Dissertation

submitted to the

Faculty of Bio Sciences of the
Ruperto-Carola University of Heidelberg, Germany

for the degree of
Doctor of Natural Sciences

Diplom-Biologin Roswitha Elisabeth Schmickl

born in Heilbronn-Neckargartach, Germany

Oral examination: 09.11.2009

Reticulate evolution in glacial refuge areas –

the genus *Arabidopsis* in the eastern Austrian Danube Valley (Wachau)

For my father

Referees: Prof. Dr. Marcus Koch

Prof. Dr. Hans-Peter Comes

TABLE OF CONTENTS

Supplementary material (CD) – Table of contents

1. An Amphi-Beringian allopolyploid *Arabidopsis* and the evolutionary history of the *Arabidopsis lyrata* complex

Abstract

Hybridisierung und Polyploidisierung tragen wesentlich zur Artbildung im Pflanzenreich bei. Innerhalb der Gattung *Arabidopsis* ist Hybridisierung nur von *Arabidopsis suecica* aus Fennoskandinavien und *Arabidopsis kamchatica* aus Japan bekannt. Diese Studie befasst sich mit den Artkomplexen von *Arabidopsis lyrata* und *Arabidopsis arenosa*. Unser Ziel war es, herauszufinden, ob und in welchem Ausmaß Hybridisierung an der Artbildung beteiligt war, und ob Polyploidisierung durch Selbstverdopplung des Genoms stattfand. Zudem waren wir an der evolutionären Historie von Di- und Tetraploiden der beiden Artkomplexe interessiert. Wir näherten uns der Lösung dieser Fragestellungen sowohl auf weltweiter Ebene der Gesamtverbreitungsareale beider Artkomplexe als auch auf regionaler Ebene einer mitteleuropäischen Kontaktzone.

Im ersten Kapitel „Amphi-beringische, allopolyploide *Arabidopsis* und die evolutionäre Historie des *Arabidopsis lyrata* Komplexes“ charakterisierten wir drei genetische Hauptlinien, eine eurasiatische, nordamerikanische und amphi-pazifische, mit den molekularen Markern ntDNA ITS, ntDNA *Pgi*C und cpDNA *trn*L/F. Allopolyploidisierung zwischen eurasiatischer *Arabidopsis lyrata* ssp. *petraea* und ostasiatischer *Arabidopsis halleri* ssp. *gemmifera* in der amphi-pazifischen Linie ereignete sich dreimal unabhängig voneinander in Japan, China und Kamtschatka. Wir identifizierten die unvergletscherten Bereiche der ostösterreichischen Alpen und das arktische Eurasien einschließlich Beringias als eiszeitliche Hauptrefugialgebiete der eurasiatischen Linie. Die nordamerikanische Linie überdauerte die Vereisungen im Südosten Nordamerikas. Genfluss zwischen der eurasiatischen und nordamerikanischen Linie fand wahrscheinlich sowohl zwischen den Perioden der Vergletscherung als auch nach der letzten Vereisung statt.

Hybridisation and polyploidisation are two major driving forces for plant speciation. In the genus *Arabidopsis* hybridisation is reported from *Arabidopsis suecica* from Fennoscandinavia and *Arabidopsis kamchatica* from Japan. Within this study we focussed on the species complexes *Arabidopsis lyrata* and *Arabidopsis arenosa*. We aimed to clarify, if and to which extent hybridisation contributed to speciation, and if polyploidisation occurred via self-doubling of the genome. Moreover, we were interested in the evolutionary history of both diploids and tetraploids of the two species complexes. We investigated this on both the worldwide scale of their distribution range and the local scale of a Central European contact zone.

In the first chapter "An Amphi-Beringian allopolyploid *Arabidopsis* and the evolutionary history of the *Arabidopsis lyrata* complex" we characterised three main genetic lineages, Eurasian, North American, and amphi-pacific, with ntDNA ITS, ntDNA *Pgi*C, and cpDNA *trn*L/F molecular markers. The latter was identified to be of threefold independent allopolyploid origin between Eurasian *Arabidopsis lyrata* ssp. *petraea* and East Asian *Arabidopsis halleri* ssp. *gemmifera*, in Japan, China, and Kamtchatca. The major glacial refugia of the Eurasian lineage were the unglaciated parts of the Eastern Austrian Alps and arctic Eurasia, including Beringia. The North American lineage survived the glacials in the southeast of North America. Geneflow between the Eurasian and North American lineage probably occurred inter- and postglacially.

