971) The genus *Hippophae* L. A taxonomic study. *Annales Botanici Fennici* **8**, 177-227.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates very among sites: Application to human mitochondrial DNA. *Genetics* **152**, 1079-1089.

Shi YF (2002) Characteristics of late Quaternary monsoonal glaciation on the Tibetan Plateau and in East Asia. *Quaternary International* **97**, 79-91.

Shi YF, Li JJ, Li BY (1998) *Uplift and Environmental Changes of Qinghai-Tibetan Plateau in the Late Cenozoic* Guangdong Science and Technology Press, Guangzhou.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* **123**, 603-613.

Sun H, McLewin W, Fay MF (2001) Molecular phylogeny of *Helleborus* (Ranunculaceae), with an emphasis on the East Asian-Mediterranean disjunction. *Taxon* **50**, 1001-1018.

Sun K, Chen X, Ma R*, et al.* (2002) Molecular phylogenetics of *Hippophae* L. (Elaeagnaceae) based on the internal transcribed spacer (ITS) sequences of nrDNA. *Plant Systematics and Evolution* **235**, 121-134.

Sun K, Ma RJ, Chen XL, Li CB, Ge S (2003) Hybrid origin of the diploid species *Hippophae goniocarpa* evidenced by the internal transcribed spacers (ITS) of nuclear rDNA. *Belgian Journal of Botany* **136**, 91-96.

Swenson U, Bartish IV (2002) Taxonomic synopsis of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **22**, 369-374.

Swofford DL (2002) PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer Associates, Sunderland, MA.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-

coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

Tang LZ, Wang LY, Cai ZY*, et al.* (2010) Allopatric divergence and phylogeographic structure of the plateau zokor (*Eospalax baileyi*), a fossorial rodent endemic to the Qinghai-Tibetan Plateau. *Journal of Biogeography* **37**, 657-668.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. 3. Cladogram estimation. *Genetics* **132**, 619-633.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876-4882.

Tu TY, Volis S, Dillon MO, Sun H, Wen J (2010) Dispersals of Hyoscyameae and Mandragoreae (Solanaceae) from the New World to Eurasia in the early Miocene and their biogeographic diversification within Eurasia. *Molecular Phylogenetics and Evolution* **57**, 1226-1237.

Wang AL, Schluetz F, Liu JQ (2008) Molecular evidence for double maternal origins of the diploid hybrid *Hippophae goniocarpa* (Elaeagnaceae). *Botanical Journal of the Linnean Society* **156**, 111-118.

Wang H, Qiong L, Sun K*, et al.* (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**, 2964-2979.

Wang LY, Abbott RJ, Zheng W*, et al.* (2009) History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnandrum* (Ranunculaceae). *Molecular Ecology* **18**, 709-721.

White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal

RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds. Innis MA, Gelfand DH, Shinsky JJ, White TJ), pp. 315-322. Academic Press, New York.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 9054-9058.

Wu LL, Cui XK, Milne RI, Sun YS, Liu JQ (2010) Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**, 1691-1704.

Wu ZY (1988) Hengduan Mountains flora and its significance. *Journal of Japanese Botany* **63**, 297-311.

Wu ZY, Wu SG (1996) A proposal for a new floristic kingdom (realm) - the E. Asiatic kingdom, its delineation and characteristics. In: *Floristic Characteristics and Diversity of East Asian Plants* (eds. Zhang AL, Wu SG), pp. 3-42. China Higher Education Press, Beijing.

Wulff EV (1943) *An Introduction to Historical Plant Geography* Chronica Botanica Company, Waltham.

Xu TT, Abbott RJ, Milne RI*, et al.* (2010) Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions. *Bmc Evolutionary Biology* **10**, 194.

Young ND, Healy J (2003) GapCoder automates the use of indel characters in phylogenetic analysis. *Bmc Bioinformatics* **4**, 6.

Yu Y, Harris AJ, He X (2010) S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution* **56**, 848-850.

Yu Y, Harris AJ, He X (2011) RASP (Reconstruct Ancestral State in Phylogenies) 2.0 beta. Available at http://mnh.scu.edu.cn/soft/blog/RASP.

Zhang LB, Comes HP, Kadereit JW (2001) Phylogeny and Quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *American Journal of Botany* **88**, 2331.

Zhang LB, Comes HP, Kadereit JW (2004) The temporal course of Quaternary diversification in the European high mountain endemic *Primula* sect. *Auricula* (Primulaceae). *International Journal of Plant Sciences* **165**, 191-207.

Zhang ML, Kang Y, Zhou LH, Podlech D (2009) Phylogenetic origin of *Phyllolobium* with a further implication for diversification of *Astragalus* in China. *Journal of Integrative Plant Biology* **51**, 889-899.

Zhang ML, Uhink CH, Kadereit JW (2007) Phylogeny and biogeography of *Evpimedium*/*Vancouveria* (Berberidaceae): Western North American - East Asian disjunctions, the origin of European mountain plant taxa, and East Asian species diversity. *Systematic Botany* **32**, 81-92.

Zheng BX, Xu QQ, Shen YP (2002) The relationship between climate change and Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and speculation. *Quaternary International* **97-98**, 93-101.

Supporting information

Table S2-1 Locations of populations of *Hippophae rhamnoides* sampled, sample sizes(n), frequencies of cpDNA haplotypes and ITS sequences per population, and estimates of gene diversity (*h*) and nucleotide diversity (π) in percent. Frequencies of chlorotypes/ITS sequences more than two are shown in parentheses. Private sequences particular to each population are underlined

P.	location	Latitude (°N)	Longitude (°E)	Alt. (m)	n	ITS			Chlorotypes (*trn*L-F, *trn*S-G)		
						ITS sequences nos.	*h* (SD)	π (SD) in %	Haplotypes nos.	*h* (SD)	π (SD) in %
ssp. *yunnanensis*					**171**	**20**	**0.74 (0.02)**	**1.88 (0.94)**	**19**	**0.90 (0.01)**	**6.24 (2.98)**
1	Linzhi XZ	29°44.510'	094°43.487'	3640	12	R50(12)	0	0	C30(9) C31(3)	0.41 (0.13)	5.69 (2.97)
2	Mozhu XZ	29°47.466'	091°48.823'	3890	10	R56(10)	0	0	C35(5) C36(5)	0.56 (0.07)	6.38 (3.40)
3	Bomi XZ	29°45.434'	095°58.211'	2950	10	R51(3) R52(3) R53(4)	0.53 (0.09)	0.40 (0.26)	C16(2) C19(2) C21(4) C31(2)	0.80 (0.09)	7.57 (4.03)
4	Litang SC	29°43.330'	100°24.287'	3640	10	R55(4) R56(6)	0.53 (0.09)	0.48 (0.30)	C14(5) C31(5)	0.56 (0.07)	8.41 (4.47)
5	Yajiang SC	30°02.493'	101°13.487'	3150	12	R09(10) R11(2)	0	0	C19(10) C20(2)	0.30 (0.15)	4.79 (2.50)
6	Cha'ou XZ	29°17.456'	097°11.079'	3560	12	R10(10) R13(2)	0.30 (0.15)	0.45 (0.28)	C14(12)	0	0
7	Deqin2 YN	28°27.884'	098°53.217'	3470	14	R10(10) R54(4)	0	0	C14(14)	0	0
8	Deqin YN	28°26.167'	098°57.332'	3570	10	R10(10)	0	0	C14(4) C15(4) C17(2)	0.71 (0.09)	2.69 (1.45)
9	Lijiang YN	27°07.797'	100°14.404'	2930	11	R09(11)	0	0	C09(11)	0	0
10	Daofu SC	30°32.741'	101°31.435'	3580	12	R12(6) R13(4) R14(2)	0	0	C18(12)	0	0
11	Daofu SC	30°32.148'	101°35.564'	3780	12	R13(12)	0	0	C32(8) C33(4)	0.48 (0.11)	2.46 (1.30)
12	Danba SC	30°38.294'	101°43.960'	2600	10	R46(4) R47(3) R48(3)	0.47 (0.13)	0.14 (0.12)	C19(8) C20(2)	0.36 (0.16)	5.62 (2.99)
13	Xiaojin SC	30°58.757'	102°44.094'	2850	12	R13(3) R14(3) R57(6)	0	0	C19(9) C34(3)	0.41 (0.13)	1.43 (0.76)
14	Hongyuan SC	31°55.789'	102°38.654'	3020	12	R13(6) R44(2) R45(4)	0	0	C27(10) C28(2)	0.30 (0.15)	4.81 (2.51)
15	Hongyuan SC	32°06.722'	102°32.709'	3440	12	R13(8) R70(4)	0	0	C29(12)	0	0
ssp. *wolongensis*					**22**	**5**	**0.52 (0.04)**	**0.08 (0.08)**	**4**	**0.71 (0.06)**	**0.06 (0.05)**
16	Wenchuan SC	30°52.400'	102°57.274'	3610	12	R15(4) R16(5) R17(3)	0	0	C10(9) C11(3)	0.41 (0.13)	0.03 (0.03)
17	Wenchuan SC	30°52.609'	102°58.573'	3620	10	R13(3) R58(7)	0	0	C12(8) C13(2)	0.36 (0.16)	0.03 (0.03)
ssp. *sinensis*					**295**	**39**	**0.49 (0.03)**	**0.14 (0.11)**	**12**	**0.63 (0.03)**	**0.32 (0.17)**
18	Yushu QH	32°53.446'	097°03.410'	3810	11	R01(11)	0	0	C01(4) C05 C07(2) C25(4)	0.76 (0.08)	0.07 (0.06)
19	Jiangda XZ	31°35.360'	098°23.090'	3240	12	R01(4) R59(2) R60(4) R61(2)	0	0	C07(12)	0	0

Table S2-1 (*continued*)

20	Dege SC	31°58.130'	098°42.500'	3620	12	R60(4) R62(8)	0.48 (0.11)	0.07 (0.08)	C07(4) C08(8)	0.48 (0.11)	2.46 (1.30)
21	Songpan SC	32°36.674'	103°35.890'	2820	16	R01(12) R16(2) R18(2)	0.43 (0.14)	0.87 (0.49)	C01(3) C02(2) C07(11)	0.51 (0.13)	0.05 (0.04)
22	Wenxian GS	33°03.107'	104°39.217'	1810	10	R05(7) R07(3)	0	0	C01(10)	0	0
23	Haiyan QH	36°52.810'	101°00.785'	3004	10	R01(6) R06(2) R34(2)	0	0	C25(10)	0	0
24	Hezuo GS	34°55.319'	102°51.734'	3034	12	R01(8) R49(4)	0	0	C02(10) C06(2)	0.30 (0.15)	0.02 (0.03)
25	Lintan GS	34°45.000'	103°33.000'	3027	11	R01(4) R06(3) R65(3) R66	0.51 (0.10)	0.15 (0.12)	C07(11)	0	0
26	Zhuoni GS	34°36.578'	103°30.695'	2670	12	R01(6) R19(3) R20(3)	0	0	C01(10) C22(2)	0.30 (0.15)	0.02 (0.03)
27	Zhangxian GS	34°50.568'	104°01.859'	2820	12	R05(6) R06(6)	0	0	C01(3) C07(9)	0.41 (0.13)	0.03 (0.03)
28	Dingxi GS	35°22.598'	104°32.665'	2350	10	R01(6) R02 R03(2) R04	0	0	C01(8) C02(2)	0.36 (0.16)	0.02 (0.03)
29	Zhengning GS	35°31.113'	108°28.908'	1444	13	R04(9) R43(4)	0	0	C01(13)	0	0
30	Yijun SN	35°20.862'	109°05.258'	1178	10	R31(10)	0	0	C01(10)	0	0
31	Ganquan SN	36°23.420'	109°24.599'	1126	11	R01(11)	0	0	C01(11)	0	0
32	Baotou IM	40°47.376'	110°17.119'	1475	12	R01 R21(3) R22(3) R23 R24 R25 R26 R27	0.44 (0.16)	0.07 (0.07)	C01(7) C23(5)	0.53 (0.08)	0.04 (0.04)
33	Fenyang SX	37°18.815'	111°34.367'	1206	10	R42(10)	0	0	C01(10)	0	0
34	Qinxian SX	36°41.004'	112°31.501'	1068	10	R27(10)	0	0	C01(10)	0	0
35	Wutai SX	38°44.929'	113°17.032'	1168	10	R01(5) R22(2) R38(3)	0.47 (0.13)	0.07 (0.07)	C01(10)	0	0
36	Fanzhi SX	39°04.970'	113°38.766'	1922	10	R01(2) R22(2) R25(2) R26 R41(3)	0.20 (0.15)	0.03 (0.05)	C01(10)	0	0
37	Laiyuan HB	39°14.307'	114°39.534'	1158	12	R03(4) R22(6) R39 R40	0	0	C01(12)	0	0
38	Zhuolu HB	39°57.829'	115°04.140'	949	14	R37(4) R38(10)	0	0	C23(14)	0	0
39	Keshiketen IM	43°30.832'	117°11.699'	1336	14	R21(6) R24(5) R30(3)	0.49 (0.09)	0.07 (0.08)	C01(12) C23(2)	0.26 (0.14)	0.02 (0.02)
40	Zhaoyang LN	41°31.966'	120°16.583'	342	10	R24(10)	0	0	C01(8) C04(2)	0.36 (0.16)	0.02 (0.03)
41	Bengxi LN	41°14.980'	124°17.625'	426	10	R01(2) R04(2) R22 R24(3) R35 R36	0.64 (0.10)	0.10 (0.09)	C01(10)	0	0
42	Huanren LN	41°23.846'	125°03.090'	363	11	R22(3) R23 R24 R26 R27 R34(4)	0.47 (0.16)	0.08 (0.08)	C03(10) C24	0.18 (0.14)	0.01 (0.02)
43	Xifeng LN	46°01.196'	124°45.903'	303	10	R31(6) R32(2) R33(2)	0.62 (0.14)	0.13 (0.11)	C01(8) C23(2)	0.36 (0.16)	0.02 (0.03)
ssp. *mongolica*					**54**	**4**	**0.58 (0.05)**	**0.19 (0.14)**	**1**	**0**	**0**
44	Altai RS	50°18.383'	087°37.732'	1330	12	R28(4) R29(8)	0	0	C26(12)	0	0
45	Tuva RS	50°34.974'	096°45.562'	1700	14	R08(4) R67(10)	0.44 (0.11)	0.07 (0.07)	C26(14)	0	0
46	Bu'erjin XJ	47°46.436'	086°49.657'	475	16	R08(16)	0	0	C26(16)	0	0

Table S2-1 (*continued*)

47 Bu'erjin XJ	47°47.725'	086°52.041'	475	12	R08(12)	0	0	C26(12)	0	0
ssp. *turkestanica*				**76**	**9**	**0.53 (0.05)**	**0.11 (0.09)**	**8**	**0.77 (0.04)**	**0.53 (0.28)**
48 Gongliu XJ	43°32.446'	082°12.602'	740	12	R08(8) R63(4)	0	0	C37(6) C38(2) C40(4)	0.67 (0.09)	0.07 (0.06)
49 Zhaosu XJ	43°05.243'	081°05.164'	1730	12	R08(9) R69(3)	0.41 (0.13)	0.06 (0.07)	C38(2) C39(8) C43(2)	0.55 (0.14)	0.08 (0.06)
50 Baicheng XJ	42°13.993'	083°14.045'	1770	10	R08(10)	0	0	C41(10)	0	0
51 Akesu XJ	41°37.258'	080°46.639'	1610	10	R08(6) R64(4)	0.53 (0.09)	0.08 (0.08)	C41(10)	0	0
52 Yingjisha XJ	38°35.026'	076°07.668'	1945	12	R08(12)	0	0	C41(12)	0	0
53 Zhada XZ	31°48.702'	078°44.890'	3115	10	R64(2) R71(2) R72(3) R73 R74 R75	0.60 (0.13)	0.16 (0.13)	C44(10)	0	0
54 Zhada XZ	31°30.577'	079°45.430'	3701	10	R64(4) R71(3) R72(3)	0.47 (0.13)	0.07 (0.07)	C45(10)	0	0
ssp. *fluviatilis*				**20**	**1**	**0**	**0**	**1**	**0**	**0**
55 Wildhaus SW	47°11.590'	009°19.786'	1130	10	R68(10)	0	0	C42(10)	0	0
56 Diemtigen SW	46°35.497'	007°28.047'	1680	10	R68(10)	0	0	C42(10)	0	0
ssp. *caucasica*				**4**	**2**	**-**	**-**	**2**	**-**	**-**
57 Dagestan RS[a]	41°25.362'	047°54.802'	880	1	R77	-	-	C46	-	-
58 Tuapse RS[b]	44°08.226'	039°07.855'	30	1	R77	-	-	C46	-	-
59 Kazbegi GE[c]	42°40.207'	044°37.310'	1900	1	R77	-	-	C48	-	-
60 Hazi Mehmet TK[d]	40°56.000'	039°43.000'	230	1	R76	-	-	C46	-	-
ssp. *carpatica*				**2**	**2**	**-**	**-**	**1**	**-**	**-**
61 Gheorghe RM[e]	45°00.000'	029°30.000'	0	1	R77	-	-	C47	-	-
62 Serpeni RM[f]	46°40.000'	027°00.000'	290	1	R78	-	-	C47	-	-
ssp. *rhamnoides*				**1**	**1**	**-**	**-**	**1**	**-**	**-**
63 Koksijde BE[g]	51°07.935'	002°41.160'	20	1	R77	-	-	C49	-	-

Abbreviation: GS, Gansu; HB, Hebei; IM, Inner Mongolia; LN, Liaoning; QH, Qinghai; SC, Sichuan; SN, Shaanxi; SX, Shanxi; XJ, Xinjiang; XZ, Xizang; YN, Yunnan; RS, Russia; SW, Switzerland; RM, Romania; BE, Belgium; GE, Georgia.

[adef] Balsgård research station, Swedish University of Agricultural Sciences (voucher number: Nos: 0402-1-36, 9835-10-19, 9897-13-5 and 9898-13-16, respectively);

[b]Herbarium of Main Botanical Garden, Moscow, Russia, collected by V.P. Chernovol (voucher number: caucasica-02K)

[c]Herbarium of Department of Botany, Swedish Museum of Natural History, Stockholm, Sweden, collected by Jens Klakenberg (voucher number: 830524-11);

[g] Collected by Filip Verloove (no voucher).

Table S2-2 List of the general ITS types grouped by TCS after collapsing sequences that differ only by ambiguous sites

General ribotypes	ITS types collapsed
R01	R02, R03, R04, R05, R06, R07, R19, R20, R21, R22, R23, R25, R30, R32, R34, R39, R40, R41, R42, R43, R49, R59, R60, R61
R08	R63
R09	R11
R10	R54
R12	R13, R14, R44, R45, R46, R48, R51, R52, R57, R58, R70
R15	R16, R17
R18	
R24	R26, R35
R27	R33
R28	R29
R31	
R36	
R37	R38
R47	
R50	
R53	
R55	
R56	
R62	
R64	R71, R73, R74
R65	R66
R67	
R68	
R69	
R72	
R75	
R76	
R77	R78

Table S2-3 Divergence times (Ma, million years ago) of major cpDNA and ITS nodes (Fig. 2-2) estimated by BEAST analyses

Nodes	cpDNA		Nodes	ITS	
	Time	95% HPD		Time	95% HPD
a1	5.46	2.23-10.10	**b1**	3.02	1.42-5.18
a2	4.65	1.91-8.60	**b2**	2.56	1.22-4.33
a3	2.46	0.94-4.73	**b3**	2.44	1.07-4.19
a4	2.05	0.72-3.86	**b4**	1.87	0.82-3.25
a5	1.45	0.53-2.82			

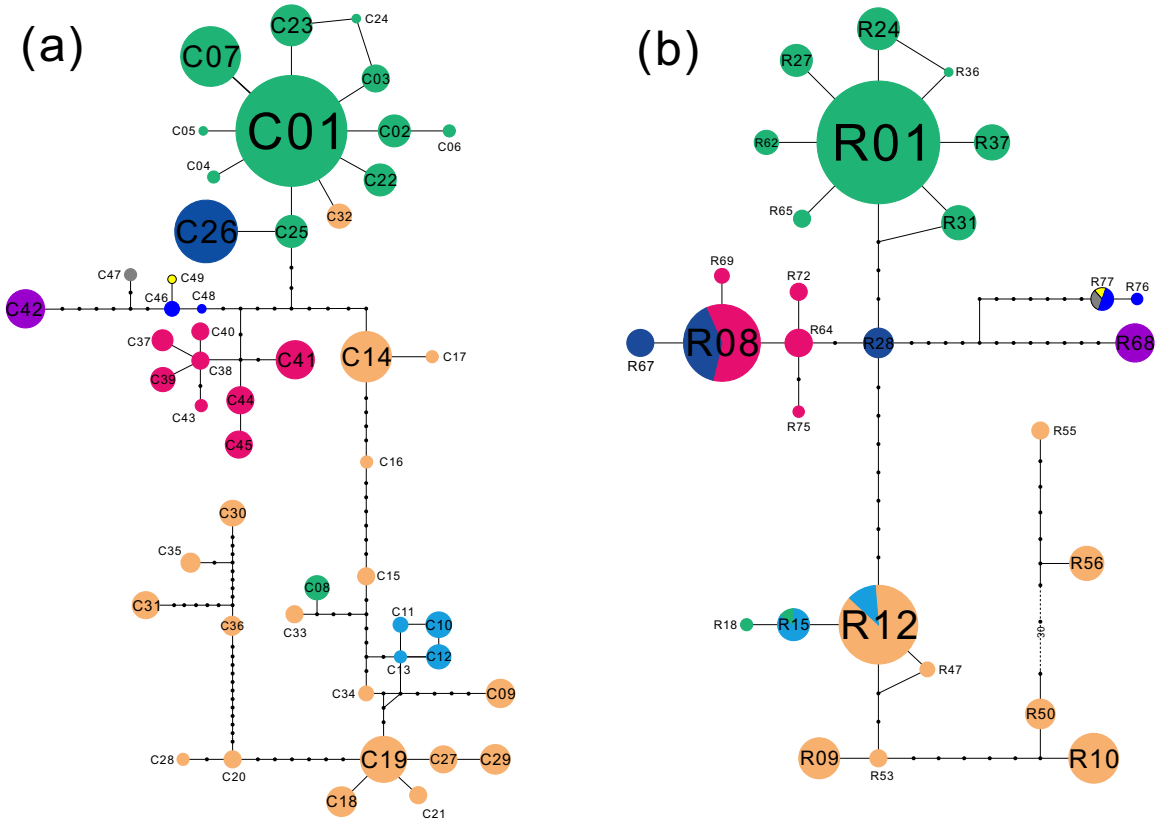


Fig. S2-1 TCS networks of (a) chlorotypes and (b) general ITS ribotypes (see Table S2-2). Circle sizes are proportional to chlorotype/ribotype frequencies within the total sample. Lines between circles represent single mutational steps and black dots represent extinct or unsampled haplotypes. Pie colors correspond to different subspecies (Fig. 2-1).

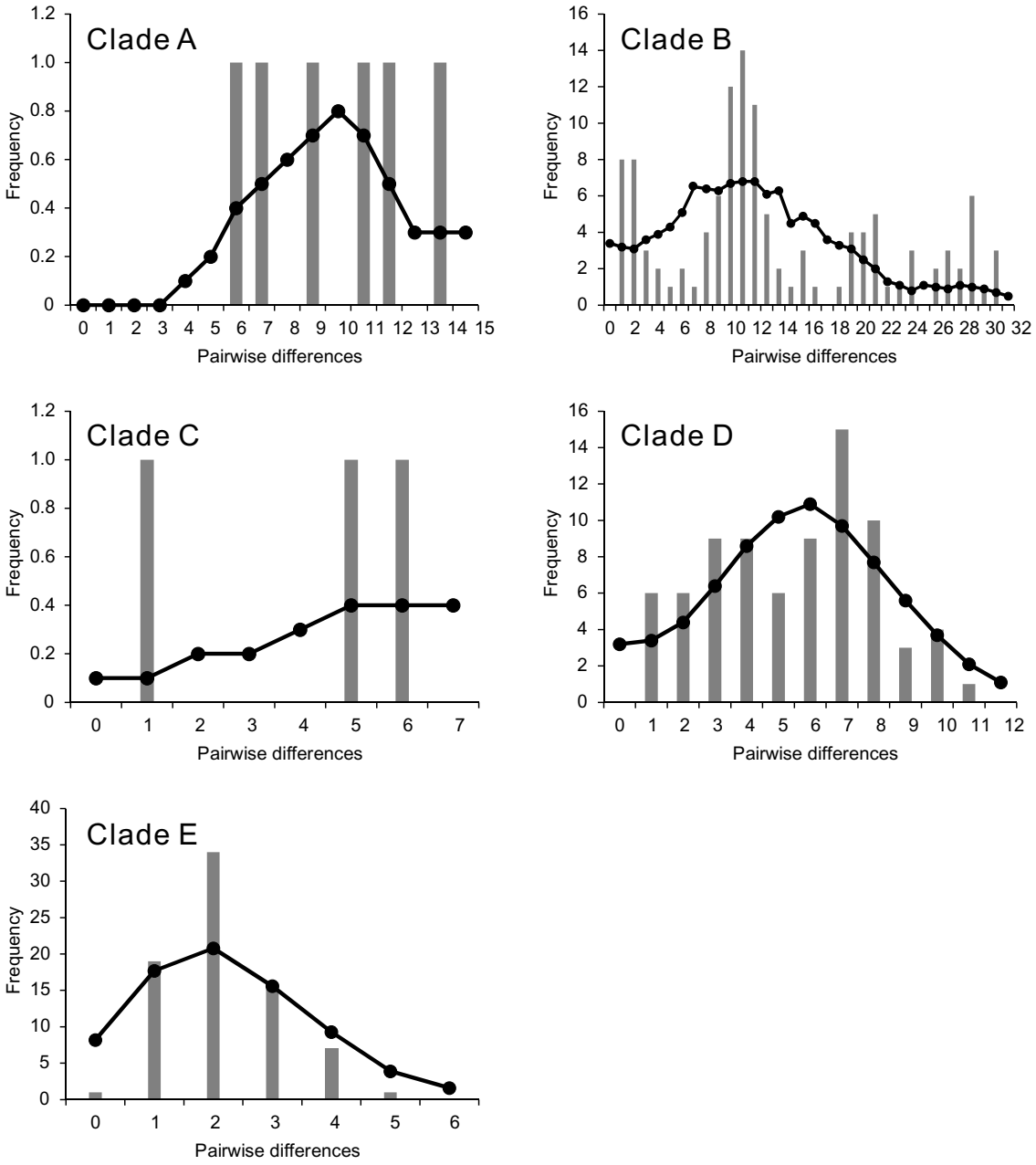


Fig. S2-2 Mismatch distribution analyses of the five cpDNA clades showing histogram of observed mismatch frequencies and best-fit curve of the sudden expansion model.

Chapter III. Climatic changes and orogenesis in the Eurasian Late Miocene: the main triggers of an expansion at a continental scale?

Abstract

Migrations from the Qinghai-Tibetan Plateau (QTP) to other temperate regions represent one of the main biogeographic patterns for Northern Hemisphere. However, the ages and routs of these migrations are largely not resolved. We aimed to reconstruct a well-resolved and dated phylogeny of *Hippophae* L. (Elaeagnaceae) and test hypothesis of a westward migration of this plant out of QTP across Eurasian mountains in the Miocene. We produced two data matrices of five chloroplast DNA (cpDNA) and five nuclear DNA (nDNA) markers for all non-hybrid taxa of *Hippophae*, each represented by three populations across the whole range of a taxon. These matrices were used to reconstruct cpDNA and nDNA phylogenies in the genus. For estimation of age of the stem node of *Hippophae* we relied on five fossil records evenly distributed across a tree of Rosales (including Elaeagnaceae and *Hippophae*). We finally used two more fossil records, cpDNA phylogeny of *Hippophae*, and age estimates from the tree of Rosales to reconstruct ages and ancestral areas of diversifications within the genus. The monophyly of *Hippophae*, all five species, and most of subspecies was strongly supported by both plastid and nuclear data sets. Both genomic trees were largely congruent. Radiation of *Hippophae* started most likely in central Himalayas/southern Tibet in the Early Miocene and gave rise to all species by the Middle Miocene. Diversification of *H. rhamnoides*

started most likely in the Late Miocene east of QTP from where this species expanded to central and western Eurasia. The crown node ages of most taxa in the genus were from the Late Pliocene/Early Pleistocene. Our findings highlight the impact of different stages in uplift of the QTP and Eurasian mountains and climatic changes in the Neogene on diversification and range shifts in highland flora on the continent. The results provide support to the idea of an immigration route for some European highland plants from their ancestral areas on QTP across central and western mountain ranges of Eurasia in the Late Miocene.

Key words: Asian highlands, ecological opportunity, homoploid hybridization, long-distance dispersal, molecular dating, Pikermian fauna, range expansion, Rosids, Tortonian, vicariance.

Introduction

The Neogene geologic processes and climatic changes had tremendous impact on evolution of biota in different regions of Eurasia (e.g. Comes & Kadereit 1998, 2003; Hewitt 2000, 2004). The Qinghai-Tibetan Plateau (QTP) was a central part of these processes (An *et al.* 2001; Guo *et al.* 2002; Li *et al.* 2001; Molnar 2005; Tang & Shen 1996; Zheng *et al.* 2000). It is also one of the most important global evolutionary hot spots (López-Pujol *et al.* 2011a, 2011b; Mittermeier *et al.* 2005; Myers *et al.* 2000). Recent phylogenetic analyses suggest that many genera of Northern Hemisphere (NH) temperate plants, which exhibit the greatest diversification in the highlands of Asia, originated in the QTP and adjacent regions, and then migrated to other NH regions where they gave rise to daughter species (e.g. Gao *et al.* 2013; Liu *et al.* 2006; Roquet *et al.* 2013; Sun *et al.* 2012; Wang *et al.* 2009b; Zhang *et al.* 2009). These studies support in general the hypotheses, which indicated the QTP as a major source of migrants to west Eurasia, especially those of highland and arid areas (Axelrod *et al.* 1996; Comes & Kadereit 2003; Wu 1988; Wulff 1943). However, ages of disjunction and migration from the QTP to west Eurasia for different ecological groups are not clear yet. Few phylogenetic analyses of taxa with Eurasian ranges have been reported (for a recent review, see Kadereit *et al.* 2008) and even less studies used dated phylogenies and ancestral area reconstructions for these taxa. Detailed information on direction and ages of migrations in Eurasia is mostly lacking (but see Gao *et al.* 2013; Jia *et al.* 2012), thus preventing analyses of impact of historical landscape and climatic changes on formation of different ecological groups of Eurasian plants.

Ecology and distribution of *Hippophae* L. are highly relevant to the problem of

migrations of NH highland plants and their evolutionary responses to the Neogene orogenesis and climatic changes across Eurasia. This shrub or small tree is diploid (2n = 24; Rousi 1965), dioecious, and wind-pollinated. Fruits of most species in the genus are juicy favouring dispersal by birds and long-distance dispersal (Lian *et al.* 1998; Rousi 1971). Roots of *Hippophae* form nitrogen-fixing nodules (Bond *et al.* 1956) and posses an efficient dual symbiosis with mycorrhiza and *Frankia* (Tian *et al.* 2002). The symbiosis enables *Hippophae* to colonize infertile and bare soil after disturbance (e.g. landslides, mountain river flash floods and dune migration). Their roots distribute rapidly and extensively, providing a non-leguminous nitrogen fixation role in surrounding soils, stabilizing riverbanks, steep slopes, and dunes. These plants are therefore early-successional pioneers with habitat engineering function for their communities. *Hippophae* can also serve as a representative of Eurasian distribution and east-central-west disjunctions for mountainous plants (Hyvönen 1996; Jia *et al.* 2012; Rousi 1971). Ranges of six species in the genus are restricted to the QTP region and adjacent areas and two of them are of hybrid origin (Bartish *et al.* 2002; Lian *et al.* 1998; Sun *et al.* 2002; Wang *et al.* 2008a). The last species, *H. rhamnoides*, contains nine subspecies (Lian *et al.* 2003; Swenson & Bartish 2002) that are distributed widely but fragmentally in eastern Asia (ssp. *yunnanensis*, *wolongensis*, and *sinensis*), central Asia (ssp. *turkestanica* and *mongolica*), Asia Minor (ssp. *caucasica*) and Europe (ssp. *carpatica*, *fluviatilis* and *rhamnoides*; see Fig. 3-1). Although monophyly of *H. rhamnoides* has been tested and supported earlier (Bartish *et al.* 2000; Jia *et al.* 2012), no comprehensive tests of monophyly of the other four non-hybrid species have been reported.

Bobrov (1962) suggested east to west migrations of *Myricaria germanica* (L.) Desv.