1.1. Introduction

Biological research within the last decade has largely focussed on model organisms like e.g. *Drosophila melanogaster*, *Caenorhabditis elegans*, or *Arabidopsis thaliana* in the plant kingdom. Now that knowledge in molecular genetics, cell and developmental biology of these organisms has greatly improved, closely related organisms are promising to study different characteristics, which make them well suited for answering biological questions, which are not possible to be elucidated with the classical model organisms (Mitchell-Olds, 2001; Benfey and Mitchell-Olds, 2008). In the plant kingdom *Arabidopsis lyrata* is a close relative of the model plant *Arabidopsis thaliana* and diverged approximately five million years ago (Clauss and Koch, 2006; Koch and Matschinger, 2007; Koch et al., 2008). *Arabidopsis lyrata* represents a small species complex with a circumpolar arctic-alpine distribution. Populations have been adapted to various ecological conditions, including the harsh environment of the arctic tundra, cryptic warm-stage refugia (exposed rocks, rocky slopes) in Central Europe, and different edaphic conditions with substrates such as dolomite, silicious bedrocks, and even heavy metal rich serpentine soil in Central Europe (Lower Austria, personal observation) and the USA (Maryland; Turner et al., 2008). Most members of the species complex are perennial outbreeders of mostly diploid ($2n = 2x = 16$), but also tetraploid cytotypes (Clauss and Koch, 2006; Schmickl et al., 2008a). There are numerous aspects of the *Arabidopsis lyrata* species complex that cannot be studied in *Arabidopsis thaliana*, as the latter is an inbreeding annual or winter annual plant with a temperate distribution range and narrow ecological amplitude. Furthermore, in *Arabidopsis thaliana* exclusively diploid cytotypes are reported with $2n = 2x = 10$. The *Arabidopsis lyrata* complex already proved to be a suitable study system for the analysis of character traits such as flowering time (Riihimäki and Savolainen, 2004; Riihimäki et al., 2005) and pathogen defense (Clauss et al., 2006). Additionally, molecular mechanisms of the function of sporophytic self-incompatibility were investigated (Kusaba et al., 2001; Schierup et al., 2001; Charlesworth et al., 2003; Mable et al., 2004; Mable et al., 2005; Charlesworth et al., 2006; Hagenblad et al., 2006; Schierup et al., 2006; Schierup et al., 2008), and comparative approaches of sporophytic self-incompatibility in diploids versus polyploids are underway (Jørgensen, unpublished data). Experimentally, research in *Arabidopsis lyrata* is facilitated by feasible crosses between *Arabidopsis thaliana* ($x = 5$) and *Arabidopsis lyrata* ($x = 8$) (Beaulieu et al., 2009). Furthermore, whole genome sequencing of *Arabidopsis lyrata* was finished last year, and data are available since few month (The *Arabidopsis lyrata* genome sequence assembly v1.0, http://genome.jgi-psf.org/Araly1/Araly1.info.html).

However, in contrast to *Arabidopsis thaliana*, where the evolutionary history has been analysed in more detail (e.g. Sharbel et al., 2000; Beck et al., 2008), evolutionary studies on the *Arabidopsis lyrata* complex were so far only restricted to a small number of populations from Central Europe (Wright et al., 2003; Balañá-Alcaide et al., 2006; Clauss and Mitchell-Olds, 2006; Muller et al., 2008) or larger sample sizes with a more general genus-wide perspective (Koch and Matschinger, 2007).