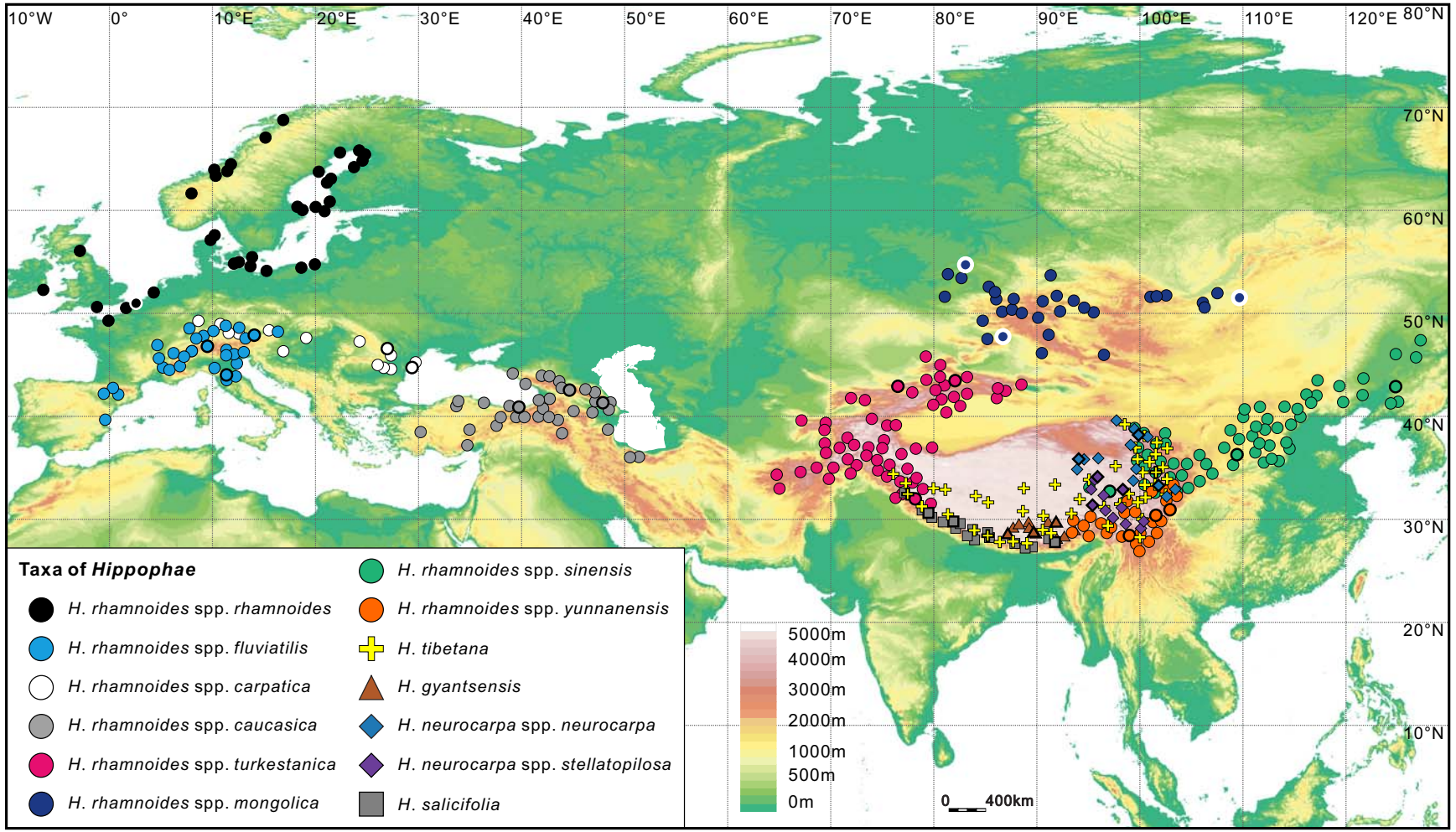


Fig. 3-1 Geographic distribution range of *Hippophae* L. in the world and the sampling sites (thick outlines; details in Table 3-1).

(Tamaricaceae) and *Hippophae rhamnoides* L. along Eurasian mountain ranges in the Late Tertiary, more specifically in the Sarmatian, which corresponds to geological age from the Middle to Late Miocene (Berrgren & Van Couvering 1974). He based this hypothesis exclusively on floristic, paleobotanic, and geologic inferences. In contrast, using phylogenetic analyses of morphological characters of most of taxa in the genus, Hyvönen (1996) suggested west to east migration followed by range fragmentation and vicariance. Jia *et al.* (2012) were the first to reconstruct ancestral areas and migration routes for all taxa in *H. rhamnoides* (using albeit a rather restricted sample of populations from central and west Eurasia) and found support for the hypothesis of the east QTP origin of the species and its migration westwards. This study suggested there was only one migration from Asia to Europe, supporting the earlier results from analyses of dominant molecular markers, which indicated a strongly supported monophyly of four western subspecies (Bartish *et al.* 2000). However, relationships within the western group could not be resolved (Bartish *et al.* 2006; Jia *et al.* 2012). Jia *et al.*(2012) also estimated ages of diversification within *H. rhamnoides* and suggested that it diversified mostly in the Quaternary. However, these age estimates should be considered as preliminary, since they were based exclusively on secondary calibrations, which can result in considerable age underestimates (Sauquet *et al.* 2012). Although the oldest fossil records of *H. rhamnoides* in Europe for long were known only from the late Pleistocene (Lang 1994), in a recent study fossil pollen grains of the species were reported from several localities from the Late Miocene, Pliocene and the early Pleistocene of southern Balkans and western Anatolia (Biltekin 2010). Thus, in view of these latest advances in paleobotany of the region it become obvious that ages of early diversifications within the genus and

age of the migration from east QTP to west Eurasia should be reconsidered.

Here we report molecular phylogenetic, dating, and ancestral area reconstructions in a thorough sample of all taxa of *Hippophae* currently accepted by Swenson & Bartish (2002), excluding those of hybrid origin (*H. goniocarpa* and *H. litangensis*). One of the primary aims of this study was to test the Borov's (1962) hypothesis of the Middle to Late Miocene migration of *H. rhamnoides* from eastern QTP to west Eurasia and his idea that mountain ranges which arose in this epoch between QTP and Europe might facilitate this migration. However, due to the unsatisfactory fossil record in the genus we opted to expand our sample to include all genera of Elaeagnaceae and all families of Rosales and test this hypothesis through reconstruction of resolved, robust, and calibrated with multiple fossil records molecular phylogenies. Phylogenetic relationships within *Hippophae* were analyzed based on a sample of three populations across the whole range of each taxon, and on a sample of five genes from each chloroplast and nuclear genomes of these plants. We believe this sampling effort and used analytical approaches are both necessary and sufficient to answer the following questions: (i) What are ages of Elaeagnaceae and its genera? (ii) Are all taxa in *Hippophae* monophyletic? (iii) What are ages and ancestral areas of all species and subspecies of *Hippophae*? (iv) What are age and route of the migration of *H. rhamnoides* from East Asia to Europe? (v) What was the impact of the Eurasian Neogene orogenesis and climatic changes on evolutionary history of *Hippophae*? We discuss later how answering these questions can contribute to better understanding of evolutionary and ecological processes in biota of Eurasian highlands.

Materials and methods

Data set of Rosales

To estimate the age of Elaeagnaceae by application of available deep primary calibration points, we used an earlier published multi-gene data set of 25 Rosales taxa (Zhang *et al.* 2011). We extended the data set by including six more species from Elaeagnaceae (*Hippophae tibetana*, *Hippophae salicifolia*, *Shepherdia argentea*, *Shepherdia canadensis*, *Elaeagnus angustifolia*, *Elaeagnus umbellata*) and one more plastid region (*trn*L-F). The final data set comprised 11 plastid (*rbc*L, *atp*B, *mat*K, the *psb*BTNH region (= 4 genes), *rpo*C2, *ndh*F, *rps*4 and *trn*L-F) and two nuclear genes (18S and 26S nrDNA) for 31 species of Rosales. Individual gene alignments were concatenated into a supermatrix of characters with unsampled values coded as missing data. We chose *Begonia* (Begoniaceae) as outgroup to root the tree of Rosales based on Zhang *et al.* (2011). Details of Genbank accession numbers are listed in Table S3-1 (Supporting information).

Data set of *Hippophae*

We sampled three populations for each of all the currently recognized non-hybrid taxa of the genus *Hippophae* (Swenson & Bartish 2002). The aim of this sampling design was to obtain representatives across whole ranges and/or main representative lineages of each taxon. The design should allow us robust tests of reciprocal monophyly and ancestral area

Table 3-1 Sampling information of *Hippophae* L. Three species from *Shepherdia* Nutt. and *Elaeagnus* L. were included

Taxon	Population	Herbarium	Voucher	Location	Collector	Year of collection
H. salicifolia	4079	IBP	PRA-00004079	Hildum, Humla, Nepal	R. Maan	2009
	09XZ040	LZU	LiuJQ-09XZ-LZT-040	Gongri, Xizang, China	J.-Q. Liu	2009
	P	AMU	P	Chamoli, Uttaranchal, India	S. Raina	2011
H. gyantsensis	06252	LZU	Liu-06252	Tingri, Xizang, China	J.-Q. Liu	2006
	06217	LZU	Liu-06217	Khangmar, Xizang, China	J.-Q. Liu	2006
	2568	LZU	2568	Lhasa, Xizang, China	J.-Q. Liu	2004
H. neurocarpa						
ssp. *neurocarpa*	YNG1	LZU	YNG1/4	Golmud, Qinghai, China	J.-Q. Liu	unknown
	1522	LZU	1522	Qilian, Qinghai, China	J.-Q. Liu	2003
	MM-31	LZU	MM-31	Jiuzhi, Qinghai, China	J.-Q. Liu	unknown
ssp. *stellatopilosa*	Ao129	LZU	Ao129	Dingqing, Xizang, China	J.-Q. Liu	2005
	QML1	LZU	QML1-2	Qumalai, Qinghai, China	J.-Q. Liu	unknown
	Ao111	LZU	Ao111	Shiqu, Sichuan, China	J.-Q. Liu	2005
H. tibetana	07138	LZU	Liu-07138	Jilong, Xizang, China	J.-Q. Liu	2007
	Henan	LZU	Henan	Henan, Qinghai, China	J.-Q. Liu	2008
	ST	AMU	ST	Spiti, Himachal Pradesh, India	S. Raina	2011
H. rhamnoides						
ssp. *yunnanensis*	06309	LZU	Liu-06309	Deqin, Yunnan, China	J.-Q. Liu	2006
	06321	LZU	Liu-06321	Daofu, Sichuan, China	J.-Q. Liu	2006
	06324	LZU	Liu-06324	Xiaojin, Sichuan, China	J.-Q. Liu	2006
ssp. *sinensis*	Ao96	LZU	Ao96	Yushu, Qinghai, China	J.-Q. Liu	2005
	05276	LZU	Liu-05276	Ganquan, Shaanxi, China	J.-Q. Liu	2005
	05179	LZU	Liu-05179	Xifeng, Liaoning, China	J.-Q. Liu	2005

Table 3-1 *(continued)*

Taxa	Population	Herbarium	Voucher	Location	Collector	Year of collection
H. rhamnoides						
ssp. *turkestanica*	GL	LZU	GL	Gongliu, Xinjiang, China	J.-Q. Liu	2005
	SKZ	AMU	SKZ	Spiti, Himachal Pradesh, India	S. Raina	2011
	12413	S	S08-12413	Almaty, Kazakhstan	L. Stenberg	2008
ssp. *mongolica*	05047	LZU	Liu-05047A	Bu'erjin, Xinjiang, China	J.-Q. Liu	2005
	6667	IBP	PRA-00006667	Novokizhiginsk, Buryatiya, Russia	I. Bartish; A. Borisyuk	2012
	6657	IBP	PRA-00006657	Berdsk, Novosibirskaya, Russia	I. Bartish; A. Borisyuk	2012
ssp. *caucasica*	0402	SUA	0402-1-36	Usukhchay, Dagestan, Russia	M. Rabadanov	1998
	21286	S	S09-21286	Kazbegi, Georgia	J. Klakenberg	1993
	9835	SUA	2898492	Hazi Mehmet, Turkey	M. Kucuk	1996
ssp. *fluviatilis*	SW1	LZU	Liu-SW1	Wildhaus, Switzerland	Y.-M. Yuan	2005
	6757	IBP	PRA-00006757	Il Mulino, Firenzuola, Italy	I. Bartish; I. Schanzer	2012
	6682	IBP	PRA-00006682	Kirchdorf an der Krems, Austria	I. Bartish; I. Schanzer	2012
ssp. *carpatica*	9897	SUA	9897-13-5	Gheorghe, Romania	P. Mladin	2007
	9898	SUA	9898-13-16	Serpeni, Romania	P. Mladin	2007
ssp. *rhamnoides*	6672	IBP	PRA-00006672	Oostduinkerke, Belgium	F. Verloove	2012
Elaeagnus triflora	IVB-29	S	IVB-29	Glen Allyn, Queensland, Australia	I. Bartish; A. Ford	2004
Elaeagnus umbellata	07078	LZU	Liu-07078	Jilong, Xizang, China	J.-Q. Liu	2007
Shepherdia argentea	6777	IBP	PRA-00006777	Canada*	I. Bartish	2011

Herbrium: IBP, Institute of Botany, Prague; LZU, Lanzhou University, China; AMU, Amity University, New Delhi; S, Museum of Natural History, Stockholm; SUA, Balsgård research station, Swedish University of Agricultural Sciences, Kristianstad.

* Cultivated in Botanical Garden of Prague City, TROJA. ACCID: 1993.00168, Canada IS 237.

reconstructions for all taxa of this genus. Geographic representation of the main lineages within taxa was inferred from literatures on cpDNA phylogeography: for *H. tibetana*, Wang *et al.* (2010) and Jia *et al.* (2011); *H. neurocarpa*, Meng *et al.* (2008); *H. gyantsensis*, Cheng *et al.* (2009); and for *H. rhamnoides* ssp. *yunnanensis* and *sinensis*, Jia *et al.* (2012). Within *H. salicifolia* and several subspecies of *H. rhamnoides* (ssp. *caucasica*, *fluviatilis*, *mongolica*, and *turkestanica*) phylogeographic relationships have not yet been resolved on informative cpDNA sequences from a comprehensive sample of populations of these taxa. We therefore selected three representative populations across geographic range of each of these taxa. Finally, *H. rhamnoides* ssp. *carpatica* and *rhamnoides* were shown to be closely related with a broad sharing of the same cpDNA haplotype among multiple populations in an earlier study (Bartish *et al.* 2002; Bartish *et al.* 2006), we thus used three populations to represent the two subspecies. Three species from the other two genera of the family Elaeagnaceae (*Elaeagnus triflora* Roxb., *E. umbellata* Thunb., and *Shepherdia argentea* (Push) Nutt.) were also used in our phylogenetic and dating analyses on *Hippophae*. Included materials, voucher information, sources are listed in Table 3-1.

We used DNeasy™ Tissue Kit (Qiagen GmbH, Hilden, Germany) to isolate total genomic DNA from silica gel dried leaves. We sequenced five chloroplast DNA (cpDNA: *trn*C^{GCA}-ycf6, *trn*D^{GUC}-*trn*T^{GGU}, *trn*L^{UAA}-*trn*F^{GAA}, *trn*S^{UGA}-*trn*fM^{CAU} and *trn*S^{GCU}-*trn*G^{UCC}) and five nuclear DNA regions (nDNA: *At*103, *G3pdh*, ITS, *Ms* and *Tpi*) for all taxa. Primers used for amplification and sequencing of these regions are provided in Table 3-2. Polymerase Chain Reactions (PCR) were performed in 25 μL reaction mixture volumes using reagents and manufacture's instruction for *Taq* DNA polymerase (Thermo

Table 3-2 Primers used for analysis of *Hippophae* L.

Locus	Primer sequence 5'–3'		Reference
*trn*C^{GCA}-ycf6	F:	CCAGTTCRAATCYGGGTG	Shaw *et al.* 2005
	R:	GCCCAAGCRAGACTTACTATATCCAT	
*trn*D^{GUC}-*trn*T^{GGU}	F:	ACCAATTGAACTACAATCCC	Demesure *et al.* 1995
	R:	CTACCACTGAGTTAAAAGGG	
*trn*L^{UAA}-*trn*F^{GAA}	F:	CGAAATCGGTAGACGCTACG	Taberlet *et al.* 1991
	R:	ATTTGAACTGGTGACACGAG	
*trn*S^{UGA}-*trn*fM^{CAU}	F:	GAGAGAGAGGGATTCGAACC	Demesure *et al.* 1995
	R:	CATAACCTTGAGGTCACGGG	
*trn*S^{GCU}-*trn*G^{UCC}	F:	GCCGCTTTAGTCCACTCAGC	Hamilton 1999
	R:	GAACGAATCACACTTTTACCAC	
*At*103	F:	CTTCAAGCCMAAGTTCATCTTCTA	Li *et al.* 2008
	R:	TTGGCAATCATTGAGGTACATNGTMACATA	
G3pdh	F:	GATAGATTTGGAATTGTTGAGG	Strand *et al.* 1997
	R:	AAGCAATTCCAGCCTTGG	
ITS	F:	AGAAGTCGTAACAAGGTTTCCGTAGG	White *et al.* 1990
	R:	TCCTCCGCTTATTGATATGC	
Ms	F:	GGAAGATGRTCATCAAYGCNCTYAAYTC	Lewis & Doyle 2001
	R:	GTCTTNACRTAGCTGAADATRTARTCCC	
Tpi	F:	AAGGTCATTGCATGTGTTGG	Strand *et al.* 1997
	R:	CTTTACCAGTTCCAATAGCCC	

Scientific, UK) in a Mastercycler (Eppendorf, Germany). PCR products were purified using QIAquick PCR Purification Kit (Qiagen, Germany) and sequencing reactions were conducted with ABI Prism Bigdye^{TM} Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Sequences were obtained using an ABI 3130 Genetic Analyzer (Applied Biosystems). CLUSTALX version 1.81 (Thompson *et al.* 1997) was used to align produced sequences, and MEGA version 4 (Tamura *et al.* 2007) was used to adjust them manually. Chloroplast and nuclear sequences were concatenated separately to make two data sets. Newly identified sequences were deposited in GenBank (Table S3-2). We chose *Rhamnus davurica* Pall. (Rhamnaceae) as outgroup to root the tree, based on Zhang *et al.* (2011) and our phylogenies of Rosales (see Results).

Available outgroup sequences were taken from Genbank. GAPCODER (Young & Healy 2003) was used to edit indels as separate characters for inclusion in Bayesian analyses.

Phylogenetic analyses

Partitioned Bayesian and Maximum Likelihood (ML) analyses were conducted on the three concatenated data sets. Bayesian inference was performed in MRBAYES version 3.2.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). We determined the best fitting model of sequence evolution for each individual region using the Akaike Information Criterion (Akaike 1974) as employed in JMODELTEST version 0.1.1 (Guindon & Gascuel 2003; Posada 2008). Two independent runs with one cold and three incrementally heated Monte Carlo Markov chains (MCMCs) were run for 5,000,000 generations, with trees sampled every 500th generation. A standard discrete model (Lewis 2001) was applied to the indel matrix. Model parameters were unlinked across partitions. We discarded the first 2,500 trees out of the 10,001 trees as burn-ins and used the remaining trees to build a 50% majority rule consensus tree.

ML analyses were performed using RAxML version 7.2.7 (Stamatakis 2006). A separate General Time Reversible + Gamma model (GTR + G) of nucleotide substitution was specified for each data partition, and 500 independent searches were conducted. Support values for nodes in the phylogenetic tree were estimated across 1000 pseudoreplicates using the GTR + CAT model (Stamatakis *et al.* 2008), and mapped thereafter onto the best-scoring tree from the 500 independent searches.

Selection of fossils for calibration

We constrained the maximum age of the crown node of Rosales using maximum age estimate for this node (103 Ma) from Wang *et al.* (2009a). This study was based on a comprehensive sample of families from Rosids (represented by 104 species), sequences of 12 genes (10 cpDNA and two nDNA), and on seven fossil records to calibrate the tree. This sampling effort indicates that estimates of ages of diversification among the sampled orders, including the stem and crown nodes of Rosales, can be fairly robust. For calibration of early diversification within Rosales we selected from paleobotanical literature five fossil records, which can define the minimum ages of the stem nodes of respective taxa and clades in our tree (Fig. 3-4; Table S3-3). *Celtis aspera* (Newburry) Manchester, Akhmetiev, and Kodrul was selected based on well-preserved endocarps and leaves from numerous localities in North America and Asia in the Late Paleocene and leaves throughout the Paleocene in North America (Manchester *et al.* 2002). This species was first described as *Viburnum asperum* Newberry, but later reclassified to *Celtis* on the basis of apomorphies in leaf and fruit characters from the same specimens (Manchester *et al.* 2002). We used the Late Paleocene records of this fossil species in our analyses to calibrate the stem node of *Celtis*. *Triorites minutipori* Muller was selected based on pollen from Turonian-Senonian of Malaysia (Muller 1968), which was considered later as a reliable representative of Ulmaceae (Muller 1981). This fossil was used therefore to calibrate the stem node of Ulmaceae. *Paliurus clarnensis* Burge & Manchester has a set of morphological apomorphies to assign this fossil species reliably to extant *Paliurus* (Burge & Manchester 2008). We calibrated by this fossil the stem node of *Hovenia*,

which in our sample represents the crown node of Paliureae (Islam & Simmons 2006; Richardson *et al.* 2000). Achenes of *Ficus* L. were described from the Early Eocene London Clay Flora by Chandler (1962) and can be used to calibrate the stem node of *Ficus* in our sample. Leaves of *Shepherdia weaveri* Becker were described from the Late Eocene flora of southwestern Montana (Becker 1960) and recently included into a revised list of paleotaxa from this locality by Lielke *et al.* (2012), who also provided revised geochronology for the locality. We used this fossil species and age of its geological stratum to calibrate the stem node of *Shepherdia* in our sample. Finally, a recent report of fossil pollen grains of *Hippophae rhamnoides* from the Late Miocene of Anatolia and south of the Balkan peninsula documents the earliest records of the genus in Asia Minor and Europe (Biltekin 2010). The fossil pollen cannot be reliably assigned to a subspecies within *H. rhamnoides* due to the lack of diagnostic characters (Sorsa 1971). Nevertheless, molecular phylogenetic analyses (Bartish *et al.* 2000; Jia *et al.* 2012) and ancestral area reconstructions (Jia *et al.* 2012) suggested that geographic localities of these records can represent an ancestral area of the strongly supported monophyletic group of four western subspecies within the species. We used therefore this fossil record to calibrate the stem node of the clade, which includes *H. rhamnoides* ssp. *carpatica*, *caucasica*, *fluviatilis*, and *rhamnoides* (Fig. 3-2; Table S3-3).

Divergence time estimates

We used both penalized likelihood (PL) and Bayesian method in R8S version 1.71 (Sanderson 2002, 2003) and BEAST version 1.7.5 (Drummond *et al.* 2002; Drummond &

Rambaut 2007), respectively, to infer divergence times. For PL analyses, the ML tree with branch lengths (phylogram) obtained with RAxML was used. The outgroup was pruned from the tree prior to all analyses. The smoothing parameter (λ) was determined by cross-validation analysis. We used the truncated-Newton (TN) algorithm as recommended for PL in R8S manual and chose the additive penalty over the log penalty due to our balanced calibrations in the tree. All fossil constraints were applied as minimum ages, while secondary calibrations were enforced as maximum age constraints. Confidence intervals for the divergence date estimates were obtained in a bootstrap-based approach. We used the best-scoring ML tree as a fixed topology and estimated sets of branch lengths from 1000 bootstrap replicates using the software RAxML (Sauquet *et al.* 2012). Divergence dates were estimated from the resulting 1000 trees using the software R8S. Settings were as described above, except that the smoothing parameter was fixed to the value obtained for the original data set.

For Bayesian analyses, two different uncorrelated lognormal (UCLN) relaxed clock models with fossils treated as being drawn from either a uniform distribution (UCLN-uniform) or a lognormal distribution (UCLN-lognormal) were implemented. For uniform priors, fossil constraints were implemented as hard minimum bounds. For lognormal priors, the ages of the fossils were set as offset values with Log (mean) = 0 and Log (standard deviation) = 1. Age ranges for secondary calibrations were all implemented as uniform priors. The optimal model of molecular evolution selected in JMODELTEST was specified for each partition, and a Yule prior was specified for the tree. We initiated MCMC analyses from ultrametric starting tree with branch lengths that satisfied the calibration constraints. We sampled all parameter once every 1000 steps from 10^7

MCMC steps, with the first 25% of samples discarded as burn-in. The program TRACER version 1.5.0 (Rambaut & Drummond 2007) was used to examine convergence of chains to the stationary distribution. Trees then were compiled into a maximum clade credibility (MCC) tree using TREEANNOTATOR version 1.7.5 (Drummond *et al.* 2012) to display mean node ages and highest posterior density (HPD) intervals at 95% (upper and lower) for each node.

Ancestral area reconstructions

We ran Bayes-DIVA analyses using RASP version 2.0b (Reconstruct Ancestral State in Phylogenies; Yu *et al.* 2010, 2011) to infer the biogeographic history of *Hippophae* based on the phylogeny constructed from cpDNA sequences. In this analysis we defined nine distribution regions for the sampled populations according to the distribution characteristics of *Hippophae* and the floristic regionalization of China (Wu & Wu 1998) and the world (Takhtajan 1986): (A) the south QTP (the central Himalayas and southern Tibet); (B) the east QTP (the Hengduan Mts.); (C) the north QTP (Tangut area); (D) the west QTP (including Tian Shan Mts., Pamir Mts., Hindu Kush Mts. and Karakoram Range); (E) northern China; (F) the Mongolian Plateau (including Altai-Sayan Mts., Tarbagatay Range and South Siberian Plain); (G) the Caucasus and Anatolian highlands; (H) Europe north of Alps (including Carpathians) and (I) the Alps and Apennines (Figs. 3-1, 3-5). We loaded 10,001 trees previously produced in BEAST and chose the F81 model for the Bayesian MCMC analyses, allowing for different rates of change among ancestral areas.

Results

Phylogenetic relationships of Rosales and Elaeagnaceae

Based on the data set consisting of 18335 aligned nucleotide sites for the 13 gene partitions (Table S3-4), we reconstructed the phylogeny of Rosales. ML and BI approaches yielded an identical topology with slightly different support values (Fig. S3-1). All the families within Rosales were monophyletic, and the relationships among them were well resolved and supported – in complete accord with the previous inferences reported by Zhang *et al.* (2011). Rhamnaceae were placed as sister to a moderately supported clade of Elaeagnaceae and Barbeyaceae plus Dirachmaceae. Within the family Elaeagnaceae, the three genera were each strongly supported as monophyletic, and *Elaeagnus* were sister to the well-supported clade of *Hippophae* and *Shepherdia*.

Phylogenetic relationships of *Hippophae* based on cpDNA

The combined cpDNA matrix of *Hippophae* included a total of 5462 aligned base pairs (Table 3-3). The ML analysis with RAxML yielded a well-resolved phylogenetic tree (Fig. 3-2). The monophyly of *Hippophae* was strongly supported (BP = 92%). All taxa of *Hippophae* except for *H. rhamnoides* ssp. *sinensis* and *yunnanensis* were recovered as monophyletic with strong statistical support (BP > 90% for all taxa except *H. neurocarpa* ssp. *stellatopilosa* with BP = 63%). *H. gyantsensis*, *H. salicifolia* and *H. nerocarpa* formed a clade with a low support (BP = 62%), which was sister to the moderately (BP =

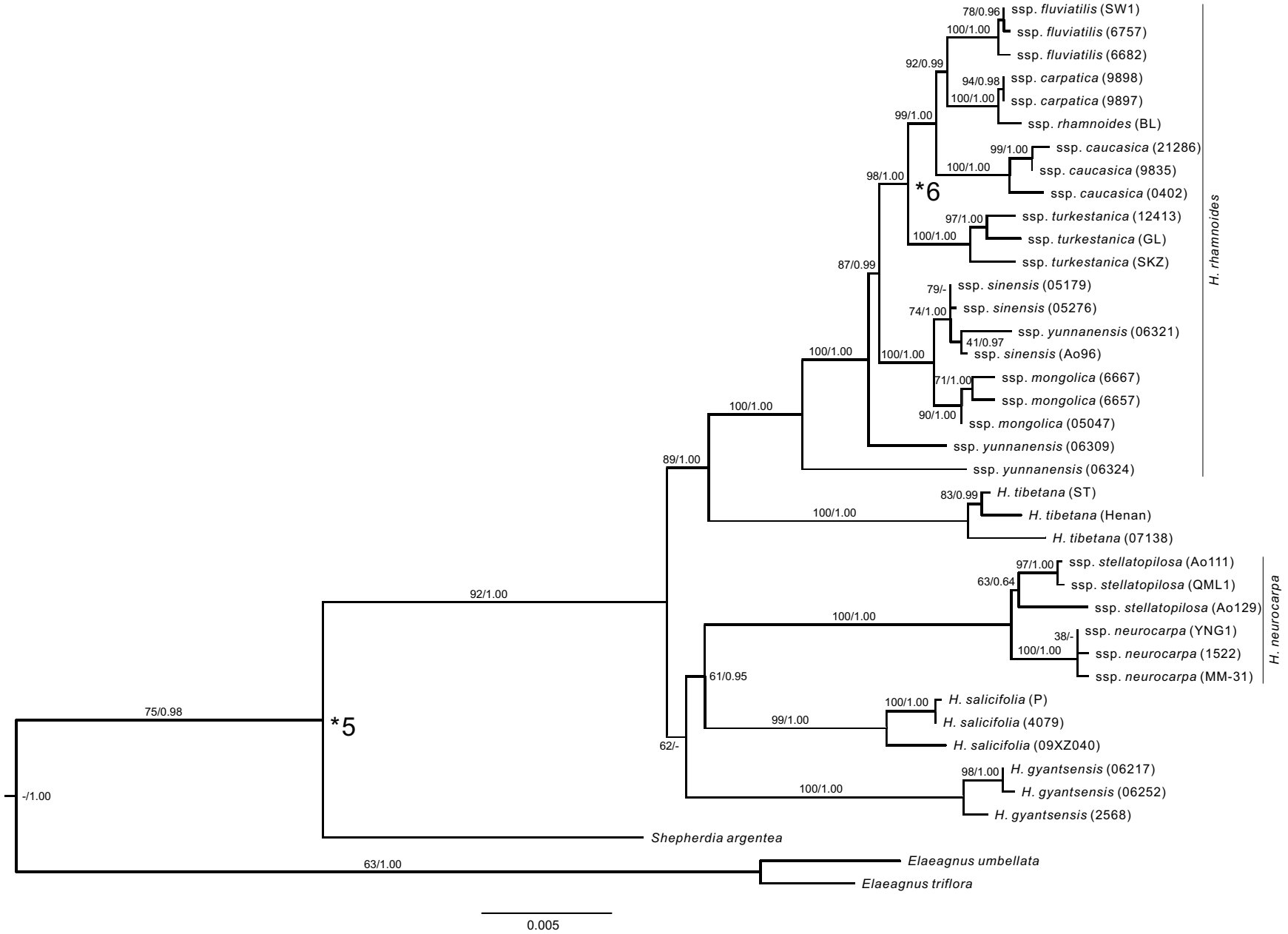


Fig. 3-2 The ML tree of *Hippophae* L. from the combined cpDNA fragments. Support values (ML Bootstrap/Bayesian posterior probability) are provided at nodes. The two fossil calibration points used for dating analyses of *Hippophae* are shown with asterisks (see Table S3-3).

89%) supported clade composed of *H. tibetana* and *H. rhamnoides*. Within *H. rhamnoides*, two lineages of ssp. *yunnanensis* created a basal grade with the clade including the other taxa. The latter clade was created by two sub-clades, both receiving a strong support (BP > 98%). One of these sub-clades was created by ssp. *sinensis* and *mongolica,* the second one by a grade of ssp. *turkestanica* and the other Asia Minor/European subspecies (ssp. *caucasica*, *carpatica*/*rhamnoides*, and *fluviatilis*, all listed here as ordered in the grade). Tree topology recovered from Bayesian analysis was largely congruent with the ML tree. The main discrepancy regards the unresolved relationships between *H. gyantsensis* and the other two clades containing *H. neurocarpa*/*H. salicifolia* and *H. tibetana*/*H. rhamnoides*, respectively.

Table 3-3 Gene/partition characteristics of the *Hippophae* L. datasets

Data set	Partition	No. of Taxa	Aligned lengths	Variable sites	Parsimony-informative sites	Model selected by AIC
cpDNA						
	*trn*C-ycf6	39	928	129	96	TPM1uf + G
	*trn*D-T	39	1396	144	99	TPM1uf + I + G
	*trn*L-F	40	959	198	54	TIM1 + G
	*trn*S-fM	38	1441	187	133	TIM1 + G
	*trn*S-G	39	738	88	60	TrN + G
	Total	40	5462	746	442	
nDNA						
	*At*103	39	524	65	46	HKY + I
	G3pdh	39	735	42	29	HKY + I
	ITS	40	714	276	174	GTR + I
	Ms	38	610	69	53	TPM1uf + G
	Tpi	36	375	40	22	TPM3uf + I
	Total	40	2958	492	324	

Phylogenetic relationships of *Hippophae* based on nDNA

The aligned length of the combined nDNA matrix of *Hippophae* was 2958 base pairs

(Table 3-3). The monophyly of *Hippophae* was strongly supported by both the ML and Bayesian methods (BP = 99%, PP = 1.00; Fig. 3-3). The clade of *H. gyantsensis*, *H. nerocarpa,* and *H. salicifolia* received strong support (BP = 92%; PP = 1.00), in which *H. salicifolia* was sister to the sub-clade of the other two taxa, and all species were monophyletic with strong support (BP > 98%; PP = 1.00). *H. tibetana* was placed sister to *H. rhamnoides* in the moderately (BP = 84%) to strongly (PP = 1.00) supported clade where each species is monophyletic with strong support (BP = 100%; PP = 1.00). Within *H. rhamnoides*, we found incongruence with cpDNA phylogenetic placement of *H. rhamnoides* ssp. *mongolica*: it formed a strongly supported (BP = 90%; PP = 1.00) clade within ssp. *turkestanica*. Additionally, monophyly of ssp. *caucasica*, *carpatica*, and *yunnanensis* was not supported.

Divergence time estimates within Rosales and Elaeagnaceae

In general, PL and two Bayesian analyses yielded similar estimates of ages with largely overlapping confidence intervals (Tables S3-5 and S3-6). We thus report the divergence times estimated under the Bayesian UCLN-lognormal model in the text. Age of the stem node for Rhamnaceae was estimated to 91.1 (95%HPD: 87.5–94.0) Ma, and ages of the stem and crown nodes for Elaeagnaceae were estimated to 89.3 (95%HPD: 85.7–93.0) and 40.6 (95%HPD: 36.9–44.0) Ma, respectively (Fig. 3-4; Table S3-5). We enforced 44.1 Ma (upper 95%HPD interval under the Bayesian UCLN-uniform model) as maximum age of the crown node of Elaeagnaceae in the dating analyses of *Hippophae*.

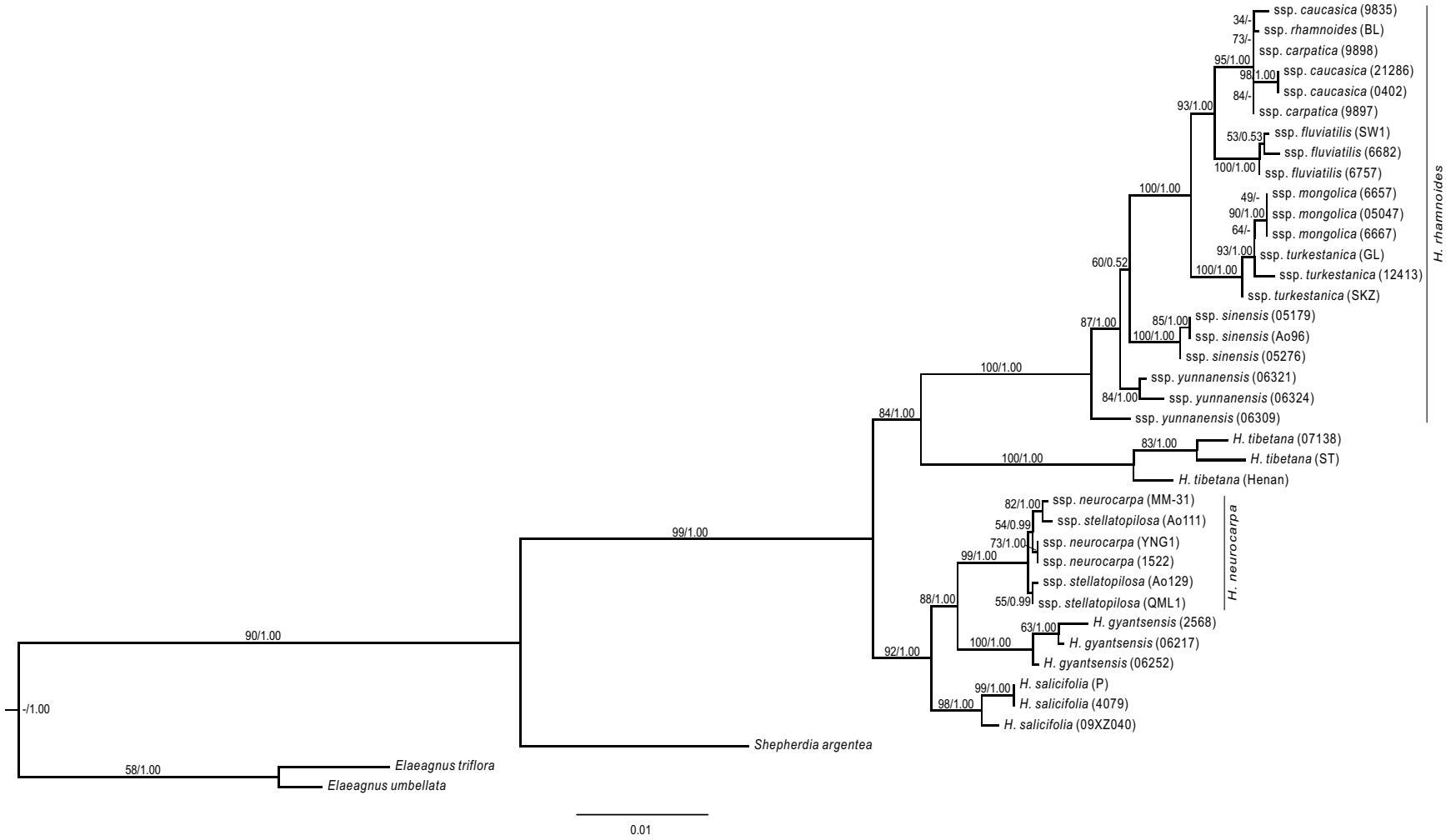


Fig. 3-3 The ML tree of *Hippophae* L. from the combined nDNA fragments. Support values (ML Bootstrap/Bayesian posterior probability) are provided at nodes.

Ancestral area reconstructions and estimation of ages of dispersal/vicariance within *Hippophae*

According to our age estimates, all species within *Hippophae* had diversified in the Early Miocene (21.2–17.6 Ma; Fig. 3-5; Table S3-6). RASP analyses based on cpDNA provided strong evidence that the ancestral area for *Hippophae* was in the south QTP with 81.1% probability (Fig. 3-5; Table S3-6). The ancestral areas of *H. neurocarpa* and *H. tibetana* optimized to the east QTP (87.4% probability) and the south QTP (86.4% probability), respectively. Additionally, results of our dating analyses suggested that ages of diversification within most of taxa in the genus are not older than the Late Pliocene or Pleistocene (3.6–0.2 Ma; Fig. 3-5; Table S3-6).

Diversification within *H. rhamnoides* started in the Middle Miocene and proceeded to the Early Pliocene (12.9–3.7 Ma). This species had a relatively high probability of ancestral area (PAA; 96.7%) in the east QTP (area B), from where several lineages had dispersed west to the west QTP (dispersal 3: 8.8 Ma; PAA_B = 94.4%), north to the Mongolian Plateau (dispersal 7: 3.7 Ma; PAA_B = 88.5%), and northeast to Northern China (dispersal 8: 1.1 Ma; PAA_B = 93.3%) (Fig. 3-5; Table 3-4). Subsequent westward dispersals out of the west QTP (dispersal 4: PAA_D = 53.6%) and Asia Minor (dispersal 5: PAA_G = 69.9%) followed by vicariance events since 7.6 Ma and 5.8 Ma, respectively, were also indicated. Finally, dispersals from Carpathians (H) to Alps (I) or in the opposite direction (both directions have almost equal probability in our reconstructions: PAA_H = 44.2% and PAA_I = 37.4, respectively) were followed by vicariance events since 4.6 Ma.

Table 3-4 Summary statistics of the 13 dispersal events detected in the *Hippophae* phylogeny

Dispersal	Area of origin				Area of Migration				Age (Ma)	
	α	%	β	%	α	%	β	%	Mean	95% HPD
1	A	52.7	B	31.8	B	96.7	AB	2.6	18.4	(15.7–21.5)
2	A	67.0	B	15.0	B	87.4	AB	4.3	17.6	(14.5–20.5)
3	B	94.4	BD	1.4	D	53.6	B	27.3	8.8	(7.8–10.0)
4	D	53.6	B	27.3	G	69.9	I	7.4	7.6	(7.2–8.2)
5	G	69.9	I	7.4	H	44.2	I	37.4	5.8	(4.7–7.1)
6	H	44.2	I	37.4	H	98.9	I	98.6	4.6	(3.2–6.0)
7	B	88.5	BF	6.7	F	93.3	BF	5.6	3.7	(2.4–5.1)
8	B	93.3	BE	4.7	E	94.5	BE	4.7	1.1	(0.3–2.0)
9	A	86.4	B	2.8	C	43.1	D	28.6	3.6	(2.1–5.3)
10	C	43.1	D	28.6	C	100.0	D	100.0	2.0	(0.8–3.3)
11	A	75.9	AD	11.0	D	97.4	AD	2.2	2.5	(1.2–4.0)
12	B	79.0	BC	11.7	C	95.7	BC	3.7	0.7	(0.2–1.3)
13	B	84.6	BC	10.8	C	100.0	–	–	0.5	(0.1–1.1)

Dispersal numbers refer to Fig. 3-4. Results from the ancestral area reconstructions are reported for the first (α) and second (β) most frequent areas with their frequency in percent. Age estimates are derived from Bayesian relaxed clock analysis treating priors on fossils as being drawn from a lognormal distribution (UCLN-lognormal).

Discussion

Ages of differentiation within Rosales

Diversification of Barbeyaceae, Dirachmaceae, Elaeagnaceae, and Rhamnaceae

Both phylogenetic reconstructions of Zhang *et al.* (2011) and this study (Fig. 3-4; Fig. S3-1) indicated that Rhamnaceae are sister to the clade of Barbeyaceae, Dirachmaceae, and Elaeagnaceae, and that diversification of these families was relatively fast. In the absence of known to us fossils in Barbeyaceae and Dirachmaceae, and a rather restricted list of fossil records in Elaeagnaceae, reliable age estimation of the stem node of Rhamnaceae is important for understanding the early evolution in these four families.

Fortunately, this large family has a relatively comprehensive fossil record, in which flowers (Calvillo-Canadell & Cevallos-Ferriz 2007; Correa *et al.* 2010), fruits (Burge & Manchester 2008), wood (Wheeler & Meyer 2012) and leaves (Burge & Manchester 2008; Correa *et al.* 2010; Manchester *et al.* 2002) are well represented in the Late Cretaceous and Paleogene. One of these fossils, a fruit assigned to *Paliurus* (Burge & Manchester 2008), was used in our analyses to constrain the stem node of *Hovenia*. This node corresponds in our tree of Rosales to the crown node of Paliureae *sensu* Richardson *et al.* (2000). Richardson *et al.* (2004) based their age estimations within Rhamnaceae on a secondary calibration of the stem node of this family (62–64 Ma) from earlier analysis of Wikström *et al.* (2001). Our estimate for the stem node of Rhamnaceae is considerably older (87.5–94.0 Ma). There are two lines of evidence indicating that the values in Wikström *et al.* (2001) and Richardson *et al.* (2004) for ages of diversifications in Rhamnaceae may be underestimates. First, within Rhamnaceae the age of the *Paliurus* fossil is in the Middle Eocene (Burge & Manchester 2008), so it should be older than 41.2 Ma (Walker *et al.* 2012). However, Richardson *et al.* (2004) reported age estimate of the crown node of Paliureae as 30.6–31.6 ± 3.1 Ma old, thus considerably younger than the minimum age of the node indicated by the fossil record. Second, outside Rhamnaceae Wang *et al.* (2009a) reported the mean age for the crown node of Rosales to be 97 Ma, which is at least 20 Ma older than the correspondent estimate for the same node by Wikström *et al.* (2001). Our findings are therefore more in line with those of Wang *et al.* (2009a) and suggest an older ages for the stem and crown nodes of Rhamnaceae. The reason for discrepancy with estimates of ages within Rhamnaceae from Richardson *et al.* (2004) can be their use of secondary constrains for calibration of the tree, which often

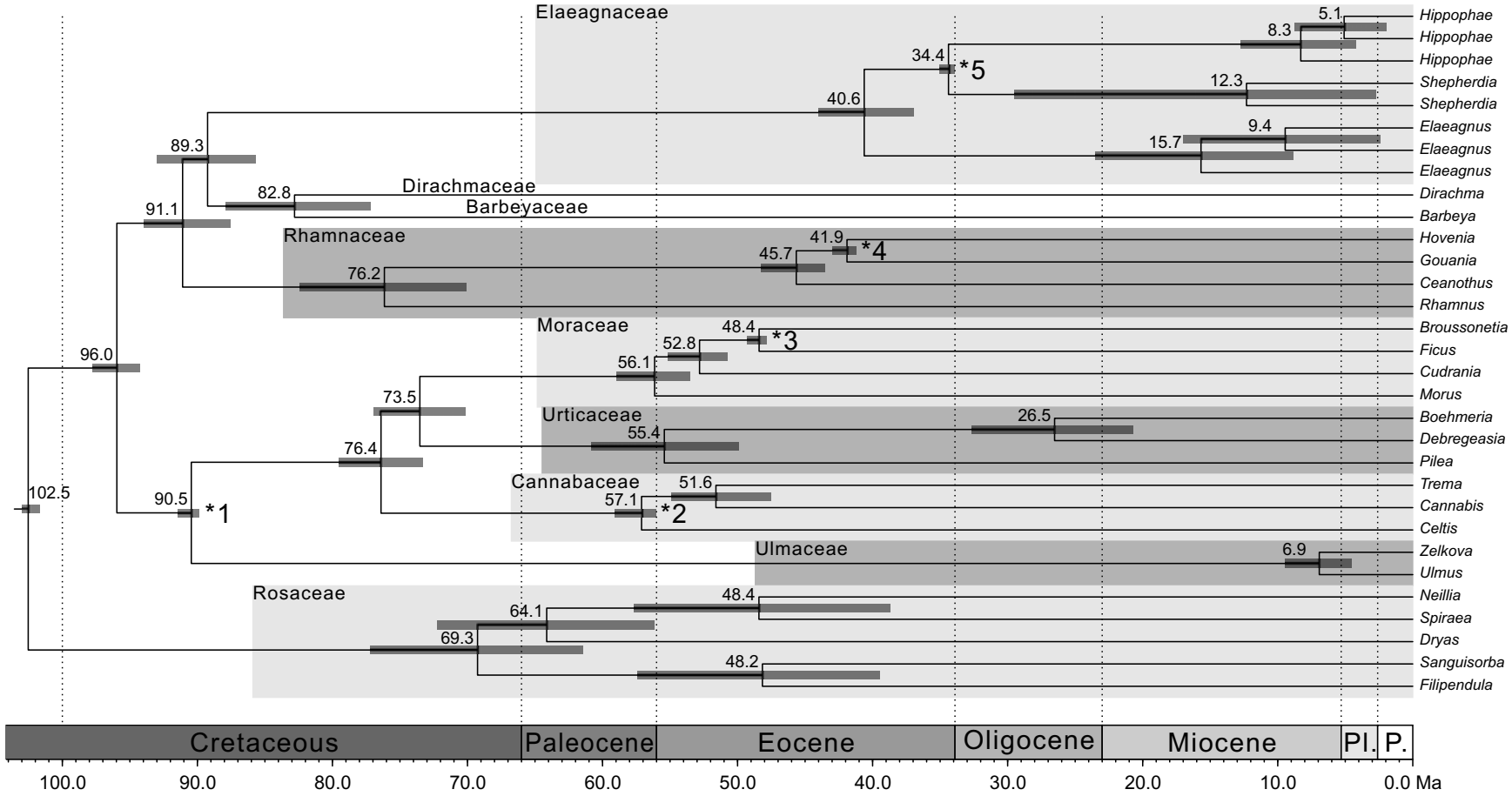


Fig. 3-4 Chronogram of divergence times within Rosales derived from Bayesian relaxed clock analysis treating priors on fossils as being drawn from a lognormal distribution (UCLN-lognormal). Mean ages of nodes are shown, with horizontal dark grey bars indicating the 95% highest posterior density intervals. The used five fossil calibration points are indicated with asterisks (see Table S3-3). Pl. = Pliocene; P. = Pleistocene.

results in underestimation of ages (Sauquet *et al.* 2012).

Our results also provide the first estimates for the stem and crown nodes of Elaeagnaceae (89.3 and 40.6 Ma, respectively), and stem node of Barbeyaceae and Dirachmaceae (82.8 Ma). We note that the oldest known to us fossil record of Elaeagnaceae (pollen of *Elaeagnacites* sp.) is from the Taizhou Formation in North Jiangsu Basin (Song & Qian 1989), which in China corresponds to Maastrichtian (Song *et al.* 2004). The age estimate for this family in our analyses thus predates its earliest fossil record by some 25–30 Ma, indicating a considerable gap in the fossil record, similar to many other families of angiosperms (Crepet *et al.* 2004).

Diversification of other families in Rosales

Age estimates for stem nodes of other families of Rosales suggest all of them had already been differentiated early in the Late Cretaceous (Fig. 3-4; Table S3-4). Differentiation of Rosaceae in the Albian has been suggested by computations of Wang *et al.* (2009a). At least the Turonian age of Ulmaceae has been indicated by pollen records of this family (Muller 1968). Differentiation in the Cenomanian for sister Moraceae and Urticaceae has been estimated by Zerega *et al.* (2005). Our analyses are mostly congruent with these inferences. We also add age estimate for the stem node of Cannabaceae/Celtidaceae clade (76.4 Ma), thus completing the set of the stem node estimates for all families in Rosales.

Phylogeny and systematics of *Hippophae*

Monophyly of species and phylogenetic relationships among them

We reconstructed the phylogenetic relationships of *Hippophae* based on thorough sampling across lineages/distribution ranges and a relatively large amount of sequence data of both nDNA and cpDNA. The genus was strongly supported as monophyletic and can be divided into two distinct sub-clades: one comprising *H. gyantsensis*, *H. salicifolia* and *H. nerocarpa* (all monophyletic with strong support in both data sets and both phylogenetic methods, BP > 98%, PP = 1.00) and the other comprising *H. tibetana* and *H. rhamnoides* (each monophyletic with strong support in both data sets and both phylogenetic methods, BP = 100%, PP = 1.00) (Figs. 3-2, 3-3). These results are not in agreement with the taxonomic classification of Lian (1988) that divided the genus into two sections based on fruit and seed characteristics: sect. *Hippophae* (*H. rhamnoides* and *H. salicifolia*) with carpodermis separated from seed coat and seed surface smooth and shining, and sect. *Gyantsensis* (*H. gyantsensis*, *H. neurocarpa*, *H. goniocarpa* and *H. tibetana*) with carpodermis adhering to seed coat and seed surface rough and dull. Indeed, the circumscription of these two sections had been rejected by two earlier phylogenetic analyses based on ITS sequences (Sun *et al.* 2002) and a combined data set of cpDNA RFLPs and morphological characters (Bartish *et al.* 2002). However, phylogenetic placements of *H. tibetana* was incongruent in these two studies. Reconstructions of Sun *et al.* (2002) placed *H. tibetana* sister to the remaining taxa consisting of two sub-clades, of which one includes subspecies of *H. rhamnoides* only. However, reconstructions of Bartish *et al.* (2002) grouped *H. tibetana* in a weekly supported sub-clade with *H. gyantsensis*, *H. salicifolia* and *H. neurocarpa*. In present analyses, the sister relationship between *H. tibetana* and *H. rhamnoides* was revealed for the first time and received a strong Bayesian (PP = 1.00) and moderate bootstrap (BP > 84%) support in analyses of

both data sets. The relationships among the three species within the other sub-clade are still poorly resolved and incongruent between the two data sets. This could result from hybridization after secondary contacts between some of the species or incomplete lineage sorting (Buckley *et al.* 2006; Doyle 1992; Liu & Pearl 2007; Maddison 1997; Pelser *et al.* 2010). For example, a secondary contact and pollen-mediated gene flow between *H. gyantsensis* and *H. neurocarpa*, followed by vicariance at some ancient period in past, could result in the observed pattern. We recommend avoiding implementation of any subgeneric ranks in *Hippophae* until relationships of *H. gyantsensis* with other two species in this sub-clade are better understood, and morphological synapomorphies are identified to define sections within the genus in accordance with robust phylogenetic relationships among all species.

Systematic problems in Hippophae rhamnoides

An earlier study on *H. rhamnoides* using a representative sample of populations of ssp. *yunnanensis* indicated this taxon as paraphyletic, with two of its lineages forming a basal grade to the remaining subspecies (Jia *et al.* 2012). This result was confirmed here, so the taxonomy of ssp. *yunnanensis* obviously needs a revision and some cryptic taxa are in need of formal description, if other names in the species should be preserved.

The four subspecies in Europe and Asia Minor (ssp. *carpatica*, *caucasica*, *rhamnoides* and *fluviatilis*) comprise a monophylogenetic lineage (Bartish *et al.* 2002; Bartish *et al.* 2006; Jia *et al.* 2012), which was sister to ssp. *turkestanica* in the Central Asia (Bartish *et al.* 2002; Jia *et al.* 2012). Our present results from both data sets are in good agreement

with the earlier reconstructions. Additionally, ssp. *rhamnoides* was shown to be closely related to ssp. *carpatica* with a broad sharing of RAPD markers (Bartish *et al.* 2002; Bartish *et al.* 2006) and cpDNA haplotypes (Bartish *et al.* 2006) among populations of these taxa. In contrast, ssp. *rhamnoides* was grouped in a clade with ssp. *caucasica* in the cpDNA phylogeny by Jia *et al.* (2012). The present cpDNA phylogeny supports the close relationship of ssp. *rhamnoides* and *carpatica*, with ssp. *fluviatilis* sister to this clade. Hence, the most likely incorrect placement of the specimen of ssp. *rhamnoides* in the earlier study (Jia *et al.* 2012) may result from sampling of an exotic population in Belgium (Filip Verloove, personal communication, 2011; Jia & Bartish, unpublished).

The final systematic problem in the species, indicated by Jia *et al.* (2012) and this study, regards ssp. *mongolica*. This taxon most likely has a hybrid origin resulting from a secondary contact between ssp. *sinensis* and *turkestanica*. The monophyletic clade of its cpDNA lineages is sister to the clade of ssp. *sinensis*, while nDNA sequences create a sub-clade within the clade of lineages from ssp. *turkestanica* (Figs. 3-2, 3-3). Unlike in some other hybrid taxa of *Hippophae* (Wang *et al.* 2008a), both cpDNA and nDNA lineages in ssp. *mongolica* are monophyletic, suggesting a unique hybridization event. Besides, ssp. *mongolica* has a distinct range non-overlapping with other taxa (Fig. 3-1) and specific climatic and biotic adaptations, imposed by this range. We believe this is a relatively rare case in plants of a unique homoploid hybridization, which likely was followed by adaptation to a new ecological niche and range expansion. It can represent an interesting evolutionary phenomenon and needs further investigation.

Biogeographic patterns in *Hippophae*

Early diversification in Hippophae *[stage I]*

The central Himalaya and southern Tibet were suggested as the ancestral area of the genus in our reconstructions (Fig. 3-5). Early diversification within the genus most likely started in this region in the Late Oligocene/Early Miocene. This placement in time and space corresponds to one of the stages of uplift of the QTP, the uplift-stage C by Zhang *et al.* (2013) at 25 to 20 Ma. Although it is tempting to link the earliest diversification in the genus with the rise of the southern QTP in the Early Miocene, we note that reconstruction of ages and paleoaltitudes of the QTP uplifts is still a controversial area of research (Harris 2006; Molnar *et al.* 2010). For example, close to current altitude of the central QTP has been suggested to exist from at least 40 Ma (DeCelles *et al.* 2007; Rowley & Currie 2006; Wang *et al.* 2008b), at least 15 Ma (Spicer *et al.* 2003), or to be no older than 2–3 Ma (Wang *et al.* 2008c). Until this controversy is resolved and robust models for early orogenesis of the QTP are developed, linking geologic and biogeographic events around this plateau in the Paleogene and Early Neogene will be possible only at a very general and approximate level. Our results can serve, however, as an indirect evidence of orogenesis and uplift in the central Himalaya/southern QTP in the Early Miocene.

Diversification and expansion of Hippophae rhamnoides *[stage II]*

After a period of rapid radiation in the central Himalaya/southern QTP in the Early to Middle Miocene and differentiation of all species, accelerated evolution in the genus shifted to the eastern QTP (Fig. 3-5). This was the ancestral area of *H. rhamnoides* in our

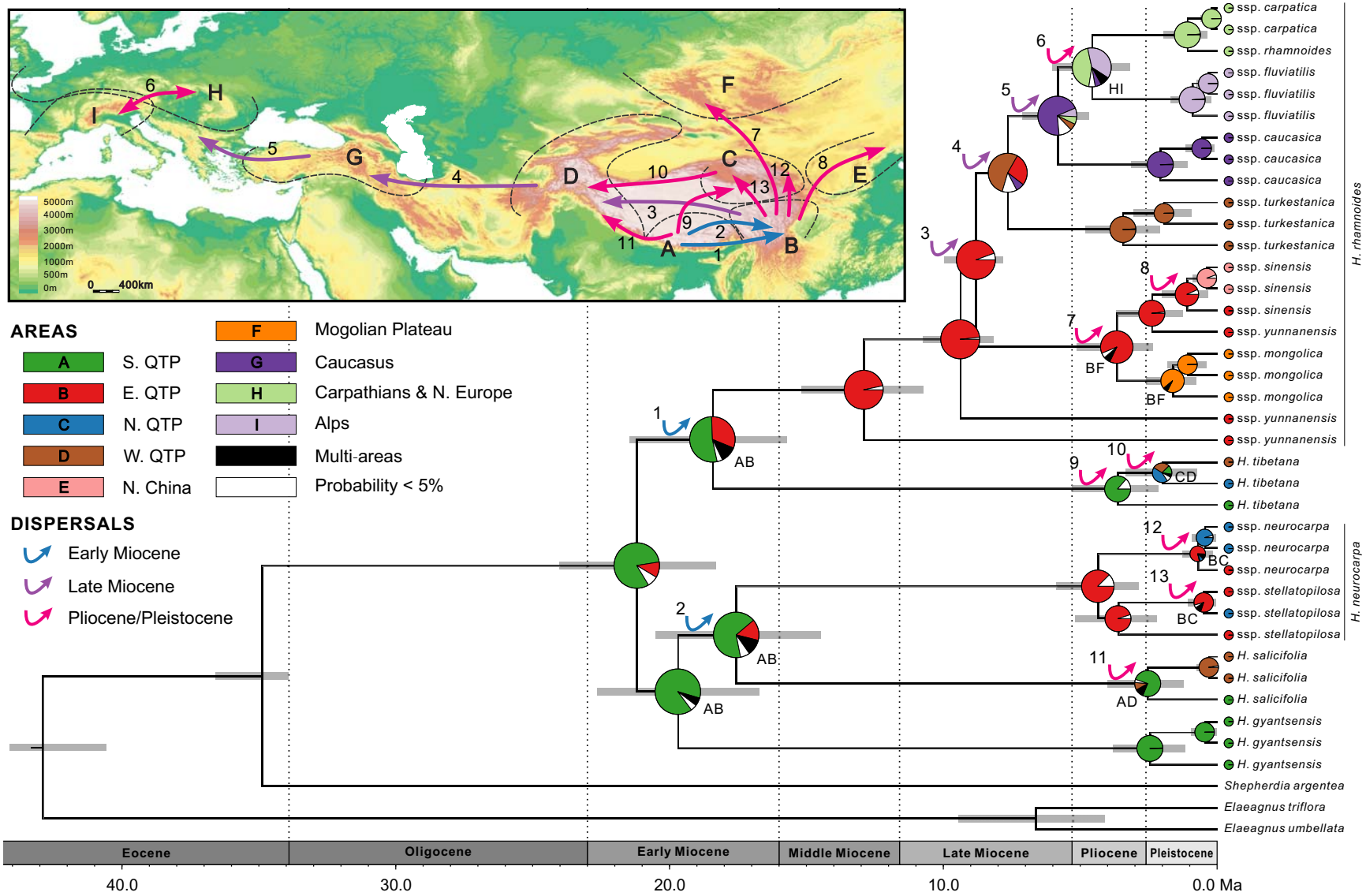


Fig. 3-5 Combined chronogram and biogeographic analysis of the genus *Hippophae* L. Grey bars represent 95% credible intervals for each node. Node charts show the relative probabilities of alternative ancestral distribution obtained by Bayes-DIVA analysis of cpDNA in RASP. The detected dispersal events are illustrated on the map.

reconstructions and several oldest lineages of this species are distributed there; some of which can represent cryptic taxa (see Phylogeny and Systematics of *Hippophae* in this Discussion). Differentiation of a common ancestor of ssp. *sinensis* and *turkestanica* (probably from one of the northern lineages within ssp. *yunnanensis*) was followed by a dispersal event across QTP in the earliest Late Miocene (the Tortonian age). This event can be associated with co-temporary uplift of the western QTP (Zhang *et al.* 2013; Zhang *et al.* 2012) and paleoclimatic changes in the area, indicating advance of more open and arid environments across QTP (Hui *et al.* 2011; Tang & Shen 1996; Zhang *et al.* 2012).

Considerable climatic and landscape modifications in area between the west QTP and Europe in the Late Miocene are indicated by geologic (Ramstein *et al.* 1997; Trifonov *et al.* 2012), paleozoologic (Eronen *et al.* 2009; Fortelius *et al.* 2006; Koufos *et al.* 2005), and paleobotanic (Kovar-Eder *et al.* 2008; Mosbrugger *et al.* 2005) data. These processes were coincident with ecosystem turnovers of this age in the area: degradation of tropical and subtropical laurophyllous forests, expansion of warm temperate sclerophyllous woodlands, scrubs and savannas (Biltekin 2010; Kovar-Eder *et al.* 2008; Mosbrugger *et al.* 2005; Syabryaj *et al.* 2007), and expansion of grasses (Strömberg *et al.* 2007). Reconstructions of evolution of bovid and hipparion megafauna (Barry *et al.* 1995; Eronen *et al.* 2009; Fortelius *et al.* 2006; Liu *et al.* 2012) and co-temporal turnovers in small mammalian fauna (Maridet *et al.* 2007; Mein 2003) across central parts of Eurasia provide a detailed evidence for these ecological and evolutionary processes in the Late Miocene. Our results indicate that *Hippophae rhamnoides* was a part of these processes. After dispersal to western QTP, this species expanded across Karakoram, Hindu Kush, Pamir, and Tian Shan, giving rise to extant ssp. *turkestanica*. Western ranges of Hindu Kush and young Kopet Dag, Elburz and Caucasus were situated relatively closely to each

other at that geological epoch (Trifonov *et al.* 2012). Their orientation created some kind of more or less continuous belt of mountain ranges and ecological bridge between highland habitats in Central Asia, Anatolia and Balkans. Western populations of ssp. *turkestanica* could use this ecological opportunity for westward expansion. Age estimate of the putative dispersal followed by vicariance corresponds to period of considerable orogenic, climatic and ecological modifications in Central Asia, eastern Anatolia and eastern Europe (Fortelius *et al.* 2006; Mosbrugger *et al.* 2005; Syabryaj *et al.* 2007; Trifonov *et al.* 2012). Thus, our results suggest that *H. rhamnoides* originated in the Middle Miocene in eastern QTP and reached Europe in three long-distance dispersals in the Late Miocene (Fig. 3-5). Ecological opportunity at that time through rise of a relatively continuous belt of mountain ranges west of QTP and opening of landscape through deforestation of these areas seems to be the main trigger of these dispersals.

Diversification within taxa [stage III]

Our dating analyses revealed an interesting pattern: regardless of the age of each taxon, differentiation within all taxa (except the paraphyletic ssp. *yunnanensis*) was largely co-temporal and relatively recent, in the Late Pliocene/Pleistocene (Fig. 3-5; Table S3-6). Results of our dating and RASP analyses suggest two dispersal/vicariance events within taxa in the Pliocene: from southern to north-eastern QTP in *H. tibetana*, and from eastern QTP to Mongolian Plateau in *H. rhamnoides* (ssp. *sinensis/yunnanensis* clade and ssp. *mongolica*). We note that both dispersals are in northern direction. Three more dispersals in northern direction followed by vicariance are indicated soon after these events in the Pleistocene, and all are from eastern QTP: one to northern China (ssp. *sinensis*), and two

to northern QTP (*H. neurocarpa* ssp. *neurocarpa* and *stellatopilosa*). Our results suggest also two almost coincident (the late Pliocene/early Pleistocene) westward dispersals along opposite slopes of the QTP: in *H. tibetana* (northern slopes) and in *H. salicifolia* (southern slopes). The pattern of coincident biogeographic events in the Pliocene and Pleistocene with similar direction of dispersals (mostly from south to north) within different regions of QTP, central, and western Eurasia is conspicuous. It indicates a relatively uniform impact of climatic fluctuations and orogenesis of these geologic epochs on evolutionary processes across Eurasian highlands.

References

Akaike H (1974) A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19**, 716-723.

An Z, Kutzbach JE, Prell WL, Porter SC (2001) Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature* **411**, 62-66.

Axelrod DI, Al SI, Raven PH (1996) History of the modern flora of China. In: *Floristic characteristics and diversity of East Asian plants* (eds. Zhang AL, Wu SG), pp. 43-55. Springer, Hongkong.

Barry JC, Morgan ME, Flynn LJ*, et al.* (1995) Patterns of faunal turnover and diversity in the Neogene Siwaliks of northern Pakistan. *Palaeogeography, Palaeoclimatology, Palaeoecology* **115**, 209-226.

Bartish IV, Jeppsson N, Bartish GI, Lu R, Nybom H (2000) Inter- and intraspecific genetic variation in *Hippophae* (Elaeagnaceae) investigated by RAPD markers. *Plant Systematics and Evolution* **225**, 85-101.

Bartish IV, Jeppsson N, Nybom H, Swenson U (2002) Phylogeny of *Hippophae* (Elaeagnaceae) inferred from parsimony analysis of chloroplast DNA and morphology. *Systematic botany* **27**, 41-54.

Bartish IV, Kadereit JW, Comes HP (2006) Late Quaternary history of *Hippophae rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (*Chsi*) sequences and chloroplast DNA

variation. *Molecular Ecology* **15**, 4065-4083.