In this study we present the first worldwide evolutionary history of the *Arabidopsis lyrata* species complex, covering its whole distribution range and all taxonomically defined segregates, by using a widely applied nuclear encoded marker system (ITS, internal transcribed spacer region of nuclear encoded ribosomal DNA), a maternally inherited chloroplast genome marker (*trn*L intron (*trn*L) and *trn*L/F intergenic spacer (*trn*L/F-IGS) of $tRNA^{Ser}$ and $tRNA^{Thr}$, respectively), and the nuclear encoded housekeeping gene *Pgi*C (cytosolic phosphoglucoisomerase). We aim to focus on the following four aspects: (1) Unravelling general phylogeographic patterns with the main genetic lineages, and interpreting genetic variation in space and time (Avise et al., 1987), both in the context of climatic and geological events throughout Pleistocene glaciation and deglaciation cycles. (2) Evaluating the role of hybridisation and polyploidisation in *Arabidopsis kamchatica*, an Amphi-Beringian member of the *Arabidopsis lyrata* complex, and discussing, if these polyploids might have a broader adaptive potential. (3) Explaining Pleistocene and postglacial migration routes by analysing genetic diversity statistics: The arctic-alpine *Arabidopsis lyrata* complex is one of the rare examples for higher plants with a Central European-circumpolar-North American distribution. Here the effects of late quaternary climate oscillations, especially the last glaciation with its maximum extent, the Last glacial maximum (LGM) about 11500 years ago, can be studied on a global scale. (4) Studying the role of Beringia as a refuge area for populations of arctic *Arabidopsis lyrata*: Amongst other organisms numerous higher plants survived the glacials in refugia, which were locally unglaciated regions of constant environment (Tribsch and Schönswetter, 2003). Beringia is assumed to be one of the major refugia for arctic plants during Pleistocene glaciations (Tremblay and Schoen, 1999; Abbott et al., 2000; Brochmann et al., 2003; Abbott and Comes, 2004; Alsos et al., 2005; Oliver et al., 2006; Eidesen et al., 2007).

With our example focussing on the evolutionary history of the *Arabidopsis lyrata* species complex we want to supply the framework for future studies. The knowledge about centres of genetic diversity, different genetic lineages and their contact zones, and hybrid speciation is essential for any comprehensive evolutionary study.

1.2. Material and methods

1.2.1. Plant material

Altogether 493 accessions of the *Arabidopsis lyrata* complex were analysed: 141 accessions from Canada, China, Germany, Japan, Russia, Taiwan, and USA, 158 accessions from Austria, 39 accessions from Canada, Faeroe Islands, Greenland, Iceland, Japan, Scotland, Taiwan, and USA, sequenced by Schmickl et al. (2008a), and 155 accessions from Canada, central Europe, Faeroe Islands, Greenland, Iceland, Japan, Norway, Scotland, Sweden, Taiwan and USA, sequenced by Koch and Matschinger (2007). Plant material was mainly collected from herbarium vouchers from BM (Natural History Museum, London), CAS (California Academy of Sciences, San Fransisco), DAO (Vascular Plant Herbarium, Agriculture and Agri-Food Canada, Ottawa), DH (Hobert and William Smith Colleges, New York), LI (Upper Austrian Provincial Museum, Linz), W (Natural History Museum, Vienna), and partly collected in the field, documented at HEID (Herbarium Heidelberg, Heidelberg). Taxon determination followed the voucher labels, and was verified, if possible, with floras and determination keys (e.g. Al-Shehbaz and O'Kane, 2002; Flora of North America, Al-Shehbaz, personal communication). However, due to a lack of morphological studies based on a large number of accessions including the whole distribution range of the *Arabidopsis lyrata* complex (Eurasia and North America) taxon determination remained problematic. Especially North American "*kamchatica*" is difficult to distinguish from "*lyrata*", as already proposed by Elven et al. (2007). A comparison of both concepts is given in the results section as an introductionary note (TABLE 1). In our study we follow the taxonomy of Al-Shehbaz and O'Kane (2002). The distribution of the investigated accessions is shown in Fig. 1. The accession list is provided as Supplementary material TABLE 1.

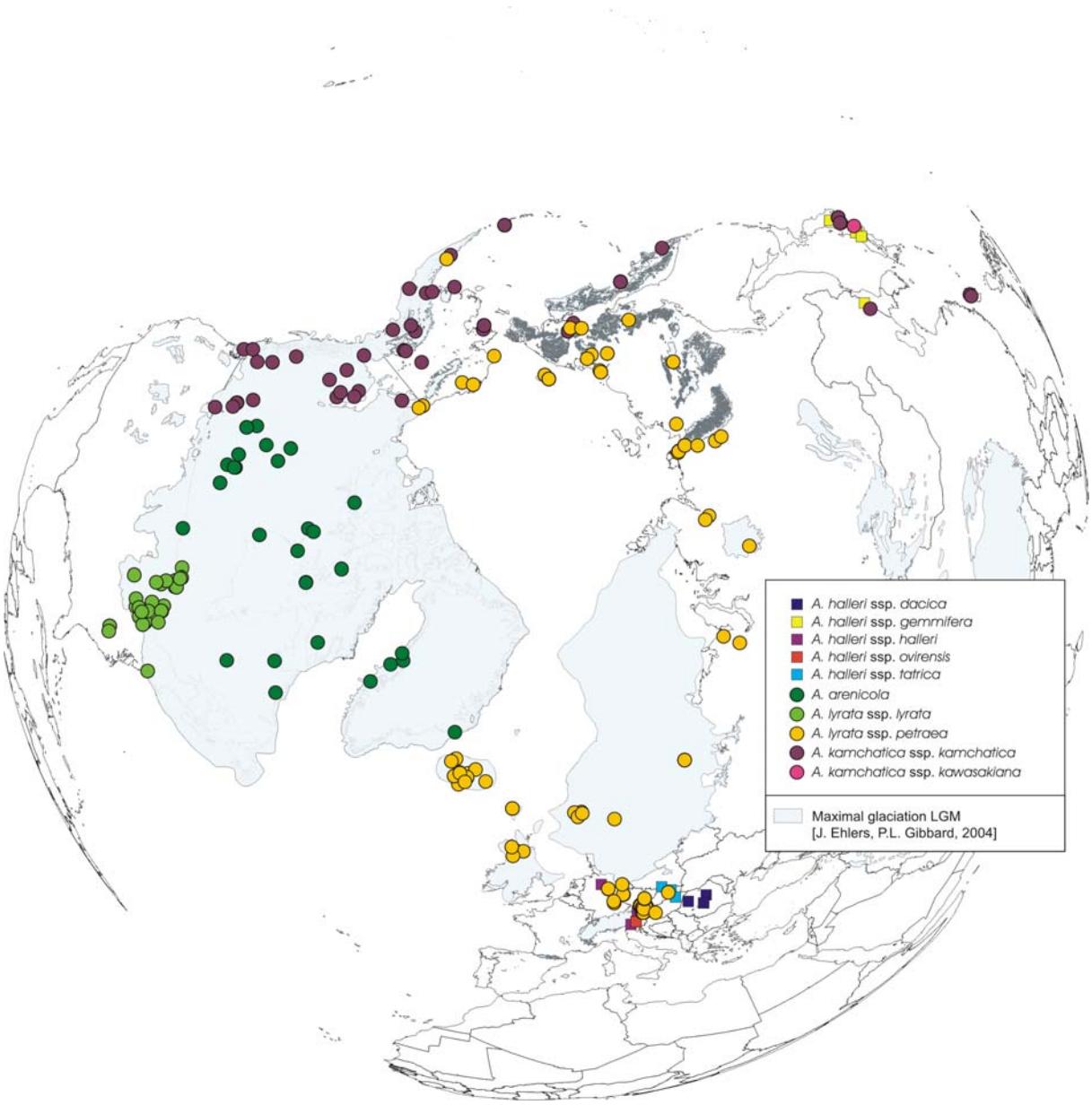


Fig. 1. Distribution of accessions investigated. Accessions of the *Arabidopsis lyrata* complex were analysed for ntDNA ITS and *Pgi*C, and cpDNA *trn*L/F. Accessions of the *Arabidopsis halleri* complex were analysed for *Pgi*C only. Maximal glaciation of the LGM was drawn according to Ehlers and Gibbard (2004).