Becker HF (1960) The Tertiary Mormon Creek flora from the upper Ruby River basin in southwestern Montana. *Palaeontographica Abteilung B* **107**, 83-126.

Berrgren W, Van Couvering J (1974) The Late Neogene: biostratigraphy, geochronology, and paleoclimatology of the last 15 million years in marine and continental sequences. *Palaeogeography, Palaeoclimatology, Palaeoecology* **16**.

Biltekin D (2010) *Vegetation and climate of North Anatolian and North Aegean region since 7 Ma according to pollen analysis* PhD Thesis, Université Claude Bernard-Lyon I and Istanbul Technical University.

Bobrov EG (1962) Obzor roda *Myricaria* Desv. i jego istorija. *Botanicheskij Zhurnal* **52**, 924-936.

Bond G, MacConnell J, McCallum A (1956) The nitrogen-nutrition of *Hippophaë rhamnoides* L. *Annals of Botany* **20**, 501-512.

Buckley TR, Cordeiro M, Marshall DC, Simon C (2006) Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada* Dugdale). *Systematic biology* **55**, 411-425.

Burge DO, Manchester SR (2008) Fruit morphology, fossil history, and biogeography of Paliurus (Rhamnaceae). *International Journal of Plant Sciences* **169**, 1066-1085.

Calvillo-Canadell L, Cevallos-Ferriz SR (2007) Reproductive structures of Rhamnaceae from the Cerro del Pueblo (Late Cretaceous, Coahuila) and Coatzingo (Oligocene, Puebla) Formations, Mexico. *American Journal of Botany* **94**, 1658-1669.

Chandler MEJ (1962) *The Lower Tertiary Floras of Southern England. II. Flora of the Pipe-Clay Series of Dorset (Lower Bagshot)* British Museum (Natural History), London.

Cheng K, Sun K, Yan WH*, et al.* (2009) Maternal divergence and phylogeographical relationships between *Hippophae gyantsensis* and *H. rhamnoides* subsp. *yunnanensis*. *Chinese Journal of Plant Ecology* **33**, 1-11.

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

Comes HP, Kadereit JW (2003) Spatial and temporal patterns in the evolution of the flora of the European Alpine System. *Taxon* **52**, 451-462.

Correa E, Jaramillo C, Manchester S, Gutierrez M (2010) A fruit and leaves of Rhamnaceous affinities from the late Cretaceous (Maastrichtian) of Colombia. *American Journal of Botany* **97**, 71-79.

Crepet WL, Nixon KC, Gandolfo MA (2004) Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *American Journal of Botany* **91**, 1666-1682.

DeCelles PG, Kapp P, Ding L, Gehrels G (2007) Late Cretaceous to middle Tertiary basin evolution in the central Tibetan Plateau: Changing environments in response to tectonic partitioning,

aridification, and regional elevation gain. *Geological Society of America Bulletin* **119**, 654-680.

Demesure B, Sodzi N, Petit R (1995) A set of universal primers for amplification of polymorphic non‐coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**, 129-134.

Doyle JJ (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany*, 144-163.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**, 1307-1320.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology* **7**, 214.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution* **29**, 1969-1973.

Eronen JT, Ataabadi MM, Micheels A*, et al.* (2009) Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proceedings of the National Academy of Sciences* **106**, 11867-11871.

Fortelius M, Eronen J, Liu L*, et al.* (2006) Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. *Palaeogeography, Palaeoclimatology, Palaeoecology* **238**, 219-227.

Gao YD, Harris AJ, Zhou SD, He XJ (2013) Evolutionary events in *Lilium* (including *Nomocharis*, Liliaceae) are temporally correlated with orogenies of the Q-T plateau and the Hengduan Mountains. *Molecular Phylogenetics and Evolution* **68**, 443-460.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.

Guo Z, Ruddiman WF, Hao Q*, et al.* (2002) Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature* **416**, 159-163.

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**, 521-523.

Harris N (2006) The elevation history of the Tibetan Plateau and its implications for the Asian monsoon. *Palaeogeography, Palaeoclimatology, Palaeoecology* **241**, 4-15.

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

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

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.

Hui Z, Li J, Xu Q*, et al.* (2011) Miocene vegetation and climatic changes reconstructed from a sporopollen record of the Tianshui Basin, NE Tibetan Plateau. *Palaeogeography, Palaeoclimatology, Palaeoecology* **308**, 373-382.

Hyvönen J (1996) On phylogeny of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **16**, 51-62.

Islam MB, Simmons MP (2006) A thorny dilemma: testing alternative intrageneric classifications within *Ziziphus* (Rhamnaceae). *Systematic botany* **31**, 826-842.

Jia DR, Abbott RJ, Liu TL*, et al.* (2012) Out of the Qinghai–Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophaë rhamnoides* (Elaeagnaceae). *New Phytologist* **194**, 1123-1133.

Jia DR, Liu TL, Wang LY, Zhou DW, Liu JQ (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* **102**, 37-50.

Kadereit JW, Licht W, Uhink CH (2008) Asian relationships of the flora of the European Alps. *Plant Ecology & Diversity* **1**, 171-179.

Koufos GD, Kostopoulos DS, Vlachou TD (2005) Neogene/Quaternary mammalian migrations in eastern Mediterranean. *Belgian Journal of Zoology* **135**, 181.

Kovar-Eder J, Jechorek H, KvačEk Z, Parashiv V (2008) The integrated plant record: An essential tool for reconstructing Neogene zonal vegetation in Europe. *Palaios* **23**, 97-111.

Lang G (1994) *Quartäre Vegetationsgeschichte Europas* Gustav Fischer Verlag, Jena, Germany.

Lewis CE, Doyle JJ (2001) Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). *Molecular Phylogenetics and Evolution* **19**, 409-420.

Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**, 913.

Li JJ, Fang XM, Pan BT, Zhao ZJ, Song YG (2001) Late Cenozoic intensive uplift of Qinghai-Xizang Plateau and its impacts on environments in surrounding area. *Quaternary Sciences* **21**, 381-391.

Li M, Wunder J, Bissoli G*, et al.* (2008) Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* **24**, 727-745.

Lian Y (1988) New discoveries of the genus *Hippophae* L.(Elaeagnaceae). *Acta Phytotaxonomica Sinica* **26**, 235-237.

Lian YS, Chen XL, Lian H (1998) Systematic classification of the genus *Hippophae* L. *Seabuckthorn Research* **1**, 13-23.

Lian YS, Chen XL, Sun K, Ma RJ (2003) A new subspecies of *Hippophae* (Elaeagnaceae) from China. *Novon* **13**, 200-202.

Lielke K, Manchester S, Meyer H (2012) Reconstructing the environment of the northern Rocky

Mountains during the Eocene/Oligocene transition: constraints from the palaeobotany and geology of south-western Montana, USA. *Acta Palaeobotanica* **52**, 317-358.

Liu J-Q, Wang Y-J, Wang A-L, Hideaki O, Abbott RJ (2006) Radiation and diversification within the *Ligularia–Cremanthodium–Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. *Molecular Phylogenetics and Evolution* **38**, 31-49.

Liu L, Pearl DK (2007) Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic biology* **56**, 504-514.

Liu L, Puolamäki K, Eronen JT*, et al.* (2012) Dental functional traits of mammals resolve productivity in terrestrial ecosystems past and present. *Proceedings of the Royal Society B: Biological Sciences* **279**, 2793-2799.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011a) Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**, 1267-1280.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011b) Mountains of Southern China as "Plant Museums" and "Plant Cradles": Evolutionary and Conservation Insights. *Mountain Research and Development* **31**, 261-269.

Maddison WP (1997) Gene trees in species trees. *Systematic biology* **46**, 523-536.

Manchester SR, Akhmetiev MA, Kodrul TM (2002) Leaves and fruits of Celtis aspera (Newberry) comb. nov.(Celtidaceae) from the Paleocene of North America and eastern Asia. *International Journal of Plant Sciences* **163**, 725-736.

Maridet O, Escarguel G, Costeur L*, et al.* (2007) Small mammal (rodents and lagomorphs) European biogeography from the Late Oligocene to the mid Pliocene. *Global Ecology and Biogeography* **16**, 529-544.

Mein P (2003) On Neogene rodents of Eurasia: distribution and migrations. *Deinsea (Rotterdam)* **10**, 407-418.

Melinte-Dobrinescu MC, Suc J-P, Clauzon G*, et al.* (2009) The Messinian Salinity Crisis in the Dardanelles region: chronostratigraphic constraints. *Palaeogeography, Palaeoclimatology, Palaeoecology* **278**, 24-39.

Meng LH, Yang HL, Wu GL, Wang YJ (2008) Phylogeography of *Hippophae neurocarpa* (Elaeagnaceae) inferred from the chloroplast DNA *trn*L-F sequence variation. *Journal of Systematics and Evolution* **46**, 32-40.

Mittermeier RA, Gil PR, Hoffman M*, et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions* Cemex, Conservation International and Agrupación Sierra Madre, Monterrey, Mexico.

Molnar P (2005) Mio-Pliocene growth of the Tibetan Plateau and evolution of East Asian climate. *Palaeontologia Electronica* **8**, 1-23.

Molnar P, Boos WR, Battisti DS (2010) Orographic controls on climate and paleoclimate of Asia: thermal and mechanical roles for the Tibetan Plateau. *Annual Review of Earth and Planetary Sciences* **38**, 77.

Mosbrugger V, Utescher T, Dilcher DL (2005) Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14964-14969.

Muller J (1968) Palynology of the Pedawan and plateau sandstone formations (Cretaceous-Eocene) in Sarawak, Malaysia. *Micropaleontology* **14**, 1-37.

Muller J (1981) Fossil pollen records of extant angiosperms. *The Botanical Review* **47**, 1-142.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**, 853-858.

Pelser PB, Kennedy AH, Tepe EJ*, et al.* (2010) Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. *American Journal of Botany* **97**, 856-873.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Rambaut A, Drummond AJ (2007) Tracer v 1.5. Available from http://beast.bio.ed.ac.uk/Tracer.

Ramstein G, Fluteau F, Besse J, Joussaume S (1997) Effect of orogeny, plate motion and land-sea distribution on Eurasian climate change over the past 30 million years. *Nature* **386**, 788-795.

Richardson JE, Chatrou LW, Mols JB, Erkens RHJ, Pirie MD (2004) Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society B-Biological Sciences* **359**, 1495-1508.

Richardson JE, Fay MF, Cronk QCB, Chase MW (2000) A revision of the tribal classification of Rhamnaceae. *Kew Bulletin* **55**, 311-340.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Roquet C, Boucher FC, Thuiller W, Lavergne S (2013) Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *Journal of Biogeography* **40**, 1874-1886.

Rousi A (1965) Observation on the cytology and variation of European and Asiatic populations of *Hippophaë ramnoides*. *Annales Botanici Fennici* **2**, 1-18.

Rousi A (1971) The genus *Hippophae* L. A taxonomic study. *Annales Botanici Fennici* **8**, 177-227.

Rowley DB, Currie BS (2006) Palaeo-altimetry of the late Eocene to Miocene Lunpola basin, central Tibet. *Nature* **439**, 677-681.

Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**, 101-109.

Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the

absence of a molecular clock. *Bioinformatics* **19**, 301-302.

Sauquet H, Ho SY, Gandolfo MA*, et al.* (2012) Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). *Systematic Biology* **61**, 289-313.

Shaw J, Lickey EB, Beck JT*, et al.* (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**, 142-166.

Song ZC, Qian ZS (1989) Palynological study of Taizhou Formation in North Jiangsu Basin. In: *Stratigraphy and Palaeontology of the Taizhou Formation and the First member of the Funing Formation, North Jiangsu Basin* (eds. Institute of Geology-Jiangsu Oil Exploration and Development Corporation, Nanjing Institute of Geology and Palaeontology-Academia Sinica), pp. 33-109. Nanjing University Press, Nanjing, China.

Song ZC, Wang WM, Huang F (2004) Fossil pollen records of extant angiosperms in China. *Botanical Review* **70**, 425-458.

Sorsa P (1971) Pollen morphological study of the genus *Hippophae* L., including the new taxa recognised by A. Rousi. *Annales Botanici Fennici* **8**, 228-236.

Spicer RA, Harris NB, Widdowson M*, et al.* (2003) Constant elevation of southern Tibet over the past 15 million years. *Nature* **421**, 622-624.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Systematic biology* **57**, 758-771.

Strand A, Leebens-Mack J, Milligan B (1997) Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* **6**, 113-118.

Strömberg CA, Werdelin L, Friis EM, Saraç G (2007) The spread of grass-dominated habitats in Turkey and surrounding areas during the Cenozoic: phytolith evidence. *Palaeogeography, Palaeoclimatology, Palaeoecology* **250**, 18-49.

Sun K, Chen X, Ma R*, et al.* (2002) Molecular phylogenetics of *Hippophae* L. (Elaeagnaceae) based on the internal transcribed spacer (ITS) sequences of nrDNA. *Plant Systematics and Evolution* **235**, 121-134.

Sun YS, Wang AL, Wan DS, Wang Q, Liu JQ (2012) Rapid radiation of *Rheum* (Polygonaceae) and parallel evolution of morphological traits. *Molecular Phylogenetics and Evolution* **63**, 150-158.

Swenson U, Bartish IV (2002) Taxonomic synopsis of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **22**, 369-374.

Syabryaj S, Utescher T, Molchanoff S, Bruch A (2007) Vegetation and palaeoclimate in the Miocene

of Ukraine. *Palaeogeography, Palaeoclimatology, Palaeoecology* **253**, 153-168.

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

Takhtajan A (1986) *Floristic regions of the world.* University of California press, Berkeley.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

Tang L, Shen C (1996) Late Cenozoic vegetational history and climatic characteristics of Qinghai-Xizang Plateau. *Acta Micropalaeontologica Sinica* **13**, 321-337.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876-4882.

Tian C, He X, Zhong Y, Chen J (2002) Effects of VA mycorrhizae and *Frankia* dual inoculation on growth and nitrogen fixation of *Hippophae tibetana*. *Forest ecology and management* **170**, 307-312.

Trifonov VG, Ivanova TP, Bachmanov DM (2012) Recent mountain building of the central Alpine-Himalayan Belt. *Geotectonics* **46**, 315-332.

Walker JD, Geissman JW, Bowring SA, Babcock LE (2012) Geologic Time Scale v. 4.0: Geological Society of America, doi: 10.1130/2012.CTS004R3C.

Wang AL, Schluetz F, Liu JQ (2008a) Molecular evidence for double maternal origins of the diploid hybrid *Hippophae goniocarpa* (Elaeagnaceae). *Botanical Journal of the Linnean Society* **156**, 111-118.

Wang C, Zhao X, Liu Z*, et al.* (2008b) Constraints on the early uplift history of the Tibetan Plateau. *Proceedings of the National Academy of Sciences* **105**, 4987-4992.

Wang H, Moore MJ, Soltis PS*, et al.* (2009a) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences* **106**, 3853-3858.

Wang H, Qiong L, Sun K*, et al.* (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**, 2964-2979.

Wang Y-J, Susanna A, Von Raab-Straube E, Milne R, Liu J-Q (2009b) Island-like radiation of *Saussurea* (Asteraceae: Cardueae) triggered by uplifts of the Qinghai–Tibetan Plateau. *Biological Journal of the Linnean Society* **97**, 893-903.

Wang Y, Wang X, Xu Y*, et al.* (2008c) Stable isotopes in fossil mammals, fish and shells from Kunlun Pass Basin, Tibetan Plateau: paleo-climatic and paleo-elevation implications. *Earth and Planetary Science Letters* **270**, 73-85.

Wheeler EA, Meyer HW (2012) A new (Hovenia) and old (Chadronoxylon) fossil wood from the Late Eocene Florissant Formation, Colorado, USA. *IAWA Journal-International Association of*

Wood Anatomists **33**, 309.

White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds. Innis MA, Gelfand DH, Shinsky JJ, White TJ), pp. 315-322. Academic Press, New York.

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**, 2211-2220.

Wu ZY (1988) Hengduan Mountains flora and its significance. *Journal of Japanese Botany* **63**, 297-311.

Wu ZY, Wu SG (1998) A proposal for a new floristic kingdom (realm): the E. Asiatic Kingdom, its delineation and characteristics. In: *Floristic characteristics and diversity of East Asian plants* (eds. Wu ZY, Wu SG), pp. 3-42. China Higher Education Press, Beijing.

Wulff EV (1943) *An Introduction to Historical Plant Geography* Chronica Botanica Company, Waltham.

Young ND, Healy J (2003) GapCoder automates the use of indel characters in phylogenetic analysis. *Bmc Bioinformatics* **4**, 6.

Yu Y, Harris AJ, He X (2010) S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution* **56**, 848-850.

Yu Y, Harris AJ, He X (2011) RASP (Reconstruct Ancestral State in Phylogenies) 2.0 beta. Available at http://mnh.scu.edu.cn/soft/blog/RASP.

Zerega NJ, Clement WL, Datwyler SL, Weiblen GD (2005) Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution* **37**, 402-416.

Zhang K, Wang G, Xu Y*, et al.* (2013) Sedimentary Evolution of the Qinghai–Tibet Plateau in Cenozoic and its Response to the Uplift of the Plateau. *Acta Geologica Sinica-English Edition* **87**, 555-575.

Zhang Q-Q, Ferguson DK, Mosbrugger V, Wang Y-F, Li C-S (2012) Vegetation and climatic changes of SW China in response to the uplift of Tibetan Plateau. *Palaeogeography, Palaeoclimatology, Palaeoecology* **363**, 23-36.

Zhang S-D, Soltis DE, Yang Y, Li D-Z, Yi T-S (2011) Multi-gene analysis provides a well-supported phylogeny of Rosales. *Molecular Phylogenetics and Evolution* **60**, 21-28.

Zhang XL, Wang YJ, Ge XJ*, et al.* (2009) Molecular phylogeny and biogeography of *Gentiana* sect. *Cruciata* (Gentianaceae) based on four chloroplast DNA datasets. *Taxon* **58**, 862-870.

Zheng H, Powell CM, An Z, Zhou J, Dong G (2000) Pliocene uplift of the northern Tibetan Plateau. *Geology* **28**, 715-718.

Supporting information

Table S3-1 Taxa and GenBank accession numbers for sequences used in reconstructing the phylogeny of Rosales

Family	Genus	18S	26S	*atp*B	*mat*K	*ndh*F	*psb*BTNH	*rbc*L	*rpo*C2	*rps*4	*trn*L-F
Rosales											
Barbeyaceae											
	Barbeya	JF317358	JF317379	JF317398	JF317418	JF317457	JF317477	JF317437	JF317497	JF317517	AJ225795
Cannabaceae											
	Cannabis	JF317360	EU002151	JF317400	AF345317	AY289250	EU002390	AJ402933	EU002481	EU002301	AF501598
	Celtis	JF317361	JF317380	JF317401	JF317420	JF317459	JF317479	JF317439	JF317499	JF317519	AY488673
	Trema	JF317376	JF317395	JF317415	JF317434	JF317474	JF317494	JF317454	JF317514	JF317532	JN040388
Dirachmaceae											
	Dirachma	JF317364	JF317383	–	JF317423	JF317462	JF317482	JF317442	JF317502	JF317522	AJ225796
Elaeagnaceae											
	Elaeagnus[1]	JF317366	JF317385	JF317405	JF317425	JF317464	JF317484	JF317444	JF317504	JF317524	–
	Elaeagnus[2]	–	–	–	–	–	–	U17038	–	–	DQ838727
	Elaeagnus[3]	L24090	–	–	AY257529	–	–	HM849968	–	–	JN102138
	Hippophae[4]	JF317370	JF317389	JF317409	JF317428	JF317468	JF317488	JF317448	JF317508	JF317526	JQ289181
	Hippophae[5]	–	–	–	JF954063	–	–	JF941954	–	–	HM769675
	Hippophae[6]	–	–	–	JF954058	–	–	U59821	–	–	HM769681
	Shepherdia[7]	–	–	–	–	–	–	AJ225787	–	–	AJ225794
	Shepherdia[8]	AF207020	AF479100	AF209675	JN966666	–	–	U17039	–	–	GQ245525
Moraceae											
	Broussonetia	JF317359	AY686805	JF317399	JF317419	JF317458	JF317478	JF317438	JF317498	JF317518	AF501602
	Cudrania	JF317362	JF317381	JF317402	JF317421	JF317460	JF317480	JF317440	JF317500	JF317520	JN102171
	Ficus	JF317367	JF317386	JF317406	JF317426	JF317465	JF317485	JF317445	JF317505	JF317525	JN040373
	Morus	L24398	AF479232	NC008359	NC008359	NC008359	NC008359	NC008359	NC008359	NC008359	AF501608

Table S3-1 (*continued*)

Family	Genus	18S	26S	*atp*B	*mat*K	*ndh*F	*psb*BTNH	*rbc*L	*rpo*C2	*rps*4	*trn*L-F
Rhamnaceae											
	Ceanothus	U42799	AF479102	AF209558	AF049815	EU002210	EU002393	U06795	EU002484	EU002304	HQ325601
	Gouania	JF317369	JF317388	JF317408	JF317427	JF317467	JF317487	JF317447	JF317507	–	AJ390344
	Hovenia	JF317371	JF317390	JF317410	JF317429	JF317469	JF317489	JF317449	JF317509	JF317527	AJ390343
	Rhamnus	JF317374	JF317393	JF317413	JF317432	JF317472	JF317492	JF317452	JF317512	JF317530	AF348565
Rosaceae											
	Dryas	JF317365	JF317384	JF317404	JF317424	JF317463	JF317483	JF317443	JF317503	JF317523	DQ851231
	Filipendula	JF317368	JF317387	JF317407	AB073684	JF317466	JF317486	JF317446	JF317506	–	AJ416463
	Neillia	JF317372	JF317391	JF317411	JF317430	JF317470	JF317490	JF317450	JF317510	JF317528	AF487228
	Sanguisorba	JF317375	JF317394	JF317414	JF317433	JF317473	JF317493	JF317453	JF317513	JF317531	AY634775
	Spiraea	U42801	AF479103	AJ235608	AF288127	EU002259	EU002451	L11206	EU002541	EU002362	AF348571
Ulmaceae											
	Ulmus	JF317377	JF317396	JF317416	JF317435	JF317475	JF317495	JF317455	JF317515	JF317533	AF501593
	Zelkova	U42819	AF479099	AF209699	AF345328	EU002273	EU002465	AF206835	EU002556	–	JX399109
Urticaceae											
	Debregeasia	JF317363	JF317382	JF317403	JF317422	JF317461	JF317481	JF317441	JF317501	JF317521	JN040372
	Pilea	JF317373	JF317392	JF317412	JF317431	JF317471	JF317491	JF317451	JF317511	JF317529	DQ179359
	Boehmeria	JF317378	JF317397	JF317417	JF317436	JF317476	JF317496	JF317456	JF317516	JF317534	AY208722
Cucurbitales(Outgroup)											
Begoniaceae											
	Begonia	AF008950	AY968403	AY968426	AY968445	EU002200	EU002384	L01888	EU002475	EU002294	GU397135

Note: [1]*Elaeagnus bockii*; [2]*Elaeagnus angustifolia*; [3]*Elaeagnus umbellata*; [4]*Hippophae rhamnoides* subsp. *sinensis*; [5]*Hippophae tibetana*; [6]*Hippophae salicifolia*; [7]*Shepherdia argentea* and [8]*Shepherdia canadensis*.

Table S3-2 Taxa, population codes and GenBank accession numbers for sequences used in analysis of *Hippophae* L.

Taxon	Population	*trn*C-ycf6	*trn*D-T	*trn*L-F	*trn*S-fM	*trn*S-G	*At*103	*G3pdh*	ITS	*Ms*	*Tpi*
H. salicifolia	4079	KF620857	KF620916	KF620652	KF620729	KF620763	KF620838	KF620955	KF620619	KF620691	KF620799
	09XZ040	KF620858	KF620917	KF620653	KF620730	KF620764	KF620839	KF620956	KF620620	KF620692	KF620800
	P	KF620856	KF620918	KF620654	KF620731	KF620765	KF620840	KF620957	KF620621	–	KF620801
H. gyantsensis	06252	KF620847	KF620886	KF620627	KF620699	KF620738	KF620808	KF620925	KF620594	KF620661	KF620769
	06217	KF620848	KF620887	KF620628	KF620700	KF620739	KF620809	KF620926	KF620595	KF620662	KF620770
	2568	KF620849	KF620888	KF620629	KF620701	KF620740	KF620810	KF620927	KF620596	KF620663	KF620771
H. neurocarpa											
ssp. *neurocarpa*	YNG1	KF620850	KF620889	KF620630	KF620702	KF620741	KF620811	KF620928	KF620597	KF620664	KF620772
	1522	KF620851	KF620890	KF620631	KF620703	KF620742	KF620812	KF620929	KF620598	KF620665	KF620773
	MM-31	KF620852	KF620891	KF620632	KF620704	KF620743	KF620813	KF620930	KF620599	KF620666	KF620774
ssp. *stellatopilosa*	Ao129	KF620853	KF620892	KF620633	KF620705	KF620744	KF620814	KF620931	KF620600	KF620667	KF620775
	QML1	KF620854	KF620893	KF620634	KF620706	KF620745	KF620815	KF620932	KF620601	KF620668	KF620776
	Ao111	KF620855	KF620894	KF620635	KF620707	KF620746	KF620816	KF620933	KF620602	KF620669	KF620777
H. tibetana	07138	KF620881	KF620919	KF620655	KF620732	KF620766	KF620841	KF620958	KF620622	KF620693	KF620802
	Henan	KF620882	KF620920	KF620656	KF620734	KF620767	KF620842	KF620959	KF620623	KF620694	KF620803
	ST	KF620880	KF620921	KF620657	–	KF620768	KF620843	KF620960	KF620624	KF620695	KF620804
H. rhamnoides											
ssp. *yunnanensis*	06309	KF620877	KF620913	KF620649	KF620726	KF620755	KF620835	KF620952	KF620616	KF620688	KF620796
	06321	KF620878	KF620914	KF620650	KF620727	KF620761	KF620836	KF620953	KF620617	KF620689	KF620797
	06324	KF620879	KF620915	KF620651	KF620728	KF620762	KF620837	KF620954	KF620618	KF620690	KF620798
ssp. *sinensis*	Ao96	KF620871	KF620907	KF620643	KF620720	KF620758	KF620829	KF620946	KF620610	KF620682	KF620790
	05276	KF620872	KF620908	KF620644	KF620721	KF620759	KF620830	KF620947	KF620611	KF620683	KF620791
	05179	KF620873	KF620909	KF620645	KF620722	KF620760	KF620831	KF620948	KF620612	KF620684	KF620792

Table S3-2 (*continued*)

Taxon	Pop. code	*trn*C-ycf6	*trn*D-T	*trn*L-F	*trn*S-fM	*trn*S-G	*At*103	*G3pdh*	ITS	*Ms*	*Tpi*
H. rhamnoides											
ssp. *turkestanica*	GL	KF620875	KF620910	KF620646	KF620723	KF620751	KF620832	KF620949	KF620613	KF620685	KF620793
	SKZ	KF620874	KF620911	KF620647	KF620724	KF620752	KF620833	KF620950	KF620614	KF620686	KF620794
	12413	KF620876	KF620912	KF620648	KF620725	KF620753	KF620834	KF620951	KF620615	KF620687	KF620795
ssp. *mongolica*	05047	KF620869	KF620903	JQ289182	KF620716	JQ289198	KF620825	KF620942	JQ289227	KF620678	KF620786
	6667	KF620867	KF620904	KF620640	KF620717	KF620757	KF620826	KF620943	KF620607	KF620679	KF620787
	6657	KF620868	KF620905	KF620641	KF620718	KF620756	KF620827	KF620944	KF620608	KF620680	KF620788
ssp. *caucasica*	0402	KF620861	KF620897	JQ663586	KF620710	JQ663592	KF620819	KF620936	JQ663578	KF620672	KF620780
	21286	KF620862	KF620898	KF620636	KF620711	KF620747	KF620820	KF620937	KF620603	KF620673	KF620781
	9835	KF620863	KF620899	KF620637	KF620712	KF620748	KF620821	KF620938	KF620604	KF620674	KF620782
ssp. *fluviatilis*	SW1	KF620866	KF620900	JQ289193	KF620713	JQ289217	KF620822	KF620939	JQ289287	KF620675	KF620783
	6757	KF620864	KF620901	KF620638	KF620714	KF620749	KF620823	KF620940	KF620605	KF620676	KF620784
	6682	KF620865	KF620902	KF620639	KF620715	KF620750	KF620824	KF620941	KF620606	KF620677	KF620785
ssp. *carpatica*	9897	KF620859	KF620895	JQ663584	KF620708	JQ663591	KF620817	KF620934	JQ663577	KF620670	KF620778
	9898	KF620860	KF620896	JQ663585	KF620709	JQ663595	KF620818	KF620935	JQ663576	KF620671	KF620779
ssp. *rhamnoides*	6672	KF620870	KF620906	KF620642	KF620719	KF620754	KF620828	KF620945	KF620609	KF620681	KF620789
Elaeagnus triflora	IVB-29	KF620844	KF620883	KF620625	KF620696	KF620735	KF620805	KF620922	KF620592	KF620658	–
Elaeagnus umbellata	07078	KF620845	KF620884	HM769678	KF620697	KF620736	KF620806	KF620923	HM769715	KF620659	–
Shepherdia argentea	6777	KF620846	KF620885	KF620626	KF620698	KF620737	KF620807	KF620924	KF620593	KF620660	–
Rhamnus davurica		–	–	AY626420	–	–	–	–	AY626441	–	–

Table S3-3 Fossils used for calibration of phylogenetic trees of Rosales and *Hippophae* in our dating analyses (see Materials and Methods for details). Absolute ages in Ma used for calibration were determined either as minimum of the geological age of stratum with the fossil from the Geological Time Scale (Walker *et al.* 2012). [a)] Refer to "intuitive" and [b)] refer to "apomorphy-based" assignments of fossils to extant taxa, as specified by Sauquet *et al.* (2012). References for "apomorphy-based" assignments provide suits of characters used to assess fossil relationships to extant taxa. [c)] The fossil record was used to calibrate the stem node of a monophyletic group of four western subspecies within *H. rhamnoides* based on vicariance event, inferred from phylogenetic and ancestral area reconstruction analyses by Jia *et al.* (2012)

Fossil	Taxon name	Family	Fossil organ	Reference for original description	Locality	Region	Geological age/epoch	Absolute age in Ma	Reference for geochronology
1	*Triorites minutipori* Muller	Ulmaceae[a)]	pollen	Muller 1968	Sarawak	South-eastern Asia	Turonian	89.8	Muller 1968
2	*Celtis aspera* Manchester, Akhmetiev & Kodrul	Cannabaceae[b)]	fruits, leaves	Manchester *et al.* 2002	Numerous localities	North America; Asia	Paleocene	56	Manchester *et al.* 2002
3	*Ficus* L.	Moraceae[a)]	fruit	Chandler 1962	Lower Bagshot	Dorset, England	Ypresian	47.8	Chandler 1962
4	*Paliurus clarnensis* Burge & Manchester	Rhamnaceae[b)]	fruit	Burge & Manchester 2008	Red Gap, Jefferson	Oregon, USA	Lutetian	41.2	Burge & Manchester 2008
5	*Shepherdia weaveri* Becker	Elaeagnaceae[a)]	leaves	Becker 1960	Mormon Creek	Montana, USA	Priabonian	33.9	Lielke *et al.* 2012
6	*Hippophae rhamnoides* L.	Elaeagnaceae[c)]	pollen	Biltekin 2010	Burhanli	Marmor sea, Turkey	Tortonian	7.2	Melinte-Dobrinescu *et al.* 2009

Table S3-4 Gene/partition characteristics of the Rosales dataset

Partition	No. of taxa	Aligned length	Variable sites	Parsimony-informative sites	Model selected by AIC
18S rDNA	28	1733	202	101	GTR + I + G
26S rDNA	27	3256	494	277	GTR + I + G
*atp*B	26	1242	286	169	TIM1 + I + G
*mat*K	30	1029	535	333	TVM + I + G
*ndh*F	26	2116	924	592	GTR + I + G
*psb*BTNH	26	2278	598	354	TVM + I + G
*rbc*L	32	1289	315	195	GTR + I + G
*rpo*C2	26	3876	1547	908	GTR + I + G
*rps*4	23	546	164	79	GTR + G
*trn*L-F	31	970	502	311	GTR + I + G
Total	32	18335	5567	3319	

Table S3-5 Divergence time estimates of Rosales. Ranges on PL divergence time represent ± 1 standard deviation of the mean based on 1000 bootstrap replicates; while ranges on Bayesian estimates represent 95% confidence limits around mean age estimates. All ages are millions of years

Node	Age					
	PL		UCLN-lognormal		UCLN-uniform	
$Rosales_{crown}$	103	(102.8– 103.1)	102.5	(101.7–103.0)	102.6	(101.7–103.0)
$Rosaceae_{crown}$	75.9	(72.7–77.6)	69.3	(61.4–77.2)	69.8	(61.5–77.8)
Subclades split	96.2	(95.6–96.9)	96.0	(94.2–97.7)	96.1	(94.3–97.9)
$Ulmaceae_{stem}$	89.8	(89.6–90.3)	90.5	(89.8–91.4)	90.6	(89.8–92.0)
$Ulmaceae_{crown}$	7.3	(5.7–9.2)	6.9	(4.5–9.5)	6.9	(4.7–9.5)
$Canabaceae_{stem}$	76.3	(75.4–77.3)	76.4	(73.3–79.5)	76.4	(72.6–79.7)
$Canabaceae_{crown}$	58.5	(57.0–60.4)	57.1	(56.0–59.1)	58.5	(56.0–62.4)
$Moraceae_{stem}$	73.7	(72.8–74.7)	73.5	(70.1–76.9)	73.6	(70.0–77.2)
$Urticaceae_{crown}$	56.6	(55.1–58.2)	55.4	(49.9–60.8)	55.6	(49.8–60.9)
$Moraceae_{crown}$	55.7	(54.8–56.6)	56.1	(53.5–59.0)	56.3	(53.4–59.1)
$Rhamnacea_{stem}$	93.2	(92.5–94.2)	91.1	(87.5–94.0)	91.3	(88.3–94.2)
$Rhamnacea_{crown}$	77.5	(75.7–79.0)	76.2	(70.1–82.4)	76.5	(70.5–82.3)
$Elaeagnaceae_{stem}$	92.3	(91.1–93.1)	89.3	(85.7–93.0)	89.5	(86.3–92.9)
$Barbeyaceae_{stem}$	86.6	(84.1–88.1)	82.8	(77.2–87.9)	83.2	(77.8–88.1)
$Elaeagnaceae_{crown}$	42.2	(40.9–43.8)	40.6	(36.9–44.0)	40.7	(37.6–44.1)
$\textit{Elaeagnus}_{crown}$	20.5	(18.1–24.0)	15.7	(8.9–23.5)	15.5	(8.8–23.0)
$\textit{Shepherdia}_{stem}$	33.9	(33.3–34.6)	34.4	(33.9–35.1)	34.4	(33.9–35.4)
$\textit{Shepherdia}_{crown}$	19.8	(16.1–24.4)	12.3	(2.7–29.5)	10.9	(2.7–28.1)
$\textit{Hippophae}_{crown}$	19.5	(16.5–23.1)	8.3	(4.2–12.8)	8.2	(4.1–12.8)

Table S3-6 Divergence time estimates and ancestral area reconstructions of *Hippophae* L. Results from the ancestral area reconstructions are reported for the first (α) and the second (β) most frequent areas. BEAST posterior probabilities (PP) were also provided for each node

Node	PP	Ancestral areas				Age					
		α	%	β	%	PL		UCLN-lognormal		UCLN-uniform	
Hippophae$_{stem}$	1.00	A	68.8	B	17.6	33.9	(32.9–35.9)	34.9	(33.9–36.6)	35.9	(33.9–39.4)
Hippophae$_{crown}$	1.00	A	81.1	B	11.2	18.8	(17.7–22.8)	21.2	(18.3–24.0)	21.6	(18.4–25.0)
H. gyantsensis$_{stem}$	1.00	A	89.8	AB	5.6	17.8	(15.1–23.0)	19.7	(16.7–22.6)	20.0	(16.7–23.4)
H. gyantsensis$_{crown}$	1.00	A	99.3	AD	0.3	1.9	(0.5–4.3)	2.5	(1.2–3.8)	2.5	(1.2–3.9)
H. salicifolia$_{stem}$	1.00	A	67.0	B	15.0	16.9	(14.2–21.9)	17.6	(14.5–20.5)	17.9	(14.5–21.2)
H. salicifolia$_{crown}$	1.00	A	75.9	AD	11.0	3.6	(2.4–5.5)	2.5	(1.2–4.0)	2.6	(1.2–4.1)
H. neurocarpa$_{crown}$	1.00	B	87.4	AB	4.3	3.6	(2.5–5.5)	4.4	(2.9–5.9)	4.4	(2.9–6.0)
ssp. *stellatopilosa*$_{crown}$	0.94	B	95.0	BC	3.7	3.1	(2.0–4.8)	3.6	(2.2–5.2)	3.7	(2.2–5.1)
ssp. *neurocarpa*$_{crown}$	1.00	B	79.0	BC	11.7	0.3	(0.0–1.2)	0.7	(0.2–1.3)	0.7	(0.2–1.3)
H. tibetana$_{stem}$	1.00	A	52.7	B	31.8	16.7	(13.4–22.5)	18.4	(15.7–21.5)	18.7	(15.7–22.0)
H. tibetana$_{crown}$	1.00	A	86.4	B	2.8	3.0	(1.6–5.2)	3.6	(2.1–5.3)	3.7	(2.1–5.3)
H. rhamnoides$_{crown}$	1.00	B	96.7	AB	2.6	11.4	(9.9–14.8)	12.9	(10.7–15.2)	13.1	(10.7–15.6)
East/West QTP split	0.91	B	94.4	BD	1.4	8.0	(7.0–9.7)	8.8	(7.8–10.0)	8.9	(7.8–10.2)
ssp. *mongolica*$_{stem}$	1.00	B	88.5	BF	6.7	2.7	(1.6–4.5)	3.7	(2.4–5.1)	3.7	(2.3–5.1)
ssp. *mongolica*$_{crown}$	1.00	F	93.3	BF	5.6	1.2	(0.5–2.1)	1.6	(0.8–2.6)	1.6	(0.7–2.5)
ssp. *sinensis*$_{crown}$	1.00	B	98.3	BE	0.7	1.2	(0.6–2.6)	2.4	(1.3–3.7)	2.4	(1.2–3.6)
ssp. *turkestanica*$_{stem}$	1.00	D	53.6	B	27.3	7.2	(6.4–8.4)	7.6	(7.2–8.2)	7.7	(7.2–8.5)
ssp. *turkestanica*$_{crown}$	1.00	D	99.1	BD	0.3	2.9	(2.1–4.2)	3.4	(2.1–4.8)	3.5	(2.2–4.9)
ssp. *caucasica*$_{stem}$	1.00	G	69.9	I	7.4	5.2	(4.5–6.5)	5.8	(4.7–7.1)	5.8	(4.5–7.1)
ssp. *caucasica*$_{crown}$	1.00	G	99.4	GH	0.1	1.8	(1.1–2.7)	2.1	(1.1–3.1)	2.1	(1.1–3.2)
ssp. *fluviatilis*$_{stem}$	1.00	H	44.2	I	37.4	4.4	(3.5–5.4)	4.6	(3.2–6.0)	4.6	(3.2–6.1)
ssp. *fluviatilis*$_{crown}$	1.00	I	98.6	HI	1.0	0.5	(0.2–1.0)	0.9	(0.2–1.7)	0.9	(0.3–1.7)
ssp. *carpatica*$_{stem}$	1.00	H	98.9	HI	0.6	0.9	(0.3–1.4)	1.1	(0.4–1.9)	1.1	(0.3–1.9)
ssp. *carpatica*$_{crown}$	1.00	H	99.7	GH	0.1	0.03	(0.0–0.2)	0.2	(0.0–0.6)	0.2	(0.0–0.6)

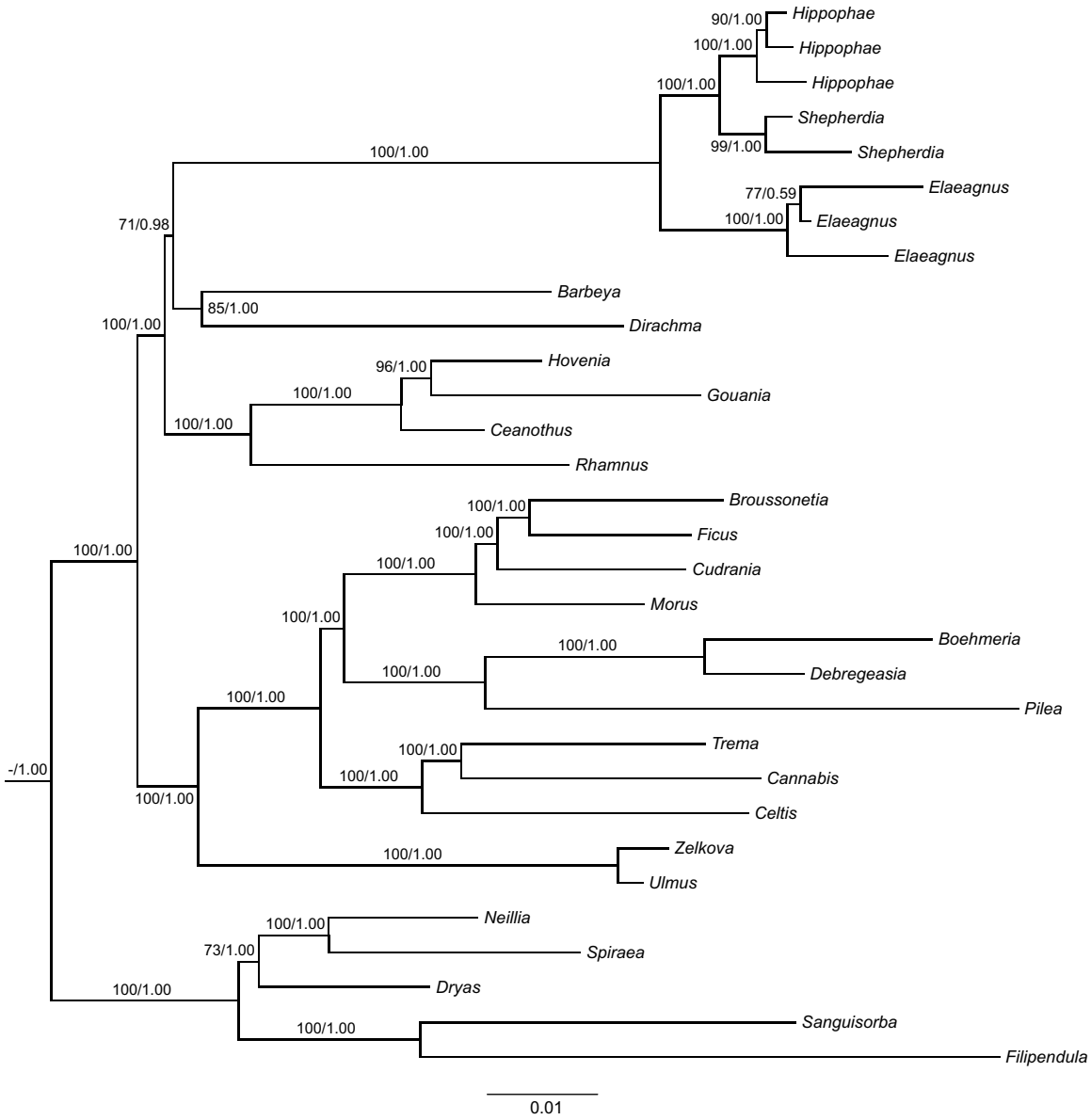


Fig. S3-1 The ML tree of Rosales. Support values (ML Bootstrap/Bayesian posterior probability) are provided at nodes.

Chapter IV. Genetic diversity of a pioneer tree is associated with both extent and direction of change in its climatic niche suitability

Abstract

Climate change potentially represents one of the most powerful drivers of range shifts for many species, resulting in considerable extinction at local, regional, and global scales. To predict the evolutionary consequences of climate change we need to understand the consequences of changes in climatic niche of species on genetic diversity within its populations. Association between habitat stability and high genetic diversity has long been assumed, but methods to test this association within species have been developed only recently with advance of Ecological Niche Models (ENM). Here we argue that estimation of both extent and direction of change in climatic niche suitability through time can be more informative than estimation of habitat stability *per se*. We test two hypotheses of association between genetic diversity and changes in climatic niche suitability: "climatic niche stability" and "climatic niche trend." For these tests we used data on ENM and microsatellite variation in a sample of populations of *H. rhamnoides* L. (Elaeagnaceae) from one of its Chinese subspecies (ssp. *sinensis* Rousi), representing the whole range of the taxon. Our analyses revealed that two distinct genetic groups were respectively distributed in the western and eastern regions of the range. Environmental factors were found to be the major drivers of the west/east split. Intra-population genetic diversity was found to have a significant negative relationship with the niche suitability

during LGM, and a significant or marginally significant positive relationship with the trend in niche suitability change after the LGM. These results suggest strong impact of historical climatic niche suitability and both extent and direction of changes in niche suitability throughout the Late Quaternary on current within-population genetic diversity in a pioneer plant species from Eastern Asia.

Key words: climatic change, ecological niche models, Elaeagnaceae, gene flow, habitat fragmentation, the Last Glacial Maximum, Loess Plateau, microsatellites, Qinghai-Tibetan Plateau, the Pleistocene

Introduction

It is becoming increasingly obvious from multidisciplinary evidence that past climatic changes, especially the climatic fluctuations of the Quaternary, have contributed greatly in shaping the current spatial patterns of species distribution and genetic diversity of organisms (Hewitt 2000, 2004). On the other hand, an ongoing rapid and profound climate change, with an average global temperatures rise of up to 4 °C by the end of this century (IPCC 2007) and associated changes in precipitation patterns, is expected to have increasing impacts and potentially present threats to biodiversity (Thuiller *et al.* 2005). Species react to climatic changes either by migrating to newly emerging suitable habitats, or by adjusting to new environments *in situ* through phenotypic plasticity and evolutionary adaptation, or by dying out in areas that have become unsuitable (Pauls *et al.* 2013). Any responses of surviving species, however, will influence the amount and structure of their intraspecific genetic diversity, which provide the main basic substrate for any evolutionary change. Intraspecific genetic diversity has been recognized to be important for maintenance of the fitness of individuals (Chapman *et al.* 2009) and of their resistance to stress (Markert *et al.* 2010; Nowak *et al.* 2007), and thus for adaptive capacity of populations (Frankham 2005). Therefore, a detailed understanding of the effects of historic climate alterations on intraspecific genetic diversity can provide valuable insights into the evolutionary consequences of past climate changes and predicting the likely direction of global warming effects on sustainability of extant populations and species (Temunović *et al.* 2013).

It has been suggested that climatically stable areas through major periods of climatic changes accumulate higher diversity of species in comparison with regions that have

experienced greater climatic shifts (Araújo *et al.* 2008; Cronk 1997; Graham *et al.* 2006; Jansson 2003). With the development of ecological niche models (ENM), this historic climatic habitat stability hypothesis was applied to both types of regions. Although higher species richness and biodiversity for regions of high climatic stability has been confirmed in general (e.g. Araújo *et al.* 2008; Graham *et al.* 2006), reports on genetic diversity within species were less conclusive. Genetic diversity could be associated with habitat stability negatively (Ortego *et al.* 2012), positively (Gugger *et al.* 2013), or no association between these parameters could be found (Edwards *et al.* 2012). The incongruence of these tests could result from different extent and, consequently, impact on biota of historical climatic changes in different regions, from different biological characteristics of studied species, or from different evolutionary histories of these species. Therefore, relationship between genetic diversity and climate change obviously requires additional investigation using a larger number of species across the whole Tree of Life sampled from different regions around the globe.

The late Quaternary ice-sheets and glaciers were probably of limited extent in China and only developed discontinuously on the Qinghai-Tibetan Plateau (QTP) and adjacent high mountains in West China (Zhou *et al.* 2006). Despite this, climatic oscillations during the late Quaternary likely had a considerable impact on most of the vegetation of China (Harrison *et al.* 2001). The QTP and adjacent areas harbour numerous ancient species as well as the newly originated species, and thus served as an important biodiversity museum and cradle, likely because of its complex topology (López-Pujol *et al.* 2011a, b; Wu 1988). For the species of this region, two main scenarios responding to climatic oscillations were suggested (see Liu *et al.* 2012; Qiu *et al.* 2011). First, species

may have retreated to the eastern QTP and southwestern China during glacial periods and re-colonized the QTP platform (e.g. Chen *et al.* 2008b; Meng *et al.* 2007; Yang *et al.* 2008; Zhang *et al.* 2005) during the interglacial phases or at the end of the Last Glacial Maximum (LGM). Second, some species may have survived the LGM in multiple refugia in the QTP and the postglacial expansions occurred mostly across adjacent populations at the local scale (e.g. Jia *et al.* 2011; Wang *et al.* 2009; Zhang *et al.* 2012). Compared to the abundant high mountains and deep valleys in the QTP region, the terrain of the northern China, extending from the Loess Plateau to the Northeast Plain, is less complex in topography. However, this region also covers diverse biomes from tropical/subtropical forests to grassland/steppe to cold forests and constitutes an important part of north-south vegetation transect in China. Similarly, for the species in northern China, some authors assumed that during the LGM cool-temperate deciduous forests shifted southwards below 30° N in the northwest and reaching 25° N in eastern China, and then re-colonized the previously unsuitable northern region occupied by steppe and even desert vegetation (Harrison *et al.* 2001; Ni *et al.* 2006; Yu *et al.* 2000). Others suggested regional expansions and multiple glacial refugia sparsely located in the mountains across northern China (Bai *et al.* 2010; Chen *et al.* 2008a; Tian *et al.* 2009; Zeng *et al.* 2011). Thus, it would be of great interest to investigate how populations in different regions of a species with a wide distribution covering both the QTP and northern China responded to the late Quaternary climatic changes.

Hippophae rhamnoides L., comprising nine subspecies (Lian *et al.* 2003; Swenson & Bartish 2002) (but see Tzvelev 2002 for an alternative taxonomy of the species) that are distributed widely and fragmentally in eastern Asia, central Asia, Asia Minor and Europe

(see Bartish *et al.* 2000; Jia *et al.* 2012), is a diploid (2n = 24; Rousi 1965), dioecious, wind-pollinated woody plant (Bartish & Swenson 2003). This shrub or small tree produces drupe-like fruits favouring dispersal by birds (Rousi 1971). It also reproduces asexually with root suckers and has symbiotic relationships with nitrogen-fixing *Frankia* (Tian *et al.* 2002), which enhance its growth in early successional, nutritionally poor, naturally disturbed, and strongly fragmented habitats, like riverbanks, mountain slopes and valleys. This species can endure extreme conditions, including temperature from -43 °C to 40 °C, high salinity and alkalinity (Li & Schroeder 1996; Lu 1992). Of the nine subspecies of *H. rhamnoides*, ssp. *sinensis* has the widest distribution in China ranging from the QTP to the Loess Plateau to the Northeast Plain. Previous studies using allozymes (Yao & Tigerstedt 1993), RAPDs (Bartish *et al.* 2000; Sun *et al.* 2006) and DNA sequences (Jia *et al.* 2012) demonstrated low levels of genetic diversity in ssp. *sinensis*. No clear spatial patterns of genetic differentiation were observed in this taxon (Jia *et al.* 2012; Sun *et al.* 2006). The QTP region was suggested as the ancestral area of ssp. *sinensis*, from where the subspecies could have experienced a significant range expansion northeastwards starting from the Last Interglacial Period (LIG) (Jia *et al.* 2012). Given its range and ecological characteristics, namely widespread and early-successional tree, *H. rhamnoides* ssp. *sinensis* can be an ideal model to study impact of historical climatic changes on genetic diversity within plant populations.

We examined variation of genetic diversity within populations of *H. rhamnoides* ssp. *sinensis* across its whole range in relation to shifts in climatic niche of this taxon throughout the Late Quaternary. The purpose of our study was to evaluate the extent to which current and past climate, and relatively mild the Late Quaternary climatic change

in China (Weaver *et al.* 1998), affected contemporary genetic variation in plants from this region. Using combined information from nuclear microsatellite markers and climate-based ENMs for both current and past periods, our aim was to (i) estimate the amount and geographic distribution of within-population genetic diversity of ssp. *sinensis*; (ii) explore putative genetic clusters within this subspecies and extent of differentiation between them; (iii) assess the role of isolation-by-distance and isolation-by-niche-difference in promoting and maintaining genetic differentiation of genetically coherent groups; (iv) evaluate how the distribution of current within-population genetic variation relates to both direction and extent of habitat suitability changes throughout the Late Quaternary. Specifically, we would like to test the (i) "climatic niche stability" (higher genetic diversity within populations is accumulated in localities, which experienced smaller change in habitat suitability through time) and (ii) "climatic niche trend" (higher genetic diversity within populations is accumulated in localities, which experienced positive trend in habitat suitability through time) hypotheses.

Materials and Methods

Plant sampling and laboratory procedures

We analyzed a total of 375 individuals of *H. rhamnoides* ssp. *sinensis*, taken from 24 natural populations covering the entire species range from the Qinghai-Tibetan Plateau to the Loess Plateau to the Northeast China Plain (Fig. 4-1, Table 4-1). We did not include

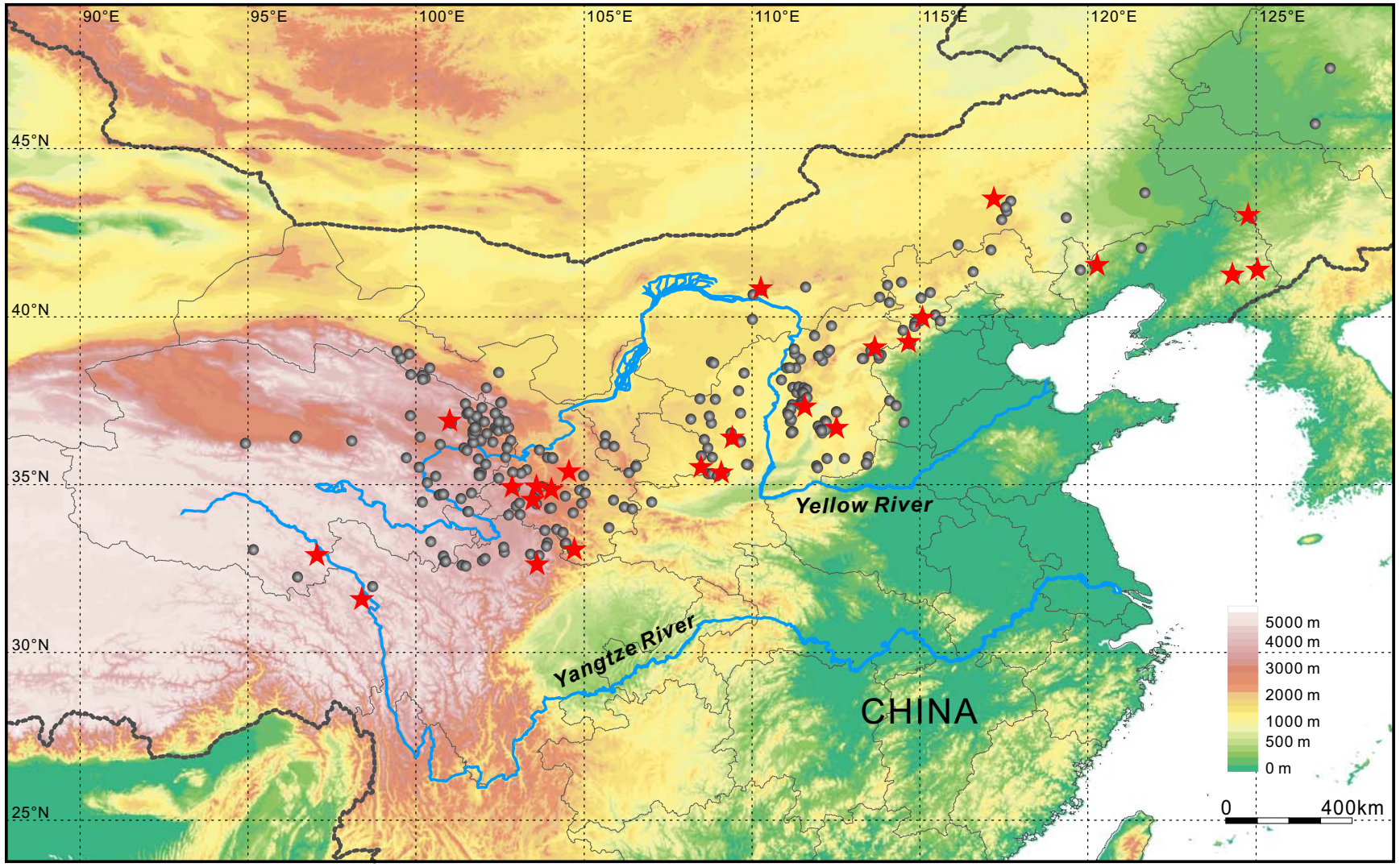


Fig. 4-1 Geographic distribution of herbarium records (grey dots) of *Hippophae rhamnoides* ssp. *sinensis* and the 24 populations (red stars) analyzed in the present study (see Table 4-1). The localities of the herbarium records were used to build the Ecological Niche Models for the subspecies.

in our analyses two populations from the total sample of the 26 populations used in the study of Jia *et al.* (2012). One of these populations revealed genetic patterns of introgression from another taxon (ssp. *yunnanensis*), so it most likely represents hybrid plants (Jia *et al.* 2012). Plant material of the second population was not available to us. The sample size of the populations varied from 9 to 19 individuals (mean = 15.6).

We extracted total genomic DNA from the silica gel dried leaves using DNeasy™ Tissue Kit (Qiagen GmbH, Hilden, Germany). For initial marker screening, we used a total of 20 SSR loci available for *Hippophae* so far (Jain *et al.* 2010; Wang *et al.* 2008). On the basis of the amplification efficiency, specificity, and observed polymorphism, 15 loci were selected for subsequent amplification (three with redesigned reverse primer; Table S4-1, Supporting information) using a three-primer system (Schuelke 2000). The PCR reactions were performed in 10 μL total volume containing 10 ng of genomic DNA, 10 × reaction buffer (Thermo Scientific, UK), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 U *Taq* DNA polymerase (Thermo Scientific, UK), 0.2 μM M13 primer (5'–TGTAAAACGACGGCCAGT–3') fluorescently labeled with either VIC, 6-FAM, NED or PET dyes (Applied Biosystems), 0.2 μM reverse primer and 0.04 μM forward primer extended with a 5'–M13 sequence tail. We used a touch-down PCR protocol conducted on a Mastercycler (Eppendorf, Germany). An initial denaturing cycle of 5 min at 95 °C was followed by 10 touchdown cycles of 30 s at 95 °C, 45 s at 60 °C (-0.5 °C per cycle) and 45 s at 72 °C. The touch-down cycles were followed by 30 cycles of 30 s at 95 °C, 45 s at 55 °C, 45 s at 72 °C, which in turn were followed by additional 10 cycles of 30 s at 95 °C, 45 s at 52 °C and 45 s at 72 °C, and a final extension at 72 °C for 10 min. Subsequently, amplification products of non-overlapping size and color were pooled

together at a ratio of 1:1:2:3 for VIC:6-FAM:NED:PET and diluted 1:50 with ddH_2O. Further, 1 μL of the diluted PCR product was added to 12 μL formamide and 0.1 μL GeneScan™-500 LIZ® size standard (Applied Biosystems, Foster City, CA, USA). We denatured samples for 5 min at 95 °C and cooled on ice before loading onto an ABI 3130 Genetic Analyzer (Applied Biosystems).

Population genetics analyses

Alleles were scored using PEAK SCANNER version 1.0 (Applied Biosystems), and binned into discrete size classes using FLEXIBIN macro (Amos *et al.* 2007). File conversions were carried out using the formatting programme CREATE version 1.37 (Coombs *et al.* 2008).

Genetic diversity

We tested for deviations from both Hardy-Weinberg equilibrium (HWE) for each locus, and genotypic linkage disequilibrium (LD) across all pairs of loci using Fisher's exact tests and the Markov chain method (10 000 dememorization steps, 1000 batches, with 10 000 iterations per batch set) in GENEPOP version 4.1.4 (Raymond & Rousset 1995; Rousset 2008). Levels of significance of the HWE and LD tests were corrected to minimize effects of type I error in case of multiple comparisons by applying a sequential Bonferroni correction with an initial α of 0.05 (Rice 1989). Where observed genotype frequencies deviated significantly from HWE expectations, MICRO-CHECKER version

2.2.3 (Van Oosterhout *et al.* 2004) was used to test for null alleles and allelic dropout using 1000 Monte-Carlo simulations and a Bonferroni corrected 95% confidence interval. Null allele frequencies were estimated using Brookfield's (1996) null allele estimator 2 (taking into account the presence of null allele homozygotes). To contrast with the uncorrected data, we generated an alternative data set, in which the observed allele and genotype frequencies were adjusted appropriately and each null allele was recorded as missing data.

The following genetic diversity parameters were measured as averages over loci within each population with the software GENALEX 6.5 (Peakall & Smouse 2006, 2012): the percentage of polymorphic loci (%*P*), number of alleles (N_A), number of effective alleles (N_E), Shannon's Information Index (*I*), observed heterozygosity(H_O), expected heterozygosity(H_E) and inbreeding coefficient F_{IS}. We also computed allelic richness (A_R) using the rarefaction procedure (El Mousadik & Petit 1996; Hurlbert 1971) implemented in HP-RARE (Kalinowski 2005) with a fixed sample size of 16 genes. This standardized parameter is independent of sample size and more suitable for comparing allelic diversity among populations of uneven sizes.

Genetic structure and differentiation

The Bayesian clustering algorithm implemented in STRUCTURE version 2.3.4 (Pritchard *et al.* 2000) was used to investigate population subdivision. We assessed likelihoods for models with K ranging from 1 to 10. For each value of K, 20 independent Markov chain Monte Carlo (MCMC) runs with a burn-in length of 100 000 iterations followed by an

additional 100 000 iterations were carried out using the admixture model with correlated allele frequencies and no prior population information. The optimal K was determined by comparing the log probabilities [ln P(K)] (Pritchard *et al.* 2000) and the standardized second-order rate change of ln P(K), ΔK (Evanno *et al.* 2005). Finally, permutations of the most likely results among different runs for each K were conducted in CLUMPP version 1.1.2 (Jakobsson & Rosenberg 2007), and the results were visualized by DISTRUCT version 1.1 (Rosenberg 2003).

Genetic relationships among populations were estimated by generating a neighbour-joining (NJ) tree based on the Nei's genetic distance D_A (Nei *et al.* 1983) with 1000 bootstrap replications, using the POPTREE2 software (Takezaki *et al.* 2010). D_A has been shown to have good discriminatory power for closely related populations (Nei *et al.* 1983) and is recommended for microsatellite data (Takezaki & Nei 1996). In addition, another distance-based clustering method, principle coordinate analysis (PCoA) was also performed on pairwise F_{ST} values among populations using GENALEX. The matrix of pairwise F_{ST} values was estimated with the software ARLEQUIN version 3.5.1.2 (Excoffier & Lischer 2010).

Hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) was carried out using ARLEQUIN (Excoffier & Lischer 2010), where variance components are estimated for three hierarchical levels: among groups, between populations within groups and within populations. Groups of populations were defined according to the results of the analyses described above.

We used the software BARRIER version 2.2 (Manni *et al.* 2004) to identify areas of genetic discontinuity or barriers to gene flow across the sampled species' range. BARRIER

connects neighbouring samples by Delaunay triangulation and then uses Monmonier's maximum-difference algorithm (Monmonier 1973) to place barriers across the network where the most significant changes in genetic continuity are found. We evaluated genetic differentiation between populations using pairwise values of F_{ST} and tested the robustness of the inferred barriers via a bootstrap analysis. We generated 100 bootstrapped datasets by resampling individuals within populations with replacement using POPTOOLS version 3.2.5 (Hood 2010). One hundred matrices of pairwise F_{ST} among populations were then calculated using an ARLEQUIN batch file and imported into BARRIER to generate bootstrapped barriers.

The spatial patterns of genetic differentiation among populations were also investigated by performing a genetic landscape shape interpolation procedure using the programme ALLELEINSPACE (AIS; Miller 2005). This method, based on Delaunay triangulations and Monmonier's maximum-difference algorithm, produces a three-dimensional graphical representation of genetic distance patterns and allows a better visualization of the heterogeneity of the genetic divergence patterns across a landscape. We used residual genetic distances that are derived from the linear regression of genetic versus geographical distance.

Finally, to determine if genetic divergence between populations was associated with their geographic distance, we conducted a Mantel test using genetic (F_{ST}) and geographic (km) distance matrices using GENALEX.

Ecological niche modeling

Ecological niche models (ENMs) for *H. rhamnoides* ssp. *sinensis* were generated using a

maximum entropy modeling technique as implemented in MAXENT version 3.3.3k (Phillips *et al.* 2006). This software combines presence-only data with ecological-climatic layers to predict species occurrence in areas where data points are unavailable.

A total of 295 localities were used to build the models, after removal of duplicate records within each pixel (30 arc-second, ~ 1 km; see below) (Fig. 4-1). These occurrence data were obtained from our own collections sites, literature records, and from herbarium record databases (Chinese Virtual Herbarium, http://www.cvh.org.cn and Global Biodiversity Information Facility, http://www.gbif.org). Nineteen environmental variables from the WorldClim database version 1.4 (http://www.worldclim.org; see Appendix) with 30 arc-second spatial resolution were used as the environmental layers (Hijmans *et al.* 2005).

We run MAXENT for 10 replicates with the default settings (i.e., convergence threshold of 10^{-5}, 10 000 background points, regularization multiplier of 1, and autofeatures) except for increasing the maximum number of iterations to allow the algorithm to converge naturally. For each run, 75% of the localities were randomly selected for model training, leaving the remaining 25% for model testing. Climatic variable importance was determined with a jackknife procedure by comparing the percent contribution of each variable to the model and the loss of predictive power when each variable was excluded. Overall model performance was evaluated using the area under the receiving operator characteristics curve (AUC), which ranges from 0.5 (randomness) to 1 (perfect discrimination).

The contemporary species-climate relationships were then projected on to past climate layers to predict the potential range of the species during the LGM (~21 000 BP) and the

LIG (~120 000–140 000 yr BP). The projection bioclimatic layers for the LIG (resolution 30 arc-second) and the LGM (resolution 2.5 arc-minute) were also obtained from the WorldClim database version 1.4 (Hijmans *et al.* 2005). For the LGM prediction, both the Community Climate System Model (CCSM) and the Model for Interdisciplinary Research on Climate (MIROC) were used. Finally, we applied a "maximum training sensitivity plus specificity" threshold for determining suitable/nonsuitable habitat, because sensitivity–specificity combined approaches were suggested as good thresholds (Liu *et al.* 2005).