1.2.2. DNA isolation, amplification and sequencing

Total DNA was obtained from dried leaf material and extracted according to the CTAB protocol of Doyle and Doyle (1987) with the following modifications: 50–75 mg of dry leaf tissue were ground in 2 ml tubes using a Retsch swing mill (MM 200), 2 units of ribonuclease per extraction were added to the isolation buffer, and the DNA pellets were washed twice with 70% ethanol. DNA was dissolved in 50 µl TE-buffer for storage and diluted 1:3 in TE-buffer before use. For the cpDNA markers *trn*L intron and *trn*L/F intergenic spacer (*trn*L/F-IGS) primer design and PCR cycling scheme followed the protocol of Dobeš et al. (2004a), using a PTC200 (MJ Research) thermal cycler. The PCR reaction volume of 50 µl contained 1x PCR buffer (10 mM TRIS/50 mM KCl buffer, pH 8.0), 3 mM $MgCl_2$, 0.4 µM of each primer, 0.2 mM of each dNTP, 1 U Taq DNA polymerase (Amersham), and approximately 1 ng of template DNA. Amplified sequences of *trn*L/F-IGS included the complete *trn*L/F-IGS and the first 18 bases of the *trn*F gene. Amplification of the nuclear marker internal transcribed spacer region (ITS) was performed according to primer design of Dobeš et al. (2004b). PCR reaction conditions were the same as for the two cpDNA markers described above, and PCR cycling scheme was 5 min at 95 °C, 35 cycles of 1 min at 95 °C, 1 min at 48 °C, and 1 min at 72 °C, 10 min extension at 72 °C, and a final hold at 4 °C. PCR products spanned the entire ITS1, 5.8 S rDNA, and ITS2 region. The nuclear marker cytosolic phosphoglucoisomerase (*Pgi*C) was amplified using the forward primer 5′-CATTCAACAGATTGTG-3′ and the reverse primer 5′-CCAGTAAACATCATGT-3′. Primer design is discussed in the following chapter. Here the PCR reaction volume of 50 µl contained 1x PCR buffer (10 mM TRIS/50 mM KCl buffer, pH 8.0), 2.5 mM $MgCl_2$, 0.13 µM of each primer, 0.2 mM of each dNTP, 1 U Taq DNA polymerase (Amersham), and approximately 1 ng of template DNA. The PCR cycling scheme was 3 min at 94 °C, 35 cycles of 20 sec at 94 °C, 30 sec at 56 °C, and 20 sec at 68 °C, 20 sec extension at 68 °C, and a final hold at 4 °C.

Before sequencing PCR products were checked for length and concentrations on 1.5% agarose gels and purified with the NucleoFast Kit (Macherey & Nagel, Germany). Cycle sequencing was performed using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences) and the original primers. Both strands were amplified in order to gain the complete sequence. PCR products were resolved in 10 µl Loading Solution and run on a MegaBace 500 sequencer.

1.2.3. Plastidic trnL/F and nuclear ITS sequence definition and map reconstruction

Plastidic *trn*L/F sequences were defined as haplotypes and suprahaplotypes, according to our previous studies (Koch and Matschinger, 2007; Koch et al., 2008; Schmickl et al., 2008a). As the 3´-region of the *trn*L/F-IGS close to the functional *trn*F gene is characterised by multiple *trn*F pseudogenes (Koch et al., 2005; Dobeš et al., 2007; Koch and Matschinger, 2007; Koch et al., 2007; Koch et al., 2008; Schmickl et al., 2008b), which evolve with a higher mutation rate than single nucleotide polymorphisms, we excluded the pseudogene region when defining *trn*L/F suprahaplotypes. Hence, haplotypes are defined as *trn*L intron - *trn*L/F-IGS *including* the 3´-end of the *trn*L/F-IGS with various copies of *trn*F pseudogenes, suprahaplotypes as *trn*L intron - *trn*L/F-IGS *excluding* the 3´-end of the *trn*L/F-IGS with various copies of *trn*F pseudogenes. Consequently, each *trn*L/F suprahaplotype comprises a varying number of haplotypes, which vary both in length and base content of their pseudogene-rich region. ITS sequences were obtained from direct sequencing of PCR products. Sequences were defined as ITS types and supratypes due to our previous studies (Koch and Matschinger, 2007; Koch et al., 2008; Schmickl et al., 2008a). Each ITS type contained ambiguous sites as a result of multiple copies due to direct sequencing. As the ITS region generally underlies concerted evolution, multiple ITS copies either indicate the natural variability among ITS loci within an individual or hybridisation between individuals.

Maps were constructed using BioOffice version 2.0.6 to create shapefiles and drawn with ArcView version 8.2. Shapefiles for visualising the maximum extent of the ice sheets during the LGM were kindly provided by Jürgen Ehlers and Phil Gibbard (2004).