Point based analysis of environmental variables and niche suitability

We performed point-based method analyses to determine: 1) whether the detected in genetic analyses groups were ecologically differentiated, and 2) whether genetic diversity of microsatellite loci was correlated with the bioclimatic variables as well as niche suitability for these populations. DIVA-GIS version 7.5 (Hijmans *et al.* 2001) was used to extract values for each of the 19 bioclimatic variables from the environmental layers and niche suitability scores from the current, LGM and LIG ENMs.

We compared the similarity of climatic niches of the detected genetic groups by conducting a principal component analysis (PCA). Subsequently, we performed multivariate analyses of variance (MANOVA) with PCA axis scores as dependent variables and genetic groups as categorical predictors, to assess whether the separation in the ecological niche was statistically significant. We also plotted the least-squares means (ls-means) to visually compare the differences in environmental variables between groups.

We regressed niche suitability of LGM and current (N_{LGM}, N_{CUR}), niche stability (N_{STA} = 1 - |N_{CUR} - N_{LGM}|) (Gugger *et al.* 2013; Ortego *et al.* 2012) and niche suitability change (N_{CHA} = N_{CUR} - N_{LGM}) from the LGM to current on genetic diversity in general linear models using STATISTICA version 8 (Statsoft Inc. Maisons-Alfort, France) to depict which variables contributed to explain genetic diversity. In order to reduce multi-collinearity among predictive variables we conducted a best-subset search, using adjusted R^2 as a search criterion. We also employed multiple univariate linear regression models to evaluate whether genetic diversity was correlated with the 19 bioclimatic variables and niche suitability variables (N_{CUR}, N_{LGM}, N_{LIG}, N_{STA} and N_{CHA}).

Results

Hardy-Weinberg Equilibrium and null alleles

The 15 microsatellite loci in our analyses of populations of *H. rhamnoides* ssp. *sinensis* varied widely in number of alleles (2–24; mean = 9.8) and expected heterozygosities (0.007–0.731; mean = 0.478) across all populations (Table S4-1). Significant departures from HWE were observed in 56 of 360 population-locus combinations (24 populations at 15 loci) after sequential Bonferroni corrections. Of 2520 locus-pair tests for LD, only 61 were significant and deviations were not consistently observed either among specific pairs of loci or within specific populations.

Failed amplifications occurred at loci Hrms003, Hrms004, Hrms014 and Hrms026 (Table S4-2), and their percentages to the total number of genotypes were 5.33%, 1.87%, 0.80% and 0.27%, respectively. Analysis with the second method of Brookfield (1996) as

implemented in MICRO-CHECKER revealed no evidence of large allele dropout, while presence of null alleles was suggested for all loci except for the least polymorphic locus Hrms028. The existence of stuttering was only indicated at locus Hrms014 with a di-tetra complex repeat motif ($(TG)_n(TA)_n$; Table S4-1). Estimated frequencies of null alleles per locus per population ranged from 0 to 0.699, with variation of frequencies averaged over loci from 0.022 to 0.167 depending on the population. The mean null allele frequency over all populations and loci was 0.106 (Table S4-3). To evaluate a possible bias in genetic structure due to the presence of null alleles, we calculated pairwise F_{ST} matrices among populations for both the original and the null allele corrected datasets, and then compared the two matrices in a Mantel's permutation test implemented in GENALEX. A high degree of correlation between the two matrices was indicated ($r^2 = 0.971$, $r = 0.985$, $P < 0.0001$). In addition, we compared the values of A_R produced by the two data sets in a simple linear regression model, and found a significant linear relationship between them confirmed by high correlation coefficient ($r^2 = 0.999$, $r = 0.999$, $P < 0.0001$). In the case of stuttering, both comparisons were conducted between the datasets with and without locus Hrms014. Similarly, strong and significant correlations between pairwise F_{ST} matrices ($r^2 = 0.989$, $r = 0.994$, $P < 0.0001$) and A_R values ($r^2 = 0.991$, $r = 0.996$, $P < 0.0001$) of the two datasets could be confirmed. Therefore, as the possibility that the presence of null alleles and stuttering had affected the detected genetic pattern could be ruled out, we finally used the original data set for all the subsequent analyses.

Genetic diversity of populations

A total of 147 alleles were detected across the 15 loci (Table S4-1), defining 294 unique

multilocus genotypes (Table 4-1). Of the 30 multilocus genotypes that were shared between individuals, only one occurred in both populations 14 and 15, while the others were restricted to a single population. The number of genotypes detected per population ranges from three to 18, and the smallest proportions of distinct genotypes (G/N) were found in populations 11 and 13 (0.18; Table 4-1).

Measures of genetic variation varied across populations, with the percentage of polymorphic loci (%*P*) ranging from 46.67% (population 13) to 93.33% (populations 05, 16–18 and 21–24); the average number of different alleles across loci (N_A) from 1.47 (population 13) to 5.33 (population 24); the number of effective alleles (N_E) from 1.40 (population 13) to 3.60 (population 08); and the Shannon's information index (*I*) from 0.29 (population 13) to 1.22 (population 08). The observed heterozygosity (H_O) varied from 0.23 (population 15) to 0.43 (population 11), while the expected heterozygosity (H_E) from 0.20 (population 13) to 0.59 (population 08). After rarefaction with a fixed sample size of 16 genes, allelic richness ranged from 1.43 in population 13 to 4.45 in population 08 with a mean value of 3.49; and private allele richness ranged from 0 in populations 06, 11 and 13 to 0.24 in population 10, with a mean value of 0.06. Overall, the three populations located in the central part of the range, i.e. populations 11–13, tended to maintain the lowest levels of intrapopulation genetic diversity (Fig. 4-3, Table 4-1). An excess of heterozygotes ($H_O > H_E$) was detected only in populations 11 and 13, leading to negative values for the fixation index (F_{IS} = -0.28 and -0.58, respectively; Table 4-1), which could be explained by the occurrence of fixed heterozygosity at some loci in the two populations. Furthermore, significant linear relationships were observed between A_R and the other genetic diversity indices with high correlation coefficients (A_R:N_A: r^2 =

Table 4-1 Geographical location and genetic diversity estimates for the studied populations of *Hippophae rhamnoides* ssp. *sinensis*. Number of individuals sampled (*N*), number of distinct multilocus genotypes (*G*), proportions of distinct genotypes (*G*/*N*), percentage of polymorphic loci (%*P*), number of different alleles (N_A), number of effective alleles (N_E), Shannon's information index (*I*), expected heterozygosity (H_E), fixation index (F_{IS}), allelic richness (A_R) and private allele richness (PA_R)

Pop	location	latitude(°N)	longitude(°E)	*N*	*G*	*G*/*N*	%*P*	N_A	N_E	*I*	H_O	H_E	F_{IS}	A_R	PA_R
01	Yushu, QH	32.89	97.06	11	11	1.00	86.67%	4.73	3.32	1.17	0.29	0.58	0.46	4.39	0.13
02	Jiangda, XZ	31.59	98.38	13	13	1.00	80.00%	3.20	2.14	0.80	0.24	0.44	0.45	3.00	0.06
03	Haiyan, QH	36.88	101.01	9	9	1.00	80.00%	3.87	2.71	0.97	0.30	0.51	0.44	3.75	0.02
04	Hezuo, GS	34.92	102.86	9	9	1.00	86.67%	4.20	2.95	1.05	0.37	0.53	0.26	4.08	0.03
05	Zhuoni, GS	34.61	103.51	12	12	1.00	93.33%	3.93	2.19	0.88	0.26	0.46	0.43	3.50	0.05
06	Lintan, GS	34.75	103.55	17	16	0.94	86.67%	4.40	2.90	1.02	0.29	0.51	0.38	3.75	0.00
07	Songpan, SC	32.61	103.60	17	13	0.76	80.00%	4.13	2.66	0.98	0.35	0.50	0.28	3.59	0.07
08	Zhangxian, GS	34.84	104.03	16	16	1.00	86.67%	5.20	3.60	1.22	0.36	0.59	0.42	4.45	0.18
09	Dingxi, GS	35.38	104.54	17	16	0.94	86.67%	5.20	3.38	1.11	0.33	0.53	0.37	4.21	0.04
10	Wenxian, GS	33.05	104.65	11	10	0.91	80.00%	4.00	2.64	0.94	0.28	0.47	0.44	3.72	0.24
11	Zhengning, GS	35.52	108.48	17	3	0.18	66.67%	2.07	1.67	0.49	0.43	0.31	-0.28	1.98	0.00
12	Yijun, SN	35.35	109.09	18	8	0.44	86.67%	2.60	1.81	0.59	0.29	0.36	0.18	2.22	0.11
13	Ganquan, SN	36.39	109.41	17	3	0.18	46.67%	1.47	1.40	0.29	0.34	0.20	-0.58	1.43	0.00
14	Baotou, IM	40.79	110.29	15	13	0.87	86.67%	4.67	3.00	1.07	0.35	0.53	0.33	4.01	0.02
15	Fenyang, SX	37.31	111.57	17	8	0.47	80.00%	3.87	1.86	0.74	0.23	0.38	0.45	3.03	0.04
16	Qinxian, SX	36.68	112.53	18	15	0.83	93.33%	4.07	2.85	0.98	0.32	0.49	0.39	3.54	0.07
17	Fanzhi, SX	39.08	113.65	17	13	0.76	93.33%	4.00	2.39	0.93	0.36	0.49	0.22	3.36	0.03
18	Laiyuan, HB	39.24	114.66	16	8	0.50	93.33%	2.87	2.14	0.80	0.35	0.47	0.37	2.73	0.05
19	Zhuolu, HB	39.96	115.07	17	11	0.65	86.67%	3.93	2.48	0.91	0.33	0.46	0.40	3.41	0.05
20	Keshiketeng, IM	43.51	117.19	18	18	1.00	86.67%	4.60	2.90	1.03	0.40	0.52	0.23	3.73	0.04
21	Zhaoyang, LN	41.53	120.28	18	17	0.94	93.33%	4.67	2.76	1.06	0.38	0.53	0.26	3.82	0.02
22	Bengxi, LN	41.25	124.29	18	18	1.00	93.33%	5.27	3.17	1.13	0.39	0.55	0.28	4.13	0.04
23	Xifeng, LN	43.02	124.77	19	17	0.89	93.33%	4.27	2.83	1.00	0.38	0.51	0.20	3.57	0.05
24	Huanren, LN	41.40	125.05	18	17	0.94	93.33%	5.33	3.16	1.19	0.40	0.58	0.27	4.31	0.07
Total				**375**	**294**	**0.78**	**84.00%**	**4.02**	**2.62**	**0.93**	**0.33**	**0.48**	**0.30**	**3.49**	**0.06**

Abbreviation: GS, Gansu; HB, Hebei; IM, Inner Mongolia; LN, Liaoning; QH, Qinghai; SC, Sichuan; SN, Shaanxi; SX, Shanxi; XZ, Xizang.

0.926, r = 0.962, $P < 0.0001$; A_R:N_E: $r^2 = 0.893$, r = 0.945, $P < 0.0001$; A_R:I: $r^2 = 0.971$, r = 0.985, $P < 0.0001$; A_R:H_E: $r^2 = 0.915$; r = 0.957, $P < 0.0001$), we thus used A_R standardized for sample size as the intrapopulation genetic diversity in further analyses.

Genetic structure and differentiation

In the Bayesian clustering analysis implemented in STRUCTURE, Ln P(K) values increased progressively from K = 1 to K = 10, and did not show a clear plateau (Fig. S4-1). The ΔK statistic of Evanno *et al.* (2005), however, showed clearly presence of two distinct genetic groups corresponding to the geographic regions: the western one was composed of populations 1–10 and 13; and the eastern one consisted of populations 11–12 and 14–24. Increasing K above K = 2 did not yield a progressive breakdown of the cluster definition of the eastern and western groups, but only introduced heterogeneity in membership proportions (especially in the three central populations 11–13), with the individuals having an assignment probability greater than 90% decreasing from 72.27% at K = 2 to 50.40% at K = 5. The neighbor-joining tree of the 24 populations based on D_A distances (Nei *et al.* 1983) supported the split between the western and eastern populations, and separated the three "outsider" central populations 11–13 as an additional group with long branch lengths, although the nodes of this tree were not supported with high bootstrap values (Fig. S4-2). In addition, we found from the NJ tree that the populations of the western group were connected by longer branches compared to those populations of the eastern group. The other distance-based clustering method, principle coordinate analysis

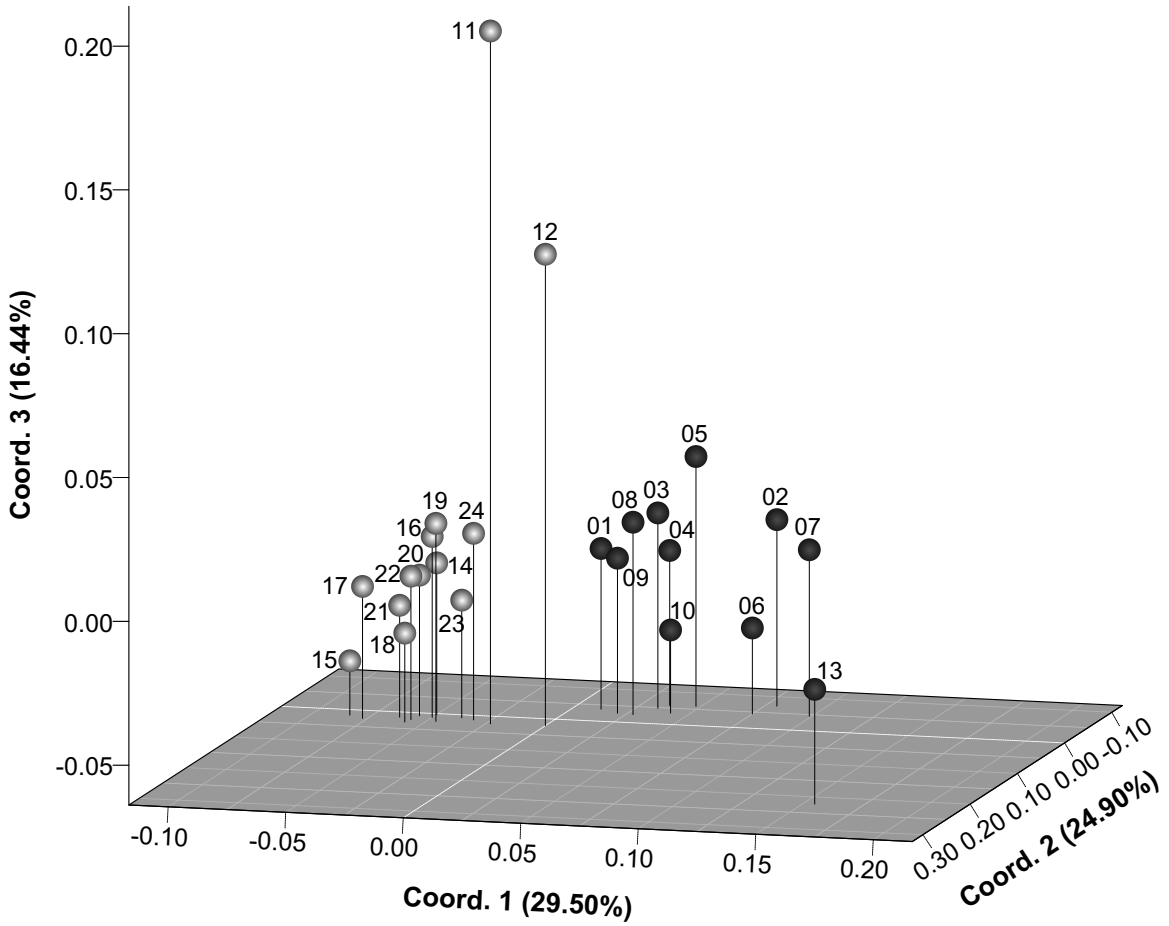


Fig. 4-2 Plot of the first three principal coordinates obtained from the PCoA analysis of the 24 populations based on pairwise F_{ST} values. Numbers above circles correspond to population codes (as in Table 4-1).

(PCoA) on pairwise F_{ST} values showed similar patterns to the STRUCTURE analyses and the populations could also be divided largely into western and eastern groups on the right and left sides of the first PCoA axis (explaining 29.50% of the variation) (Fig. 4-2). Population 13 differentiated from the western group along the second PCoA axis (24.90% of the variation), while populations 11 and 12 were separated from the eastern group along the third PCoA axis (16.44% of the variation).

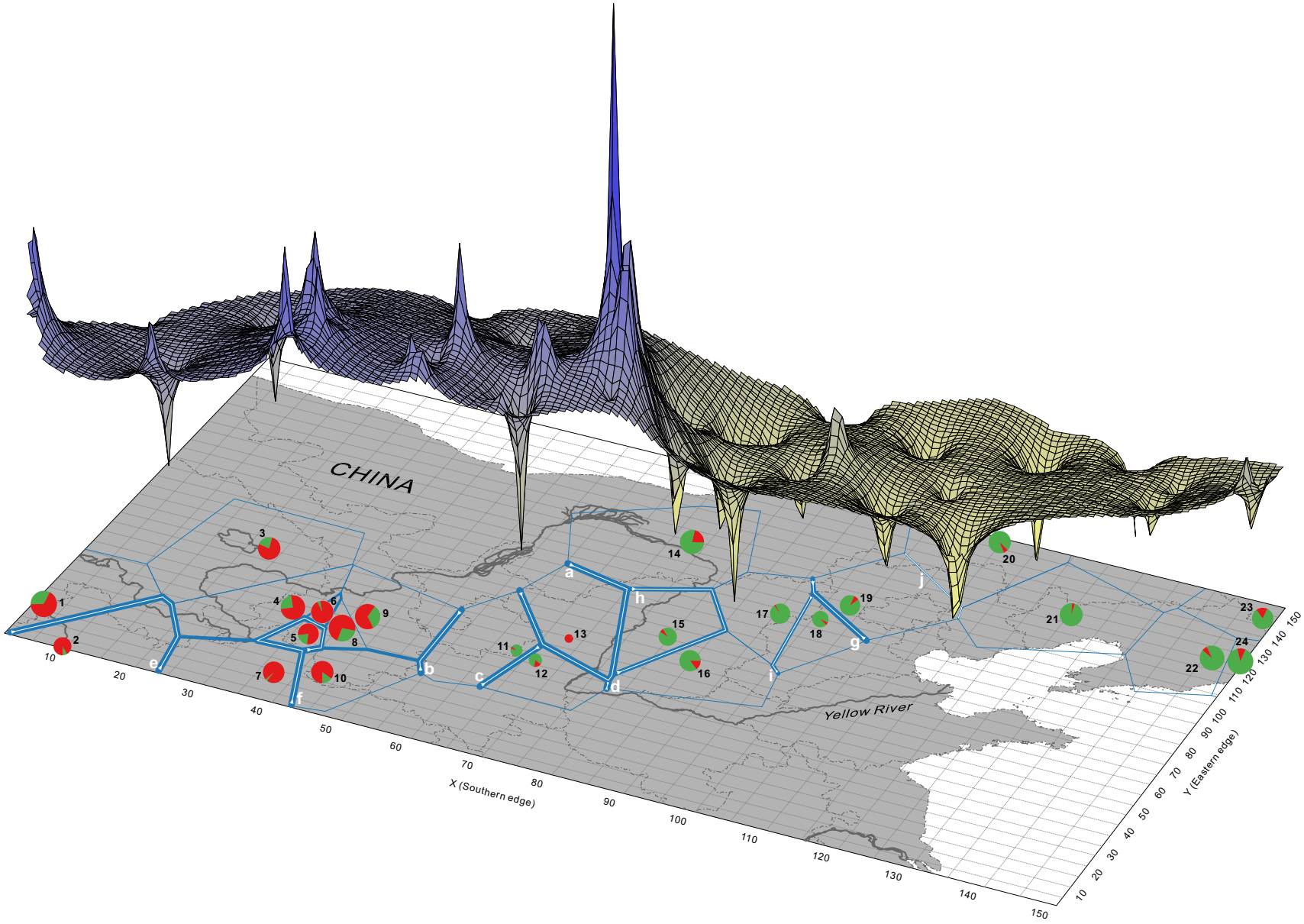


Fig. 4-3 Frequency distribution of the two genetic clusters as revealed in Bayesian STRUCTURE analysis with K = 2, the results of the Genetic Landscape Shape Interpolation (GLSI) analysis, and the BARRIER analysis. Populations are numbered as in Table 4-1 and presented as pie diagrams, with slice size proportional to the frequency of individuals assigned to each of two groups (red for the "western", and green for the "eastern" group). The GLSI surface plot heights represent genetic distance. The barriers are numbered (small letters a–j) in order of importance and their thickness is proportional to their bootstrap support. Circle size is proportional to the allelic richness (A_R) at each location.

Analysis of Molecular Variance (AMOVA) based on all samples indicated that majority of genetic variation (81.64%) occurred within populations (Table S4-4). The AMOVA with three hierarchical levels assigned 79.12% of the total variance within populations and only 6.18% between the two groups. Additionally, two-level AMOVA tests calculated for each of the two groups indicated higher differentiation within the western group (19.35%) than within the eastern group (13.06%). When the three central populations 11–13 were excluded from the analyses, the same trend was observed.

Genetic Landscape Shape Interpolation analysis showing the geographic pattern of distribution of genetic differentiation is presented in Fig. 4-3. The largest genetic distance occurred between populations 13 and 15. Two main areas were clearly present: one area of high differentiation (indicated by positive peaks) in the western and central portions of the range, and one area of prevailing homogeneity (indicated by troughs) in the eastern portion. Correspondingly, most genetic discontinuities were identified in the western and central portions of the range of ssp. *sinensis* (barriers a–f; Fig. 4-3). The four most strongest barriers (a–d), with segments having the highest bootstrap values (100%), separated central populations 11–13 from all other samples. The Mantel test involving all populations revealed no significant pattern of isolation by distance ($r^2 = 0.004$, r = -0.063, $P = 0.297$) (Fig. S4-3A). However, exclusion of populations 11–13 resulted in a weak but significant IBD ($r^2 = 0.056$, r = 0.236, $P = 0.0006$) (Fig. S4-3B).

Ecological niche modeling

The ENM based on the entire set of 19 bioclimatic variables yielded high AUC scores for

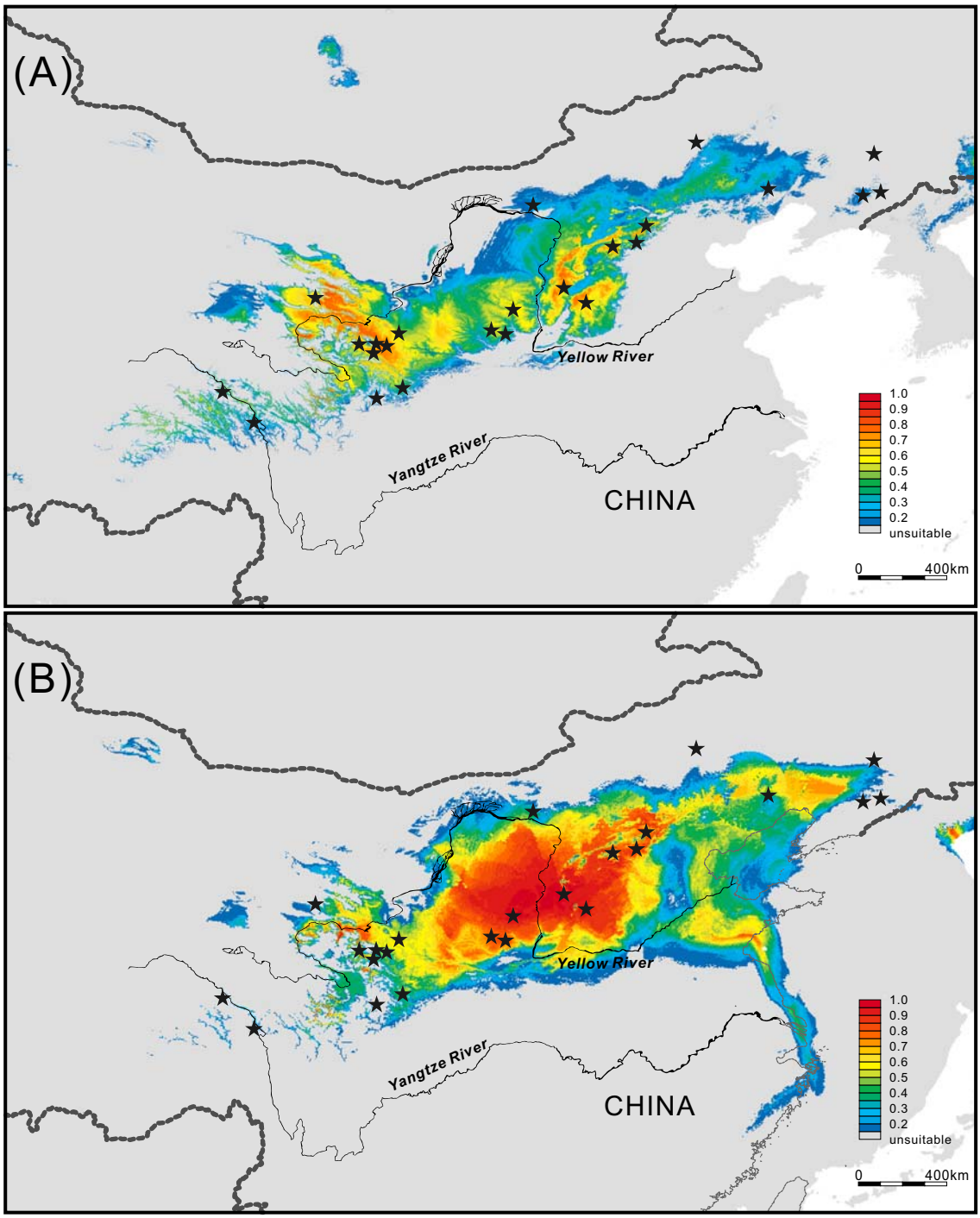
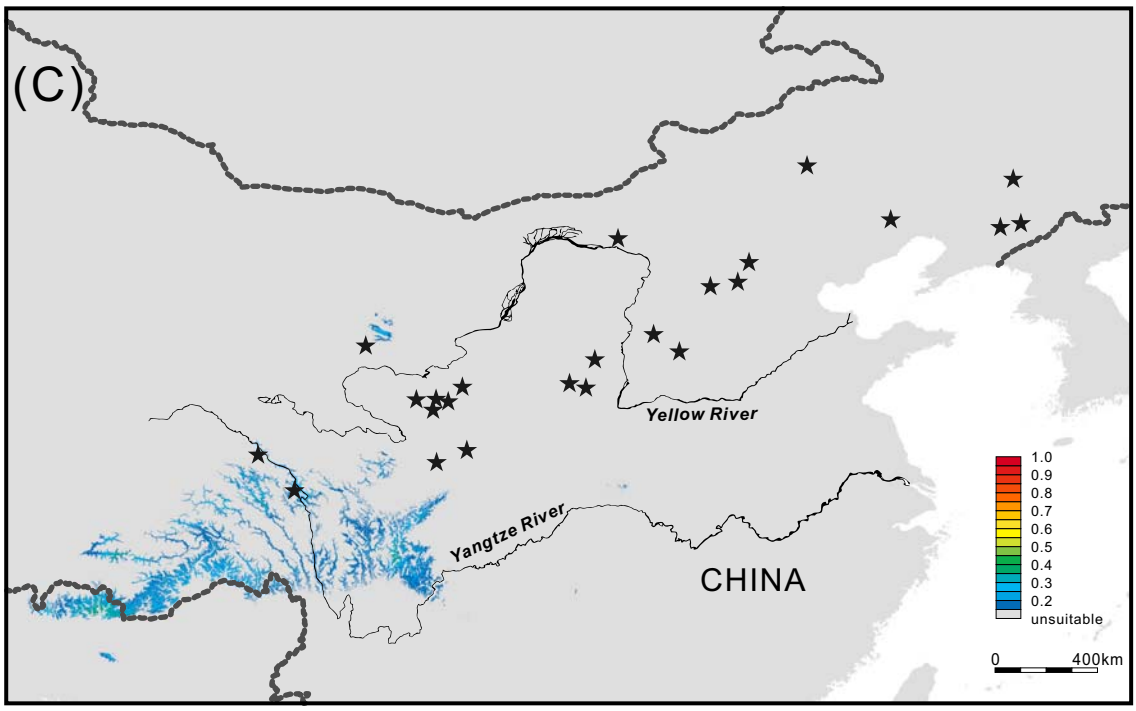


Fig 4-4 Ecological niche models (ENMs) for *Hippophae rhamnoides* ssp. *sinensis*, estimated from current conditions (A), for the Last Glacial Maximum based on Community Climate System Model (CCSM) (B), and for the Last Interglacial Period (C). Warmer colors show areas with higher suitability. The logistic threshold under the "maximum training sensitivity plus specificity" criterion for presence–absence prediction is 0.130. Black stars mark locations of the 24 sampled populations (as in Fig. 4-1 and Table 4-1).



both the training (0.967 ± 0.002 SD; 10 replicates) and the test data (0.953 ± 0.009 SD), indicating a high fit between the modeled and the actually observed current distribution. The jackknife test using the AUC on test data revealed that variable absences did not affect strongly the model, as omitting each one variable in turn changed the AUC only slightly (with the partial-model AUC mean values ranging from 0.952 to 0.954) (Fig. S4-4). When the models were calculated only with one variable, the contribution varied strongly among variables. Annual precipitation (BIO12), precipitation of warmest quarter (BIO18), precipitation of wettest quarter (BIO16) were the strongest predictive variables, while mean diurnal range (BIO2) was the weakest.

The predicted current distribution (Fig. 4-4A) showed largely continuous suitable habitat for ssp. *sinensis*, with two distinct zones of higher suitability being located in the west and east areas of its range. Although suitable habitat in the LGM based on CCSM

and MIROC differed slightly, both approaches suggested an overall retention of suitable habitat with eastward and westward shifts at the both edges of the range of ssp. *sinensis* (Fig. 4-4B and Fig. S4-5). The niche model reconstructions for LGM indicate the most suitable areas mainly in the central portion of the range. The projection of the model over the LIG layers indicated that suitable areas were strongly reduced and mostly limited to the southwest part of China, i.e. the southeast Qinghai-Tibetan Plateau (QTP) and adjacent areas (Fig. 4-4C). In addition, the prediction also suggested a general decrease in habitat suitability over the range of ssp. *sinensis* during the LIG period.

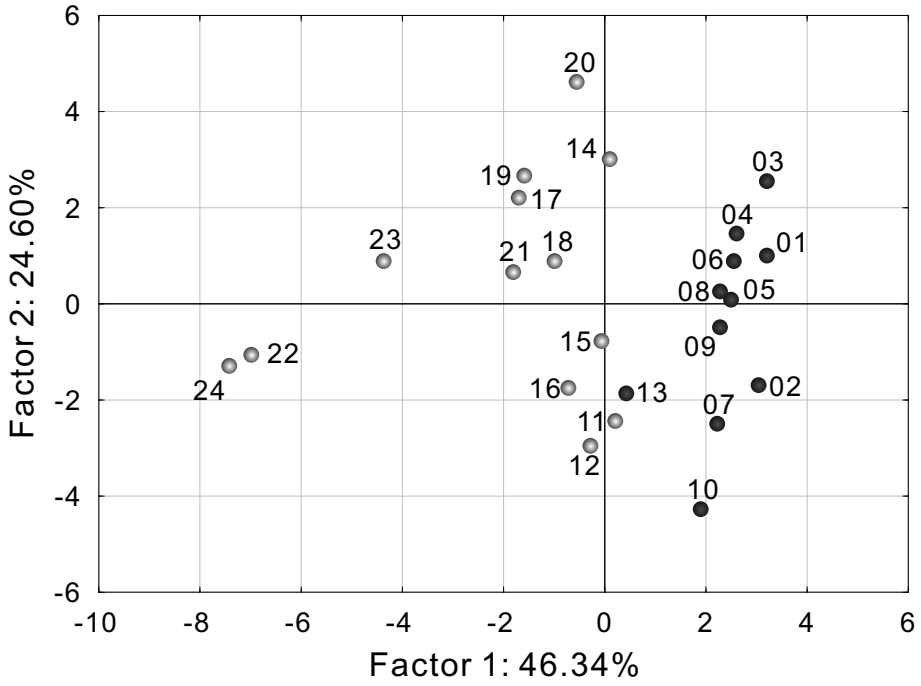


Fig. 4-5 Graphical representation of principal component analysis (PCA) of climatic niche differences between 24 populations of *Hippophae rhamnoides* ssp. *sinensis*. The first two largest principal components are plotted. The first principal component explains 46.34% and the second one 24.60% of the total variance of climatic variables. Populations from the "western" group are indicated in black, and from the "eastern" group – in grey. Numbers correspond to population codes (as in Table 4-1).