1.2.4. Network analyses and genetic diversity statistics

Network analyses and genetic diversity statistics were exclusively performed with suprahaplotypes. The *trn*L/F network was constructed using TCS version 1.21 (Clement et al., 2000), indels (except polyT stretches) were coded as additional binary characters. Newly characterised suprahaplotypes were added to the *trn*L/F network of the *Arabidopsis lyrata* complex published earlier (Schmickl et al., 2008a). Genetic diversity statistics were performed with Arlequin version 3.11 (Excoffier and Schneider, 2005). Pairwise genetic differentiation was calculated among the following nine taxonomic and regional groups: *Arabidopsis lyrata* ssp. *petraea* from (1) unglaciated Central Europe, (2) glaciated northern Europe, (3) northern Russia and western Beringia, (4) eastern Beringia, *Arabidopsis lyrata* ssp. *lyrata* from (5) unglaciated North America and glaciated Great Lakes region, (6) glaciated

North America and Greenland (*Arabidopsis arenicola* nom. prov.), and *Arabidopsis kamchatica* from (7) Japan, (8) Russia Far East, (9) Alaska and western Canada. F_{ST} values, regarding haplotype frequencies only, and Φ_{ST} values, additionally regarding haplotype sequence information, were calculated based on a permutation test with 1000 permutations. Additionally genetic diversity was estimated as nucleotide diversity π.

ITS data were analysed based on ITS supratypes. From both the *Arabidopsis lyrata* and *Arabidopsis halleri* complex a strict consensus tree was constructed using Maximum Parsimony with MEGA version 4.1 (Kumar et al., 2008).

1.2.5. Primer design for the nuclear marker PgiC

Various higher plants are known to have a duplicated locus of the cytosolic enzyme phosphoglucoisomerase (Gottlieb, 1977; Ghatnekar, 1999; Ford and Gottlieb, 2002), both loci are normally unlinked. Also within the genus *Arabidopsis* sequencing of the *Pgi*C locus unrevealed a duplicated locus, both in the *Arabidopsis lyrata* and *Arabidopsis halleri* complex, based on a Maximum Likelihood strict consensus tree (Fig. 2; Jørgensen, unpublished data). Both loci were initially amplified with the general forward primer 5′-TGCTGTSAGCACTAATCTTGCG-3′ and the general reverse primer 5′-TCGAACCCGGGAGAGGTAGACCA-3′ (Wright et al., 2003), following the protocol of Wright et al. (2003). We used the duplicated *Pgi*C locus and its various alleles to discriminate between the *Arabidopsis lyrata* and *Arabidopsis halleri* complex. According to the sequence data we obtained, the duplicated *Pgi*C locus (denoted here as *Pgi*C2) was only weakly differentiated between the two species complexes. However, the *Pgi*C1 locus showed a deletion of 7 bp length in alleles carried only by members of the *Arabidopsis lyrata* complex in contrast to alleles without this deletion carried by members of the *Arabidopsis halleri* complex (Supplementary material Fig. 1). Both groups of alleles are also substantially differentiated by various SNPs. Consequently, a primer pair with the forward primer located partly within the indel (5′-CATTCAACAGATTGTG-3′) and the reverse primer 5′-CCAGTAAACATCATGT-3′ was developed to amplify an approximately 50 bp fragment within the *Pgi*C1 locus, specific for alleles carried by *Arabidopsis halleri* and its various subspecies. With this taxon-specific primer pair we aimed to proof, if a member of the *Arabidopsis halleri* complex was one parental species of tetraploid *Arabidopsis kamchatica* throughout its whole distribution range.