Point-based analysis of climatic variables

The PCA revealed four components with eigenvalues greater than 1 that collectively explained > 97% of the total variance, with 46.34%, 24.60%, 17.88% and 8.31% of the total variance allocated to PCs 1–4, respectively (Table S4-5). The scatter plot of the first two principal component scores indicated that the two groups of populations are distributed in distinct environments (Fig. 4-5). The result also suggests that the eastern group covers a wider environmental niche compared with the western group. The MANOVA analysis of the scores of the first four components suggested the separation in the ecological niche between the two groups of populations was statistically significant (Wilks' lambda = 0.249; F (4, 19) = 14.298; P = 0.00002), which was largely driven by PC1 (F (1, 22) = 28.949, P = 0.00002). All the variables related to rainfall patterns loaded strongly negatively on PC1 (Fig. S4-6 and Table S4-5). The variables that measure seasonal temperature variation (Bio4) and annual temperature range (Bio7) also had a high negative loading on PC1, while isothermality (Bio3) that quantifies diurnal temperature oscillation had a high positive loading. Additionally, individual ANOVAs indicated that with exception of annual mean temperature (Bio1) and annual precipitation (Bio12), the rest 17 bioclimatic variables differed significantly or marginally significantly between the two groups (Table S4-6).

Regression analyses of association of intrapopulation genetic diversity with climatic variables, niche suitability estimates, and estimates of trend in niche suitability change after the LGM

Table 4-2 General linear models for allelic richness (A_R) in relation to current niche suitability (N_{CUR}), Last Glacial Maximum (LGM) niche suitability (N_{LGM}) derived from CCSM and MIROC models, niche suitability change ($N_{CHA} = N_{CUR} - N_{LGM}$) and niche stability ($N_{STA} = 1 - |N_{CUR} - N_{LGM}|$) from the LGM to present. Significant p-values ($P < 0.05$) are shown in bold type

	CCSM				MIROC			
	Estimate	SE	*t*	*P*	Estimate	SE	*t*	*P*
(Intercept)	4.1065	0.2043	20.1032	< **0.0001**	3.9345	0.1932	20.3645	< **0.0001**
N_{CUR}	–				–			
N_{LGM}	-1.5264	0.4003	-3.8130	**0.0010**	-1.6272	0.5124	-3.1756	**0.0044**
N_{CHA}	–				–			
(Intercept)	2.1136	1.0200	2.0723	0.0514	3.9345	0.1932	20.3645	< **0.0001**
N_{CUR}	1.6338	0.8452	1.9331	0.0675	–			
N_{LGM}	-2.0178	0.4681	-4.3103	**0.0003**	-1.6272	0.5124	-3.1756	**0.0044**
N_{STA}	1.8830	1.0195	1.8470	0.0796	–			

SE denotes standard error; –, parameters rejected in the model.

Allelic richness (A_R) variation within populations of ssp. *sinensis* was strongly negatively associated with LGM niche suitability (N_{LGM}) (Table 4-2). Current niche suitability (N_{CUR}) and niche stability (N_{STA}) also appeared in the model and had a positive but marginally significant relationship on A_R only when CCSM model was adopted. No relationship was found between A_R and niche suitability change (N_{CHA}). Simple linear regression analyses demonstrated that allelic richness was negatively correlated with annual mean temperature (Bio1) and minimum temperature of the coldest month (Bio6) (Figs. S4-7A and S4-7B). In the case of the niche suitability estimates, allelic richness had a significant negative relationship with the niche suitability during LGM (N_{LGM}) (Fig. 4-6A; Fig. S4-7E), and a significant or marginally significant positive relationship with the trend in niche suitability change after the LGM ($N_{CHA} = N_{CUR} - N_{LGM}$) (Fig. 4-6B; Fig. S4-7F). However, we did not observe a significant correlation between allelic richness

and the niche stability since the LGM (N_{STA} = 1 -|N_{CUR} - N_{LGM}|) (Figs. S4-7G and S4-7H).

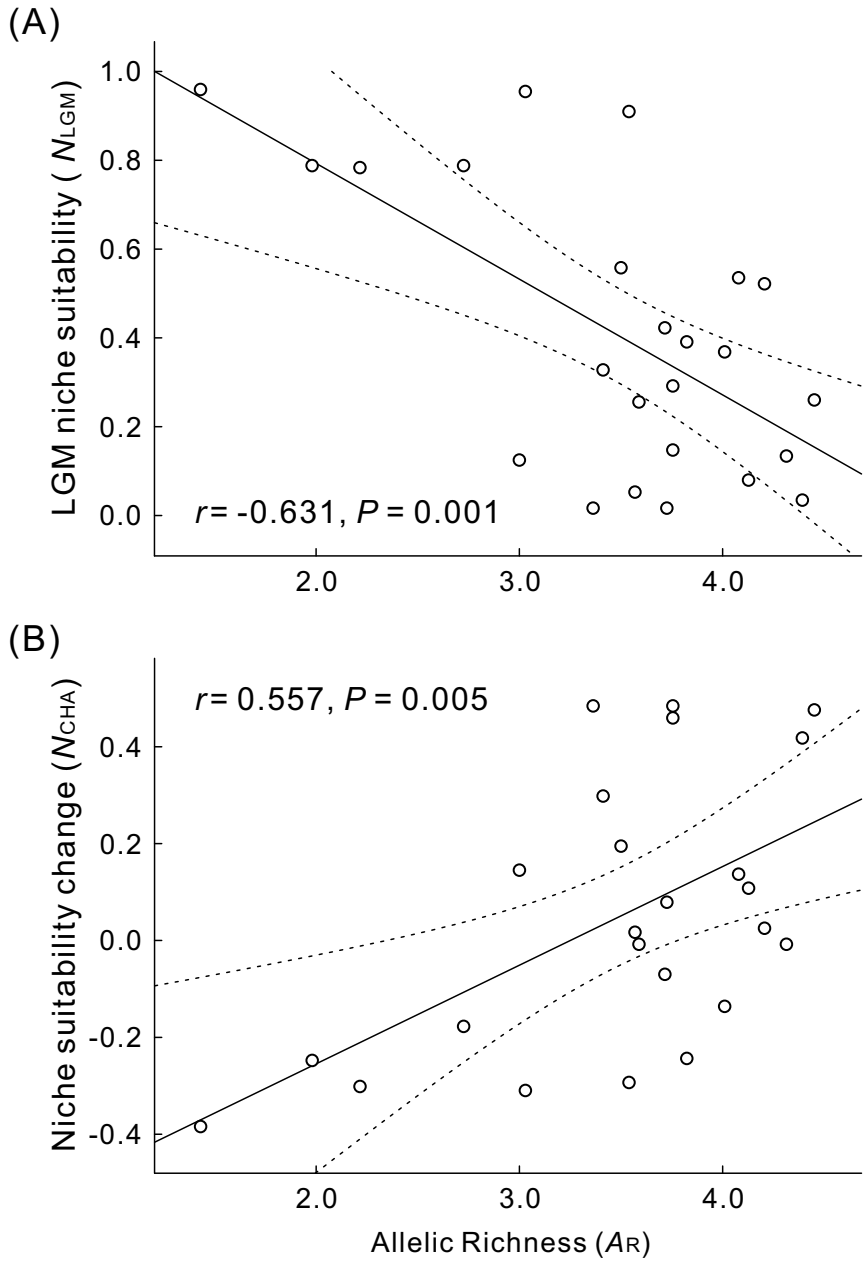


Fig. 4-6 Relationships between allelic richness (A_R) and niche suitability during the Last Glacial Maximum (LGM) based on Community Climate System Model (CCSM) (A) and the change of niche suitability after the LGM (B).

Discussion

Genetic structure and differentiation

In the earlier work on analyses of genetic diversity in cpDNA and rDNA sequences of the same populations of *H. rhamnoides* ssp. *sisnensis* as in our study, both datasets of gene sequences showed no explicit spatial pattern of genetic variation across the range of this taxon. This absence of genetic structure was reflected in the star-like haplotype network characterized by a predominant haplotype (chlorotype C01 and ribotype R01, respectively) and accompanied by a number of low frequency peripheral haplotypes differing by one or two mutational steps from the main haplotype (see Table S2-1 and Fig. S2-1 in Chapter II; Jia *et al.* 2012). In contrast, the present study employing SSRs with higher mutation rates than non-repetitive DNA (Schlötterer 2000) detected a clear-cut differentiation among the sampled populations. Both frequency-based Bayesian clustering STRUCTURE and distance-based PCoA algorithms suggested the presence of two distinct genetic groups that were respectively distributed in the western and eastern regions of the range (Fig. 4-2; Fig. S4-1). Although the population NJ tree based on population D_A distances behaved a little differently, where the three populations in the central portion showing extremely low genetic diversity (populations 11–13) were clustered into an additional group with long branch lengths, it clearly supported the split between the western and eastern populations (Fig. S4-2). However, hierarchical AMOVA attributed a relatively low percentage of the total genetic variation (6.18%) to differences among the two population groups (Table S4-4), indicating a weak genetic divergence

between the groups. Additionally, for each of the two groups, a smaller part of total genetic variation was allocated among populations, with relatively higher differentiation observed within the western group. This pattern of relatively homogenous populations in the eastern part and more heterogeneous populations in the western part was further illustrated by the two analyses based on Delaunay triangulations and Monmonier's maximum-difference algorithm (Fig. 4-3). Nevertheless, comparable levels of genetic diversity were found in the two groups (one-way ANOVA: $F(1, 22) = 0.635$, $P = 0.434$). The differences of geographical and topographical features between the two regions could account for this situation. Subspecies *sinensis* could originate in the QTP region, and colonize the Loess Plateau and then the northern China plain during the Last Interglacial Period, as suggested by Jia *et al.* (2012) and our ecological niche modeling (Fig. 4). During the eastward migration, low genetic diversity and high differentiation would be expected in the newly colonized areas due to sequential genetic drift and founder effects (Hewitt 2000; Hewitt 2004; Petit *et al.* 2003). However, several factors could contribute to violation of this expectation from then on. First, the Loess Plateau and the northern China are topographically plainer compared to the abundant high mountains and deep valleys in the QTP region. Second, life-history traits of seabuckthorn, e.g. outcrossing mating system, long-lived perennial, wind pollination and bird-mediated seed dispersal (Gams 1943; Rousi 1971), could favor more effective and extensive gene dispersal amongst the newly established eastern populations (Austerlitz *et al.* 2000; Duminil *et al.* 2007; Hamrick & Godt 1996). Subsequent genetic homogenization could both increase the intra-population genetic diversity and reduce the inter-population genetic differentiation, so that the genetic (founder effects) signature of the colonization

could gradually be erased (Austerlitz & Garnier-Géré 2003; Pannell & Dorken 2006).

An alternative hypothesis would be that each of the two groups of populations derived its current distribution from a separate origin and represents a different remnant of an ancient centre of diversity. In this scenario, the eastern group of ssp. *sinensis* could survive the LGM in multiple glacial refugia that were maintained across its current range in northern China (Tian *et al.* 2009), or in southerly located refugia outside its current range in northern China (Chen *et al.* 2008a). Consequently, postglacial expansion and admixture of divergent lineages originating from different refugia could result in the comparable diversity within the eastern group (Petit *et al.* 2003). Nevertheless, the ENMs for ssp. *sinensis* did not suggest a substantial contraction or fragmentation in range during the LGM, but depicted instead a continuous distribution of suitable habitat (Fig. 4-4B; Fig. S4-5). These results and biological characteristics of the plant suggest it is likely that continuous dispersal and gene flow between newly established populations and source populations was facilitated by habitat connectivity during the LGM. Under this scenario, genetic diversity could be fairly evenly distributed across the subspecies range after its expansion in the Late Pleistocene.

Isolation by distance and isolation by environment

Spatial genetic structure and divergence among populations can result from isolation by distance, in which landscape barriers and geographical distances cause restricted gene flow (e.g. Alexander Pyron & Burbrink 2009; Avise 2000; Jenkins *et al.* 2010; Wright 1943), and isolation by environment, in which gene flow among populations experiencing

heterogeneous environmental conditions is limited by natural selection against non-locally adapted dispersers and/or individual preference to remain in a particular environment (e.g. Dobzhansky & Dobzhansky 1937; Endler 1986; Freeland *et al.* 2010; Ortego *et al.* 2012; Pease *et al.* 2009; Sacks *et al.* 2004; Sork *et al.* 2010; Wang *et al.* 2013). Testing for IBD across the whole range did not show a significant relationship, which was mainly influenced by the three significantly different central populations 11–13. However, we found a weak but significant IBD pattern after excluding them (Fig. S4-3). On the other hand, our findings indicated that environmental factors appeared to be the major drivers of the west/east split. During the LGM, ssp. *sinensis* occupied a continuous distribution of suitable habitat, with the areas of the highest suitability concentrated in the central part (Fig. 4-4B; Fig. S4-5). Although its current habitat remained continuous, the most suitable areas turned to be fragmented into two parts corresponding to geographic distribution of the two genetic groups (Fig. 4-4A). The area of less suitable habitats between the western and eastern regions is coinciding with deeper interpopulation genetic divergences and acting as a main barrier to gene flow between the two regions of higher within-population diversity (Fig. 4-3). A clear niche structuring revealed by the PCA analysis further suggested that the two genetic groups are currently distributed in distinct environments (Fig. 4-5). This finding was further supported by the MANOVA analysis with the first four PC scores and individual ANOVAs on the 19 bioclimatic variables (Table S4-6). Based on the factor loadings (Fig. S4-6 and Table S4-5), significant separation in the ecological niche between the western and eastern groups was largely driven by the first component axis, which represented a gradient from wet, annually variable and low isothermal environmental conditions (low values) to dry,

annually stable and high isothermal environmental conditions (high values). That said, ssp. *sinensis* could have adapted to environments in the Loess Plateau and the northern China plain, which are wetter, more annually variable, but with lower isothermality, after colonising these regions.

Correlation between Genetic diversity and climatic changes

In addition to population genetic structure, environmental factors can also provide an important impact on the current pattern of genetic diversity. In our study high levels of genetic diversity were observed in all populations except for the three central populations (Table 4-1). We found that among the individual bioclimatic variables, the genetic diversity was negatively correlated with annual mean temperature and minimum temperature of coldest month, though the relationships were only marginally significant (Figs. S4-7A and S4-7B). We then analysed the relationship between genetic diversity and local niche suitability since the LGM, and found that genetic diversity was negatively associated with past habitat suitability but not with current one. Further, the relationship between genetic diversity and the change of niche suitability after the LGM was also tested, and it turned to be positively significant. This pattern indicates that genetic diversity decreased with the deteriorating, and increased with the improving environments. We stress here that "deteriorating" and "improving" environments were determined from postglacial trends in niche suitability for populations of ssp. *sinensis*, as indicated by niche modeling based on current distribution of this taxon and reconstructions of LGM environments for corresponding areas. Fast environmental

deterioration leads to the loss of population density and decrease of population size, which in turn can lead to loss of genetic diversity and alter of the genetic composition through genetic drift and inbreeding (Young *et al.* 1996). Ortego *et al.* (2012) found that genetic diversity of Engelmann oak (*Quercus engelmannii*) is negatively associated with local climatic stability since the LGM, but not with current or past habitat suitability. However, Gugger *et al.* (2013) revealed a positive correlation between genetic diversity of valley oak (*Quercus lobata*) with niche stability from the LGM to present and demonstrated for the first time that LGM climate is associated with present genetic variation. Our present study on ssp. *sinensis* thus provided new insights into the role of both historic climate and climatic changes in shaping present population genetic structure and diversity.

Conclusions

Unlike subspecies of *H. rhamnoides* from the western part of its range (Europe), ssp. *sinensis* experienced after the Last Glacial Maximum (LGM) not a range expansion, but rather a range fragmentation. The most considerable decline in niche suitability was indicated in the middle of the range of this subspecies, across Loess Plateau. Populations from this region also represented the lowest levels of both within-population genetic diversity and gene flow among populations. These results suggest that the Late Quaternary climatic changes could trigger both a moderate decline in niche suitability of ssp. *sinensis* in the middle and also a moderate increase at both edges of its range. Simultaneous climatic changes at the eastern and western edges of the range of ssp. *sinensis* could result in its expansion into less suitable regions at the LGM. Thus,

negative and positive trends in niche suitability have probably caused habitat fragmentation and expansion, respectively, in different parts of the range of this taxon. These changes in population structure and demography were likely associated with correspondent changes in patterns of genetic drift and gene flow, and probably resulted in correspondent decrease and increase of genetic diversity.

References

Alexander Pyron R, Burbrink FT (2009) Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. *Molecular Ecology* **18**, 3443-3457.

Amos W, Hoffman JI, Frodsham A*, et al.* (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes* **7**, 10-14.

Araújo MB, Nogués-Bravo D, Diniz-Filho JAF*, et al.* (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* **31**, 8-15.

Austerlitz F, Garnier-Géré P (2003) Modelling the impact of colonization on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity* **90**, 282-290.

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.

Avise JC (2000) *Phylogeography: the history and formation of species* Harvard University Press.

Bai W-N, Liao W-J, Zhang D-Y (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. *New Phytologist* **188**, 892-901.

Bartish IV, Jeppsson N, Bartish GI, Lu R, Nybom H (2000) Inter- and intraspecific genetic variation in Hippophae (Elaeagnaceae) investigated by RAPD markers. *Plant Systematics and Evolution* **225**, 85-101.

Bartish IV, Swenson U (2003) Elaeagnaceae. In: *The Families and Genera of Vascular Plants (Vol. VI)* (ed. Kubitzki K), pp. 131-134. Springer-Verlag, Berlin, Heidelberg, New York.

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote

deficiency. *Molecular Ecology* **5**, 453-455.

Chapman J, Nakagawa S, Coltman D, Slate J, Sheldon B (2009) A quantitative review of heterozygosity–fitness correlations in animal populations. *Molecular Ecology* **18**, 2746-2765.

Chen K, Abbott RJ, Milne RI, Tian XM, Liu J (2008a) Phylogeography of *Pinus tabulaeformis* Carr.(Pinaceae), a dominant species of coniferous forest in northern China. *Molecular Ecology* **17**, 4276-4288.

Chen S, Wu G, Zhang D*, et al.* (2008b) Potential refugium on the Qinghai-Tibet Plateau revealed by the chloroplast DNA phylogeography of the alpine species *Metagentiana striata* (Gentianaceae). *Botanical Journal of the Linnean Society* **157**, 125-140.

Coombs J, Letcher B, Nislow K (2008) CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. *Molecular Ecology Resources* **8**, 578-580.

Cronk Q (1997) Islands: stability, diversity, conservation. *Biodiversity & Conservation* **6**, 477-493.

Dobzhansky TG, Dobzhansky T (1937) *Genetics and the Origin of Species* Columbia University Press.

Duminil J, Fineschi S, Hampe A*, et al.* (2007) Can Population Genetic Structure Be Predicted from Life‐History Traits? *The American Naturalist* **169**, 662-672.

Edwards DL, Keogh JS, Knowles LL (2012) Effects of vicariant barriers, habitat stability, population isolation and environmental features on species divergence in the south-western Australian coastal reptile community. *Molecular Ecology* **21**, 3809-3822.

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.

Endler JA (1986) *Natural selection in the wild* Princeton University Press.

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Frankham R (2005) Stress and adaptation in conservation genetics. *Journal of evolutionary biology* **18**, 750-755.

Freeland J, Biss P, Conrad K, Silvertown J (2010) Selection pressures have caused genome‐wide population differentiation of Anthoxanthum odoratum despite the potential for high gene flow. *Journal of evolutionary biology* **23**, 776-782.

Gams H (1943) Der Sanddorn (*Hippophae rhamnoides* L.) im Alpengebiet. *Beihefte zum Botanischen Centralblatt, Abteilung B* **2**, 68-96.

Graham CH, Moritz C, Williams SE (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 632-636.

Gugger PF, Ikegami M, Sork VL (2013) Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née. *Molecular Ecology* **22**, 3598-3612.

Hamrick J, Godt M (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **351**, 1291-1298.

Harrison S, Yu G, Takahara H, Prentice I (2001) Palaeovegetation (Communications arising): diversity of temperate plants in east Asia. *Nature* **413**, 129-130.

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

Hewitt GM (2004) The structure of biodiversity–insights from molecular phylogeography. *Frontiers in Zoology* **1**, 1-16.

Hijmans R, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* **127**, 15-19.

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.

Hood G (2010) PopTools version 3.2.5. *Available on the internet. URL http://www.poptools.org*.

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577-586.

IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jain A, Ghangal R, Grover A, Raghuvanshi S, Sharma PC (2010) Development of EST-based new SSR markers in seabuckthorn. *Physiology and Molecular Biology of Plants* **16**, 375-378.

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.

Jansson R (2003) Global patterns in endemism explained by past climatic change. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 583-590.

Jenkins DG, Carey M, Czerniewska J*, et al.* (2010) A meta‐analysis of isolation by distance: relic or reference standard for landscape genetics? *Ecography* **33**, 315-320.

Jia D-R, Abbott RJ, Liu T-L*, et al.* (2012) Out of the Qinghai–Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of Hippophaë rhamnoides (Elaeagnaceae). *New Phytologist* **194**, 1123-1133.

Jia D-R, Liu T-L, Wang L-Y, Zhou D-W, Liu J-Q (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* **102**, 37-50.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187-189.

Li TS, Schroeder W (1996) Sea buckthorn (*Hippophae rhamnoides* L.): a multipurpose plant. *HortTechnology* **6**, 370-380.

Lian YS, Chen XL, Sun K, Ma RJ (2003) A new subspecies of *Hippophae* (Elaeagnaceae) from China. *Novon* **13**, 200-202.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385-393.

Liu J-Q, Sun Y-S, Ge X-J, Gao L-M, Qiu Y-X (2012) Phylogeographic studies of plants in China: Advances in the past and directions in the future. *Journal of Systematics and Evolution* **50**, 267-275.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011a) Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**, 1267-1280.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011b) Mountains of Southern China as "Plant Museums" and "Plant Cradles": Evolutionary and Conservation Insights. *Mountain Research and Development* **31**, 261-269.

Lu R (1992) Seabuckthorn: A multipurpose plant species for fragile mountains. International Center for Integrated Mountain Development, Occasional Paper No. 20, Kathmandu, Nepal.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**, 173-190.

Markert JA, Schelly RC, Stiassny ML (2010) Genetic isolation and morphological divergence mediated by high-energy rapids in two cichlid genera from the lower Congo rapids. *BMC evolutionary biology* **10**, 149.

Meng LH, Yang R, Abbott RJ*, et al.* (2007) Mitochondrial and chloroplast phylogeography of *Picea crassifolia* Kom. (Pinaceae) in the Qinghai-Tibetan Plateau and adjacent highlands. *Molecular Ecology* **16**, 4128-4137.

Miller M (2005) Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* **96**, 722-724.

Monmonier MS (1973) Maximum‐Difference Barriers: An Alternative Numerical Regionalization

Method*. *Geographical Analysis* **5**, 245-261.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Ni J, Harrison SP, Colin Prentice I, Kutzbach JE, Sitch S (2006) Impact of climate variability on present and Holocene vegetation: a model-based study. *Ecological modelling* **191**, 469-486.

Nowak C, Jost D, Vogt C*, et al.* (2007) Consequences of inbreeding and reduced genetic variation on tolerance to cadmium stress in the midge *Chironomus riparius*. *Aquatic Toxicology* **85**, 278-284.

Ortego J, Riordan EC, Gugger PF, Sork VL (2012) Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology* **21**, 3210-3223.

Pannell JR, Dorken ME (2006) Colonization as a common denominator in plant metapopulations and range expansions: effects on genetic diversity and sexual systems. *Landscape Ecology* **21**, 837-848.

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology* **22**, 925-946.

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

Pease KM, Freedman AH, Pollinger JP*, et al.* (2009) Landscape genetics of California mule deer (Odocoileus hemionus): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology* **18**, 1848-1862.

Petit RJ, Aguinagalde I, de Beaulieu J-L*, et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.

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.

Qiu Y-X, Fu C-X, Comes HP (2011) Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* **59**, 225-244.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rice WR (1989) ANALYZING TABLES OF STATISTICAL TESTS. *Evolution* **43**, 223-225.

Rosenberg NA (2003) DISTRUCT: a program for the graphical display of population structure.

Molecular Ecology Notes **4**, 137-138.

Rousi A (1965) Observations on the cytology and variation of European and Asiatic populations of *Hippophae rhamnoides*. *Annales Botanici Fennici* **2**, 1-18.

Rousi A (1971) The genus *Hippophae* L. A taxonomic study. *Annales Botanici Fennici* **8**, 177-227.

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

Sacks BN, Brown SK, Ernest HB (2004) Population structure of California coyotes corresponds to habitat‐specific breaks and illuminates species history. *Molecular Ecology* **13**, 1265-1275.

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**, 365-371.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature biotechnology* **18**, 233-234.

Sork VL, Davis FW, Westfall R*, et al.* (2010) Gene movement and genetic association with regional climate gradients in California valley oak (Quercus lobata Nee) in the face of climate change. *Molecular Ecology* **19**, 3806-3823.

Sun K, Chen W, Ma R*, et al.* (2006) Genetic Variation in *Hippophae rhamnoides* ssp. *sinensis* (Elaeagnaceae) Revealed by RAPD Markers. *Biochemical Genetics* **44**, 186-197.

Swenson U, Bartish IV (2002) Taxonomic synopsis of *Hippophae* (Elaeagnaceae). *Nordic Journal of Botany* **22**, 369-374.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389-399.

Takezaki N, Nei M, Tamura K (2010) POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular biology and evolution* **27**, 747-752.

Temunović M, Frascaria-Lacoste N, Franjić J, Satovic Z, Fernández-Manjarrés JF (2013) Identifying refugia from climate change using coupled ecological and genetic data in a transitional Mediterranean-temperate tree species. *Molecular Ecology* **22**, 2128-2142.

Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8245-8250.

Tian B, Liu R, Wang L*, et al.* (2009) Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. *Journal of Biogeography* **36**, 2148-2155.

Tian C, He X, Zhong Y, Chen J (2002) Effects of VA mycorrhizae and *Frankia* dual inoculation on growth and nitrogen fixation of *Hippophae tibetana*. *Forest ecology and management* **170**, 307-312.

Tzvelev NN (2002) On the genera *Elaeagnus* and *Hippophae* (Elaeagnaceae) in Russia and adjacent states. *Botanical journal (Ботанический журнал)* **87**, 74-86.

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.

Wang A, Zhang Q, Wan D, Liu J (2008) Nine microsatellite DNA primers for Hippophae rhamnoides ssp. sinensis (Elaeagnaceae). *Conservation Genetics* **9**, 969-971.

Wang IJ, Glor RE, Losos JB (2013) Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology letters* **16**, 175-182.

Wang L, Abbott RJ, Zheng WEI*, et al.* (2009) History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnandrum* (Ranunculaceae). *Molecular Ecology* **18**, 709-721.

Weaver AJ, Eby M, Fanning AF, Wiebe EC (1998) Simulated influence of carbon dioxide, orbital forcing and ice sheets on the climate of the Last Glacial Maximum. *Nature* **394**, 847-853.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114.

Wu Z (1988) Hengduan mountain flora and her significance. *Journal of Japanese Botany* **63**, 297-311.

Yang F-S, Li Y-F, Ding XIN, Wang X-Q (2008) Extensive population expansion of *Pedicularis longiflora* (Orobanchaceae) on the Qinghai-Tibetan Plateau and its correlation with the Quaternary climate change. *Molecular Ecology* **17**, 5135-5145.

Yao Y, Tigerstedt PA (1993) Isozyme studies of genetic diversity and evolution in *Hippophae*. *Genetic Resources and Crop Evolution* **40**, 153-164.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* **11**, 413-418.

Yu G, Chen X, Ni J*, et al.* (2000) Palaeovegetation of China: a pollen data‐based synthesis for the mid‐Holocene and last glacial maximum. *Journal of Biogeography* **27**, 635-664.

Zeng Y-F, Liao W-J, Petit RJ, Zhang D-Y (2011) Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Molecular Ecology* **20**, 4995-5011.

Zhang F-Q, Gao Q-B, Zhang D-J*, et al.* (2012) Phylogeography of *Spiraea alpina* (Rosaceae) in the Qinghai-Tibetan Plateau inferred from chloroplast DNA sequence variations. *Journal of Systematics and Evolution* **50**, 276-283.

Zhang Q, Chiang TY, George M, Liu JQ, Abbott RJ (2005) Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. *Molecular Ecology* **14**, 3513-3524.

Zhou S, Wang X, Wang J, Xu L (2006) A preliminary study on timing of the oldest Pleistocene glaciation in Qinghai–Tibetan Plateau. *Quaternary International* **154**, 44-51.

Supporting information

Table S4-1 Characteristics of the 15 microsatellite loci screened in all 24 populations of *Hippophae rhamnoides* ssp. *sinensis*. *A*, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity

Locus	*A*	H_O	H_E	Size range (bp)	Repeat motif	Primer origin	Modified primer sequence
Hr01	13	0.773	0.730	105–129	$(GA)_n$	Wang *et al.* 2008	R: 5'–GCCGTTCAAAGTGTCAAA–3'
Hr02	6	0.360	0.452	87–97	$(AG)_n$	Wang *et al.* 2008	R: 5'–ATTTTTCCCCAACTGCGG–3'
Hr03	12	0.467	0.617	80–106	$(AG)_n$	Wang *et al.* 2008	R: 5'–TCCCCAACTGCGGACAAC–3'
Hr06	11	0.749	0.731	76–96	$(CA)_n$	Wang *et al.* 2008	
Hrms003	7	0.323	0.531	333–351	$(TCA)_n$	Jain *et al.* 2010	
Hrms004	9	0.203	0.560	283–307	$(TC)_n$	Jain *et al.* 2010	
Hrms010	4	0.132	0.126	223–232	$(TGG)_n$	Jain *et al.* 2010	
Hrms012	13	0.672	0.637	139–175	$(CTT)_n$	Jain *et al.* 2010	
Hrms014	17	0.090	0.395	218–254	$(TG)_n(TA)_n$	Jain *et al.* 2010	
Hrms018	7	0.338	0.606	188–206	$(ATG)_n$	Jain *et al.* 2010	
Hrms021	2	0.002	0.007	182–185	$(GAA)_n$	Jain *et al.* 2010	
Hrms023	4	0.394	0.446	226–241	$(AAAAT)_n$	Jain *et al.* 2010	
Hrms025	24	0.398	0.706	310–364	$(AG)_n$	Jain *et al.* 2010	
Hrms026	16	0.076	0.603	209–257	$(CAC)_n$	Jain *et al.* 2010	
Hrms028	2	0.031	0.029	92–95	$(TTG)_n$	Jain *et al.* 2010	

Table S4-2 Percentage of failed amplifications across loci and populations of *Hippophae rhamnoides* ssp. *sinensis*

Pop	Hrms003	Hrms004	Hrms014	Hrms026
02	7.69	7.69	23.08	
03	11.11			
04	11.11			
05	8.33			
06	5.88			
07		29.41		
08	6.25			
09	11.76			
10	27.27	9.09		9.09
12	44.44			
22	5.56			
Total	5.33	1.87	0.80	0.27

Table S4-3 Indicated existence and frequencies of null alleles across 15 microsatellite loci and 24 populations for *Hippophae rhamnoides* ssp. *sinensis*, as suggested by the program MICRO-CHECKER

Pop	Hr01	Hr02	Hr03	Hr06	Hrms003	Hrms004	Hrms010	Hrms012	Hrms014	Hrms018	Hrms021	Hrms023	Hrms025	Hrms026	Hrms028
01	0.134	0.210	0.138	0.247	0.148	0.279			0.261	0.249			0.376	0.457	
02		0.123		0.186	0.426	0.302			0.552	0.119		0.207		0.411	
03		0.166	0.259		0.455				0.331	0.191		0.182	0.268	0.350	
04			0.170			0.247							0.464	0.294	
05	0.180	0.063	0.172			0.302			0.308	0.188		0.111		0.327	
06			0.133	0.167	0.380	0.313			0.305	0.181			0.262	0.255	
07		0.124	0.136		0.307	0.666			0.230					0.287	
08					0.346	0.189			0.321	0.210		0.162	0.368	0.400	
09					0.440	0.317			0.129	0.262			0.347	0.402	
10						0.417		0.193	0.361			0.243	0.181	0.405	
11												0.099	0.225		
12					0.699	0.333				0.256				0.319	
13						0.333									
14		0.132	0.137			0.190			0.158	0.257		0.114	0.149	0.385	
15			0.234	0.154		0.255			0.172	0.199			0.276	0.297	
16			0.151		0.200	0.152	0.095		0.095	0.377		0.283		0.308	
17			0.114		0.147	0.180				0.296				0.370	
18			0.176			0.234	0.105	0.144	0.337	0.337	0.105			0.363	
19		0.205			0.206	0.172			0.172	0.306			0.131	0.321	
20					0.111	0.192			0.342	0.176			0.164	0.353	
21					0.224	0.237			0.211	0.144			0.287	0.384	
22			0.190		0.242	0.230			0.202	0.128		0.178	0.133	0.348	
23			0.253		0.181	0.199			0.130	0.205				0.363	
24		0.205			0.168	0.245			0.417	0.125			0.147	0.331	

Table S4-4 Analysis of molecular variance (AMOVA) of microsatellite loci for the populations of *Hippophae rhamnoides* ssp. *sinensis*

Source of variation	d.f.	SS	VC	PV (%)	Fixation Indices
Group West					
Among populations	10	280.458	0.901	19.35	F_{ST} = 0.193*
Within populations	287	1078.529	3.758	80.65	
Total	297	1358.987	4.659		
Group East					
Among populations	12	275.778	0.555	13.06	F_{ST} = 0.131*
Within populations	439	1621.671	3.694	86.94	
Total	451	1897.449	4.249		
All populations					
Among populations	23	685.754	0.836	18.36	F_{ST} = 0.184*
Within populations	726	2700.199	3.719	81.64	
Total	749	3385.953	4.556		
Group West vs Group East					
Among groups	1	129.518	0.290	6.18	F_{ST} = 0.209*
Among populations within groups	22	556.236	0.691	14.7	F_{SC} = 0.157*
Within populations	726	2700.199	3.719	79.12	F_{CT} = 0.062*
Total	749	3385.953	4.701		
Group West excluding Pop 13					
Among populations	9	163.284	0.537	11.74	F_{ST} = 0.117*
Within populations	254	1026.588	4.042	88.26	
Total	263	1189.871	4.579		
Group East excluding Pops 11–12					
Among populations	10	150.010	0.320	7.64	F_{ST} = 0.076*
Within populations	371	1437.980	3.876	92.36	
Total	381	1587.990	4.196		
All populations excluding Pops 11–13					
Among populations	20	434.021	0.578	12.79	F_{ST} = 0.128*
Within populations	625	2464.567	3.943	87.21	
Total	645	2898.588	4.522		
Group West excluding Pop 13 vs. Group East excluding Pops 11–12					
Among groups	1	120.727	0.334	7.12	F_{ST} = 0.158*
Among populations within groups	19	313.294	0.409	8.72	F_{SC} = 0.094*
Within populations	625	2464.567	3.943	84.15	F_{CT} = 0.071*
Total	645	2898.588	4.686		

* $P < 0.001$ (1023 permutations).