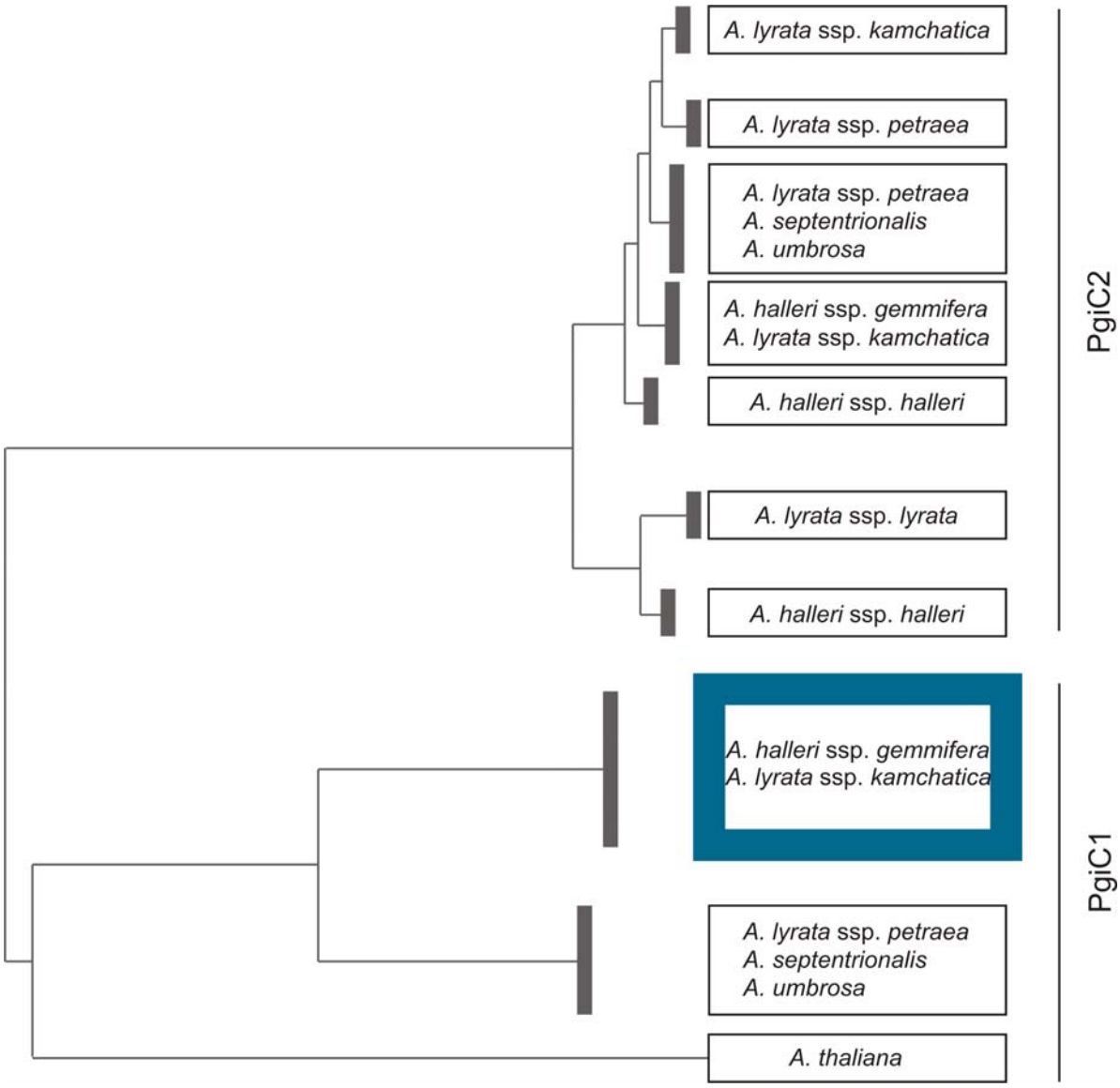


Fig. 2. One of 150 most parsimonious trees based on 130 *Arabidopsis* ntDNA *Pgi*C sequences, modified after Jørgensen et al. (unpublished data). Investigated accessions were both from the *Arabidopsis lyrata* complex (*Arabidopsis lyrata* ssp. *lyrata*, *Arabidopsis lyrata* ssp. *petraea*, *Arabidopsis septentrionalis*, and *Arabidopsis umbrosa*) and the *Arabidopsis halleri* complex (*Arabidopsis halleri* ssp. *halleri*, and *Arabidopsis halleri* ssp. *gemmifera*). Taxa with successful amplification of the chosen *Pgi*C1 fragment, and, consequently, without the deletion in the forward primer site, are marked with the blue box.