SS, sum of squares; VC, variance components; PV, percentage of variation; F_{CT}, genetic variation between groups; F_{SC}, genetic variation among populations within groups; F_{ST}, genetic variation among all populations.

Table S4-5 Results of Principal Component analysis (PCA) on the 19 bioclimatic variables

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17
Eigenvalue	8.804	4.673	3.397	1.579	0.337	0.158	0.022	0.014	0.007	0.004	0.003	0.002	0.001	0.000	0.000	0.000	0.000
% variance explained	46.335	24.597	17.877	8.308	1.773	0.830	0.114	0.073	0.039	0.021	0.016	0.010	0.004	0.002	0.001	0.000	0.000
Bio1	-0.037	-0.886	-0.439	-0.136	-0.026	0.017	0.003	-0.031	0.003	-0.001	0.014	0.001	-0.013	-0.001	0.004	-0.001	-0.001
Bio2	0.273	0.206	0.265	-0.868	-0.240	0.002	-0.003	-0.017	-0.017	-0.008	-0.005	-0.019	0.002	-0.001	0.002	0.000	0.000
Bio3	0.720	-0.178	0.427	-0.507	-0.052	-0.046	0.006	0.072	0.006	0.002	0.008	0.019	-0.005	0.004	-0.001	0.000	0.000
Bio4	-0.824	0.390	-0.405	0.045	-0.042	0.015	-0.005	0.029	0.000	-0.006	-0.001	-0.005	0.001	0.012	0.003	0.002	-0.001
Bio5	-0.555	-0.431	-0.676	-0.158	-0.152	0.020	-0.010	0.011	0.029	-0.019	-0.008	0.008	0.010	-0.006	-0.002	0.000	0.000
Bio6	0.489	-0.844	-0.146	0.109	0.062	-0.097	-0.029	0.033	0.011	-0.022	-0.010	-0.011	0.005	-0.002	0.001	0.000	0.000
Bio7	-0.791	0.444	-0.322	-0.199	-0.154	0.097	0.018	-0.021	0.010	0.006	0.003	0.015	0.002	-0.002	-0.002	0.000	0.000
Bio8	-0.683	-0.385	-0.602	-0.116	0.045	0.030	0.000	0.048	-0.044	0.030	-0.005	-0.007	0.000	-0.006	-0.002	0.000	0.000
Bio9	0.547	-0.823	-0.019	-0.134	0.030	-0.026	0.033	-0.030	-0.011	0.019	-0.006	0.007	0.018	0.008	0.001	0.000	0.000
Bio10	-0.610	-0.453	-0.645	-0.048	-0.056	0.019	-0.003	0.001	0.000	-0.006	0.002	0.000	-0.006	0.009	0.001	0.000	0.002
Bio11	0.526	-0.842	-0.027	-0.111	0.016	-0.010	0.000	-0.029	0.002	0.005	0.006	0.002	-0.007	-0.001	-0.004	0.003	0.000
Bio12	-0.630	-0.472	0.585	-0.079	0.036	0.162	-0.018	0.017	0.045	0.021	0.015	-0.014	0.004	0.001	0.000	0.000	0.000
Bio13	-0.914	-0.139	0.264	-0.197	0.179	0.007	-0.051	-0.013	-0.032	-0.023	0.027	0.003	0.006	0.002	-0.003	0.000	0.000
Bio14	-0.877	-0.258	0.295	0.186	-0.107	-0.138	0.105	0.012	-0.002	-0.009	0.018	-0.008	0.001	-0.002	-0.001	0.000	0.000
Bio15	-0.485	0.506	-0.433	-0.458	0.232	-0.239	-0.008	-0.014	0.030	0.015	0.002	-0.003	0.000	0.000	0.000	0.000	0.000
Bio16	-0.841	-0.276	0.394	-0.176	0.160	0.057	0.030	-0.008	0.004	-0.010	-0.034	-0.003	-0.008	0.004	-0.005	-0.001	0.000
Bio17	-0.804	-0.265	0.457	0.165	-0.178	-0.115	-0.044	-0.008	-0.003	0.010	-0.008	0.004	-0.002	0.000	0.000	0.000	0.000
Bio18	-0.842	-0.207	0.408	-0.223	0.171	0.040	0.023	0.002	-0.005	-0.007	-0.009	0.014	0.001	-0.006	0.008	0.001	0.000
Bio19	-0.804	-0.265	0.457	0.165	-0.178	-0.115	-0.044	-0.008	-0.003	0.010	-0.008	0.004	-0.002	0.000	0.000	0.000	0.000

Table S4-6 Results of individual ANOVAs of the 19 Bioclimatic variables between the two groups

Variables	SS	MS	F (1, 22)	*P*
Bio1	3.721	3.721	0.004	0.953
Bio2	955.630	955.630	3.795	0.064
Bio3	608.399	608.399	38.181	0.000
Bio4	84446911.189	84446911.189	52.904	0.000
Bio5	9995.469	9995.469	9.656	0.005
Bio6	10343.171	10343.171	5.005	0.036
Bio7	40674.308	40674.308	26.437	0.000
Bio8	12432.117	12432.117	13.384	0.001
Bio9	17973.577	17973.577	11.232	0.003
Bio10	12295.459	12295.459	11.710	0.002
Bio11	16402.975	16402.975	10.077	0.004
Bio12	8006.575	8006.575	0.353	0.558
Bio13	11730.485	11730.485	8.143	0.009
Bio14	51.154	51.154	11.997	0.002
Bio15	1020.000	1020.000	10.209	0.004
Bio16	30089.232	30089.232	3.728	0.067
Bio17	343.647	343.647	6.880	0.016
Bio18	31679.000	31679.000	3.709	0.067
Bio19	343.647	343.647	6.880	0.016

SS, Sum of Squares; MS, Mean Square.

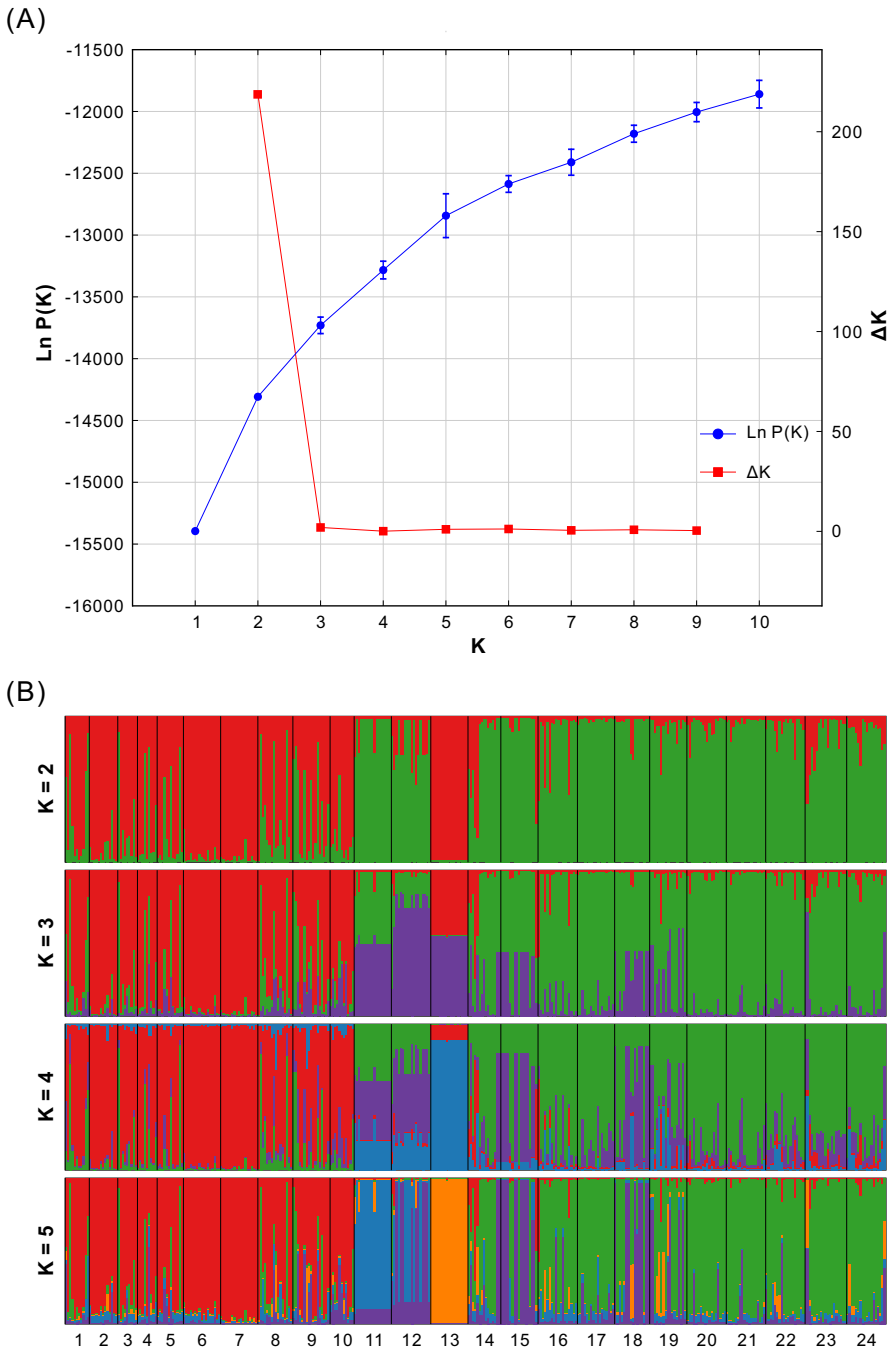


Fig. S4-1 (A) Plots of mean Ln P(K) (± SD) and corresponding ΔK statistics calculated according to Evanno *et al.* (2005) for K values from 1 to 10 over 20 independent runs in STRUCTURE analysis. (B) Histogram of the STRUCTURE analysis for the model with K from 1 to 5. Each individual is represented by a thin vertical line, which is partitioned into K colored segments that represent its estimated population group membership fractions. Black lines separate individuals from different populations, of which codes are shown at the bottom line (as in Table 4-1).

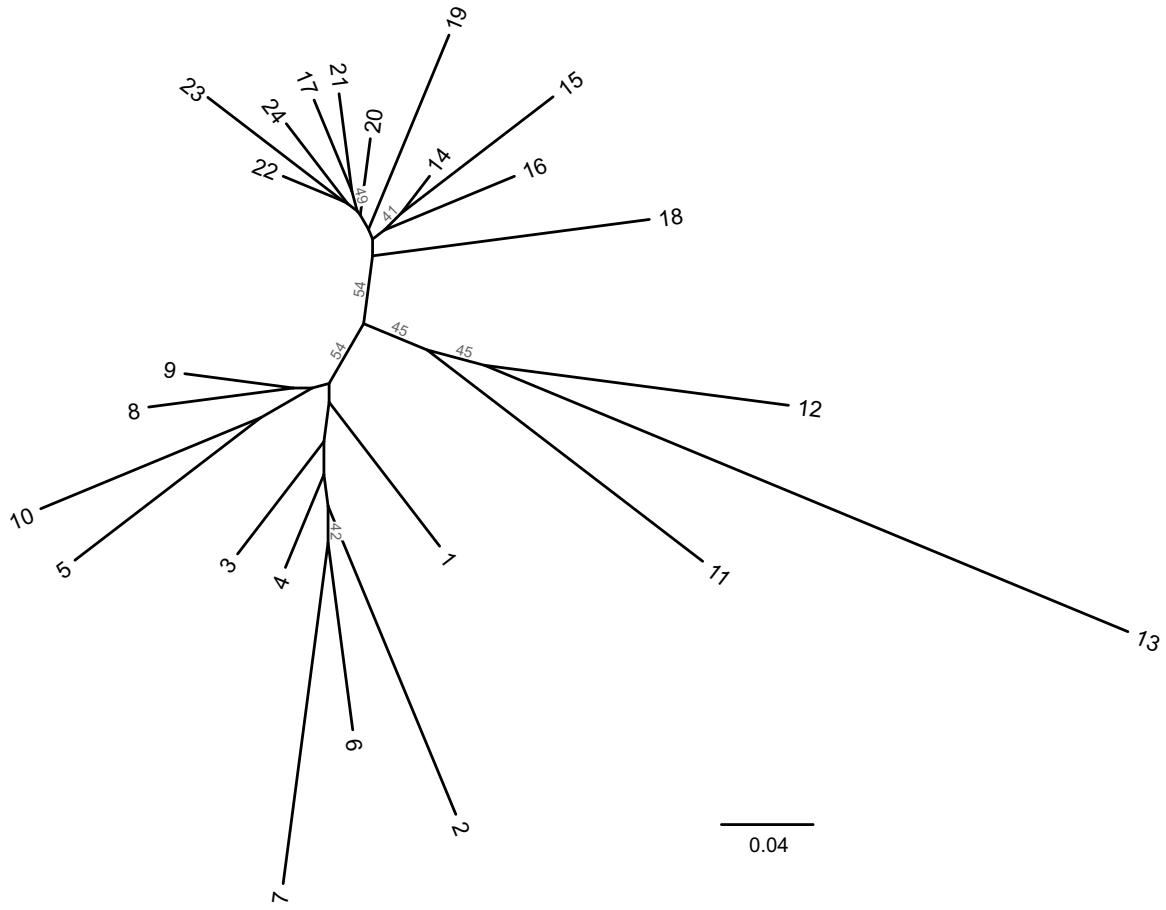


Fig. S4-2 Neighbour-joining (NJ) phylogram of the 24 populations based on Nei's D_A distances (Nei *et al.* 1983). Numbers terminating each branch correspond to population codes (as in Table 4-1). Values reported at nodes are percentages of 1000 bootstrap replicates (greater than 40%) supporting the respective nodes.

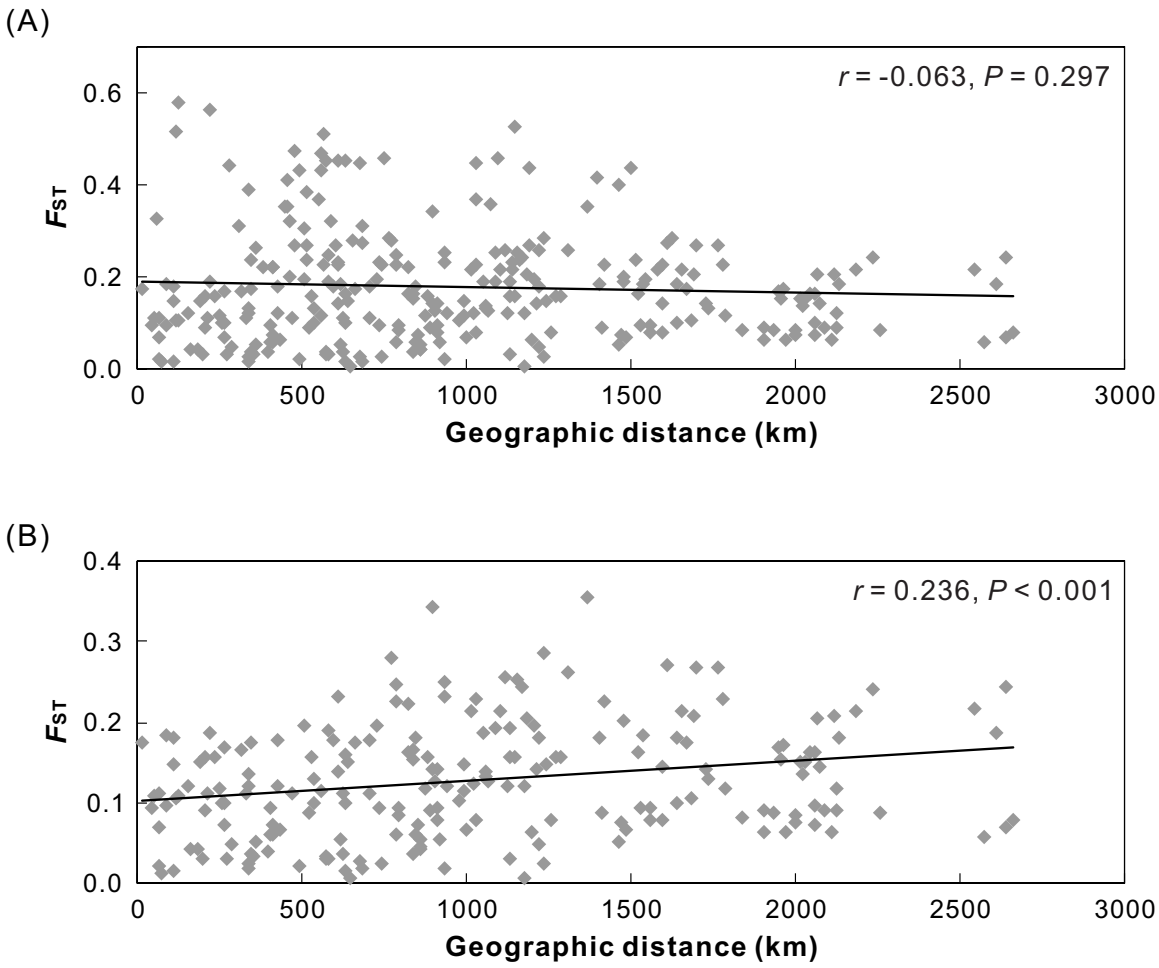


Fig. S4-3 Relationship between genetic differentiation (F_{ST}) and geographic distance (km) for all 24 populations of *Hippophae rhamnoides* ssp. *sinensis* (A) and excluding populations 11–13 (B).

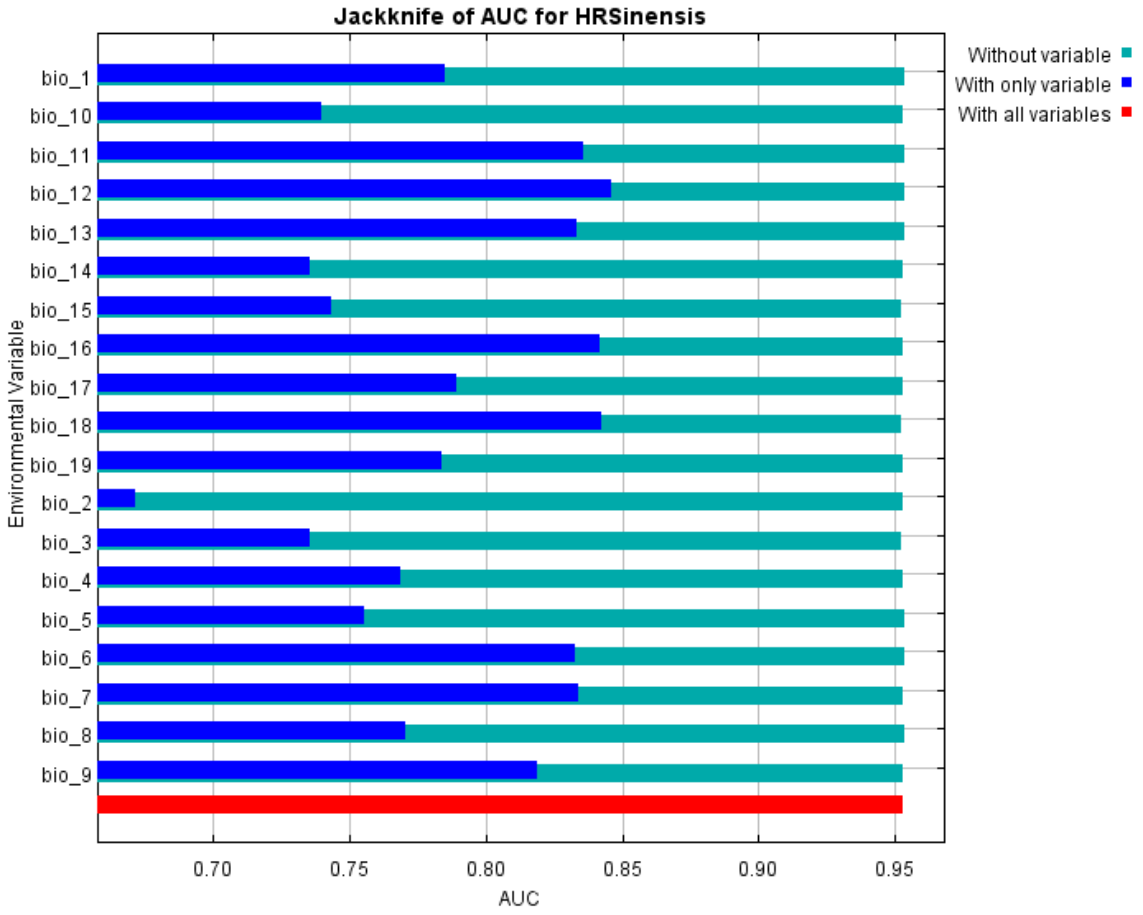


Fig. S4-4 Jackknife tests for environmental variable importance performed by MAXENT.

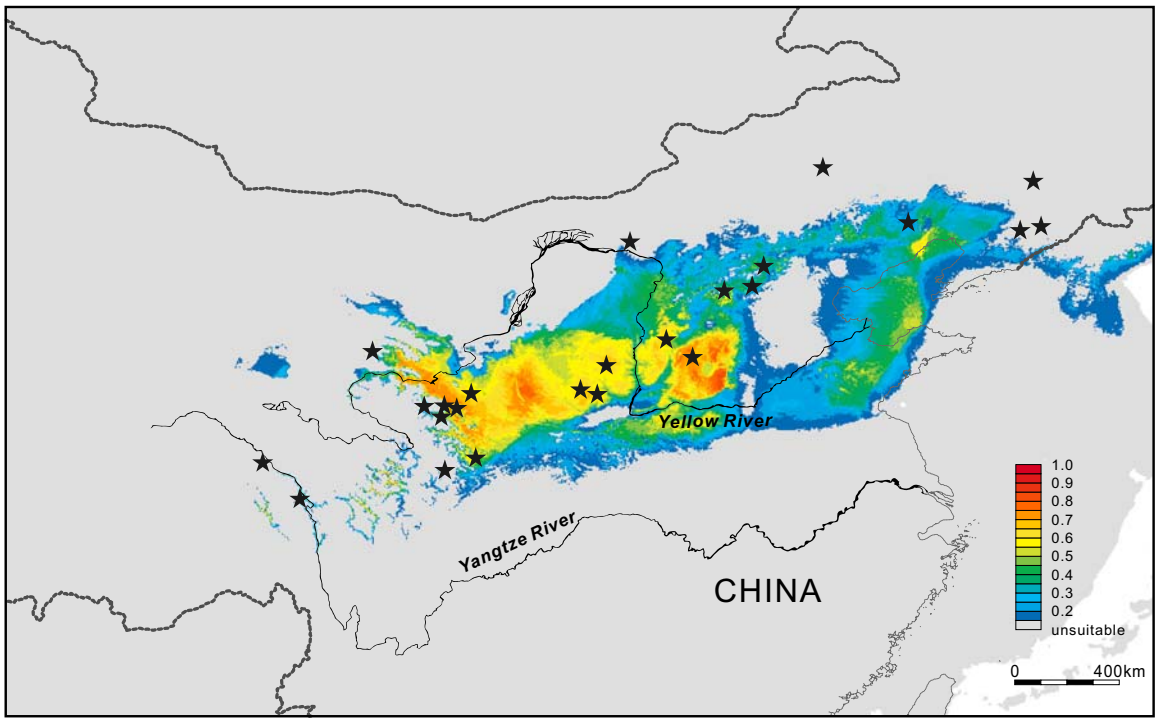


Fig. S4-5 Ecological niche model (ENM) for ssp. *sinensis* estimated from current conditions for the Last Glacial Maximum based on Model for Interdisciplinary Research on Climate (MIROC). Warmer colors show areas with higher suitability. The logistic threshold under the "maximum training sensitivity plus specificity" criterion for presence–absence prediction is 0.130. Black stars mark locations of the 24 sampled populations (as in Fig. 4-1, Fig. 4-4 and Table 4-1).

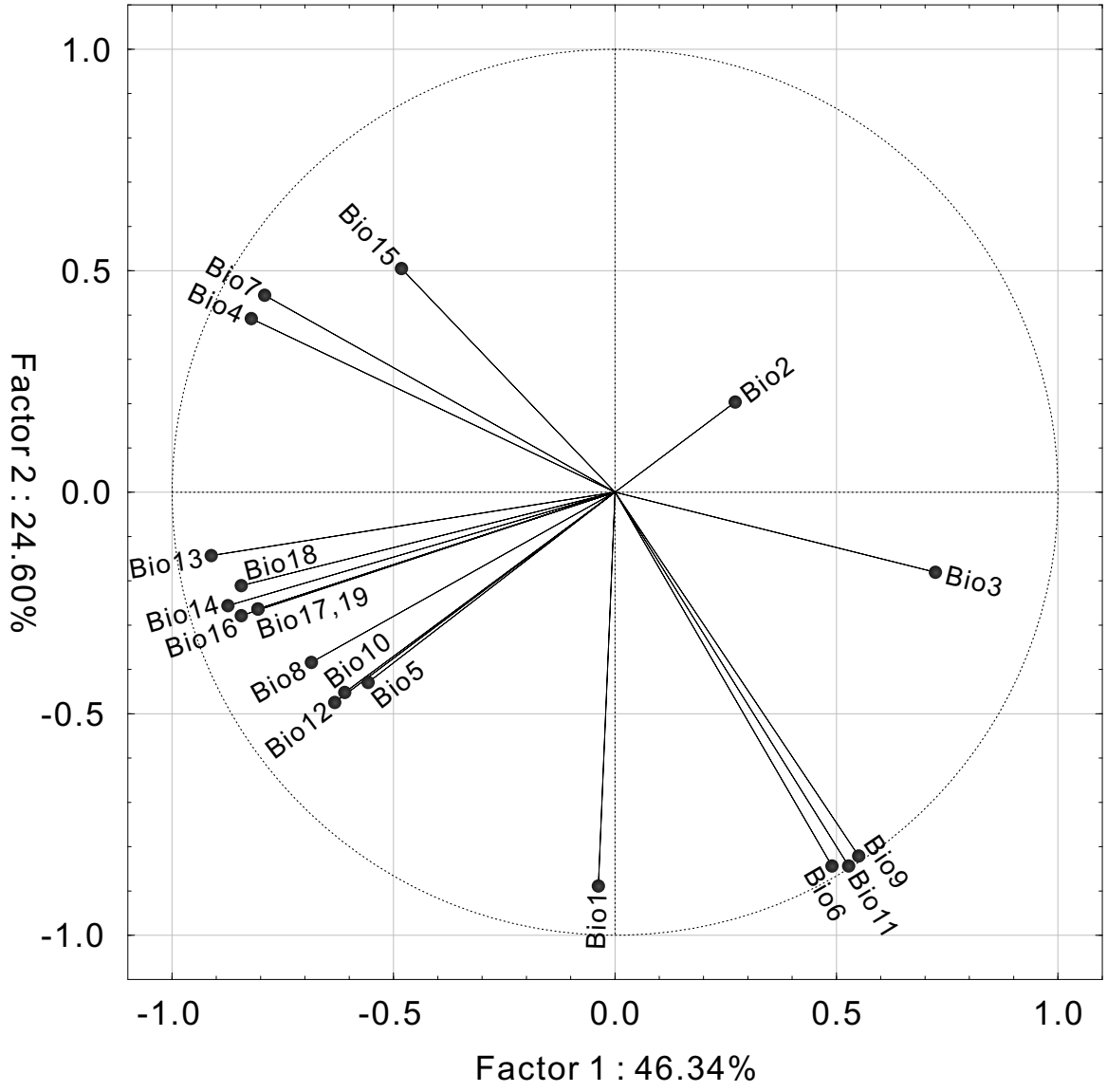


Fig. S4-6 Projection of the factor coordinates of the 19 bioclimatic variables on the 1 x 2 factor plane obtained by the principal component analysis.

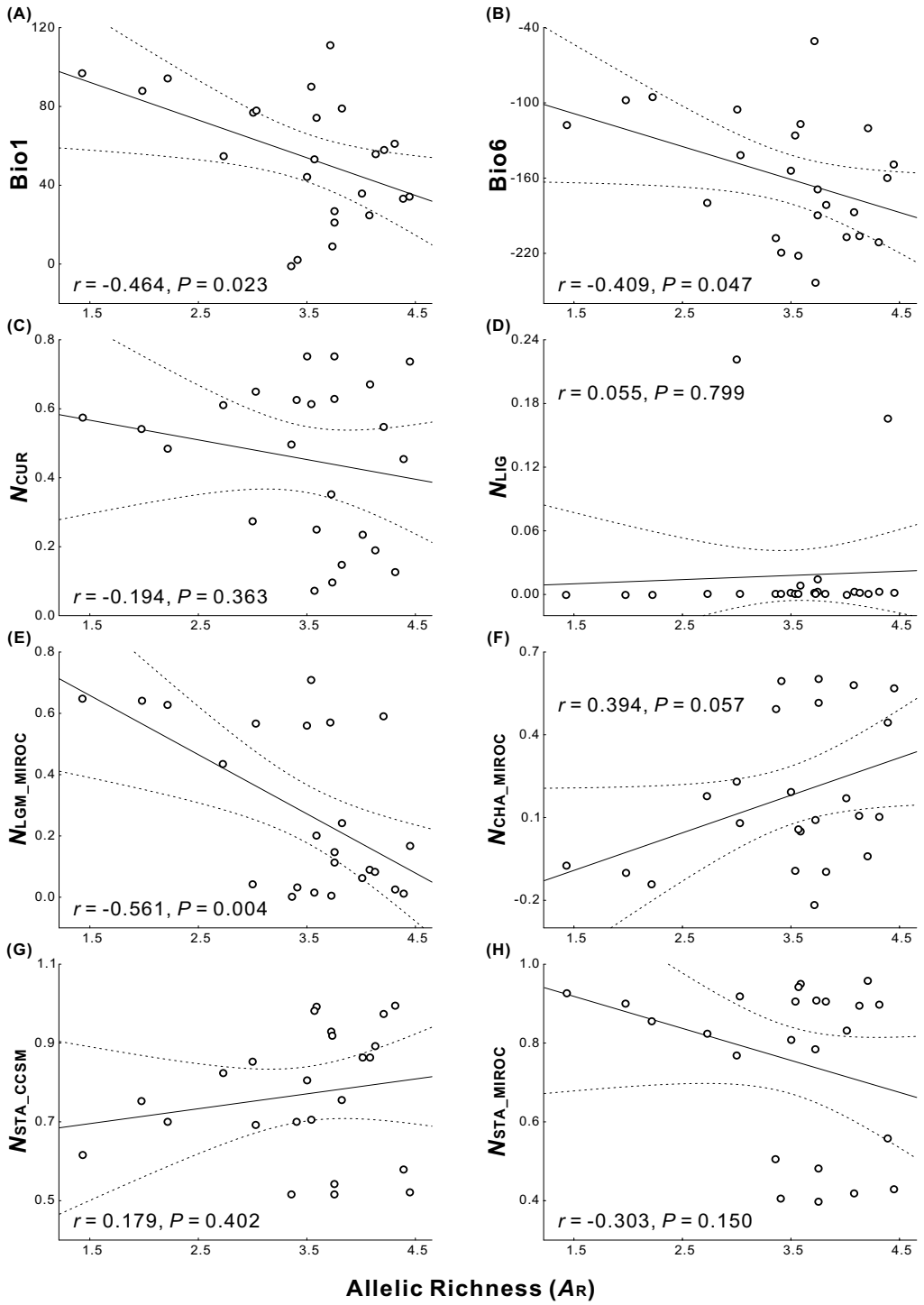


Fig. S4-7 Relationships between allelic richness (A_R) and annual mean temperature (Bio1; A), minimum temperature of coldest month (Bio6; B), current niche suitability (N_{CUR}; C), niche suitability during the Last Interglacial Period (N_{LIG}; D), niche suitability during the Last Glacial Maximum based on Community Climate System Model (N_{LGM_MIROC}; E), niche suitability change after the LGM based on MIROC (N_{CHA_MIROC}; F) and niche stability based on CCSM (G) and MIROC (H).

Appendix List of the 19 bioclimatic variables used to develop the Ecological Niche Models

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (BIO2/BIO7) (* 100)

BIO4 = Temperature Seasonality (standard deviation *100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range (BIO5 - BIO6)

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

BIO19 = Precipitation of Coldest Quarter