1.3. Results

The following short overview of *Arabidopsis halleri* and *Arabidopsis lyrata* taxonomy will introduce the two main taxonomic concepts. *Arabidopsis halleri* could be devided into five subspecies (ssp. *halleri*, ssp. *ovirensis*, ssp. *dacica*, ssp. *tatrica*, all distributed in Central Europe, and ssp. *gemmifera* in Eastern Asia). This is not only supported by morphometric analysis (Kolnik, unpublished data) but also genetic AFLP data (Marhold, unpublished data). The *Arabidopsis lyrata* complex is one of three major lineages in the genus *Arabidopsis* (Koch and Matschinger, 2007), the other two being *Arabidopsis arenosa* and *Arabidopsis halleri* lineages as described above (O'Kane and Al-Shehbaz, 1997; Al-Shehbaz et al., 1999; Clauss and Koch, 2006; Koch and Matschinger, 2007; Koch et al., 2008). The *Arabidopsis lyrata* complex is considered by different authors to include a diverse number of taxa of various ranking, distribution areas, and ploidy levels (TABLE 1). Al-Shehbaz and O'Kane (2002) treat the complex as one species, *Arabidopsis lyrata*, with three subspecies: ssp. *lyrata*, ssp. *petraea*, and ssp. *kamchatica*. The first is considered to be broadly amphi-pacific and of two ploidy levels ($2n$ = 16, 32), the second to be Northern Eurasian and Central European with the same two ploidy levels, and the third to be amphi-pacific and tetraploid ($2n$ = 32). Elven et al. (2007) rank *petraea* and *kamchatica* ($2n$ = 32) as species, the first potentially containing three different subspecies: ssp. *petraea* ($2n$ = 16), ssp. *septentrionalis* ($2n$ = 16, 32), and ssp. *umbrosa* ($2n$ = 16), the second corresponding more or less to the *lyrata* ssp. *kamchatica* sensu O'Kane and Al-Shehbaz. They consider ssp. *petraea* to be distributed in Central and Northern Europe, ssp. *septentrionalis* in Siberia, and ssp. *umbrosa* in the Russian Far East, Alaska, and Canada. The origin of the different concepts with regard to the taxonomy does not rely on different morphological traits, but the treatment is complicated by a lack of discrete and unique characters. A comprehensive quantitative morphological analysis using multivariate statistics is still not available. Furthermore, for the description of some taxonomic units information on ploidy level, which is also largely lacking, might be important. In summary, the suggested taxonomic units are largely concordant, but grouped into different subgroups or subdivided/ranked differently. As a consequence, herein, we discuss our molecular findings in the light of a simplified taxonomic concept. *Arabidopsis arenicola* has only recently been included as part of the *Arabidopsis lyrata* complex (Warwick et al., 2006). It has, for a long time, been placed within the genus *Arabis*. Elven et al. (2007) consider this taxon to be diploid ($2n$ = 16) and distributed in north-eastern North

America, and they are supported by Al-Shehbaz in the upcoming Flora of North America (personal communication).

Many of the mentioned taxa include two ploidy levels, suggesting frequent polyploidisation events within the *Arabidopsis lyrata* complex. An allopolyploid origin of *kamchatica* was suggested based on nuclear DNA sequences, with *Arabidopsis lyrata* and *Arabidopsis halleri* ssp. *gemmifera* as possible parental taxa (Shimizu et al., 2005; Koch and Matschinger, 2007; Schmickl et al., 2008a). Otherwise, little is known with regard to the number of polyploid units and their origin.

Taxonomy according to Al-Shehbaz and O'Kane (2002)			
Species	Subspecies	Ploidy	Distribution range
A. lyrata (L.) O'Kane & Al-Shehbaz	ssp. *petraea* (L.) O'Kane & Al-Shehbaz	2*n* = 16, 32	Central Europe, N Eurasia
	ssp. *kamchatica* (Fischer ex de Candolle) O'Kane & Al-Shehbaz	2*n* = 32	Amphi-Pacific
	ssp. *lyrata*	2*n* = 16, 32	North America
A. arenicola (Richardson) Al-Shehbaz, Elven, D.F.Murray & Warwick		2*n* = 16	NE North America
Taxonomy according to Elven et al. (2007), excluding *A. lyrata* from taxonomic treatment			
Species	Subspecies	Ploidy	Distribution range
A. petraea	ssp. *petraea*	2*n* = 16	Central and N Europe
	ssp. *septentrionalis* (N.Busch) D.F.Murray & Elven	2*n* = 16, 32	Siberia
	ssp. *umbrosa* (Turcz. ex Steud.) D.F.Murray & Elven	2*n* = 16	Russian Far East, Alaska, Canada
A. kamchatica		2*n* = 32	Amphi-Pacific
A. arenicola (Richardson ex Hooker) Petrovsky & Elven comb. nov.			

TABLE 1. The two main taxonomic concepts of the *Arabidopsis lyrata* complex, summarised from Al-Shehbaz and O´Kane (2002), including revision from the Flora of North America (Al-Shehbaz, personal communication), and Elven et al. (2007). *Arabidopsis arenicola* was integrated from Warwick et al. (2006). Elven et al. (2007) excluded *Arabidopsis lyrata* from taxonomic treatment, as they assumed it to be a non-arctic, boreal taxon.

1.3.1. Chloroplast sequence data indicate three main genetic lineages: Eurasia, North America, and the amphi-pacific region

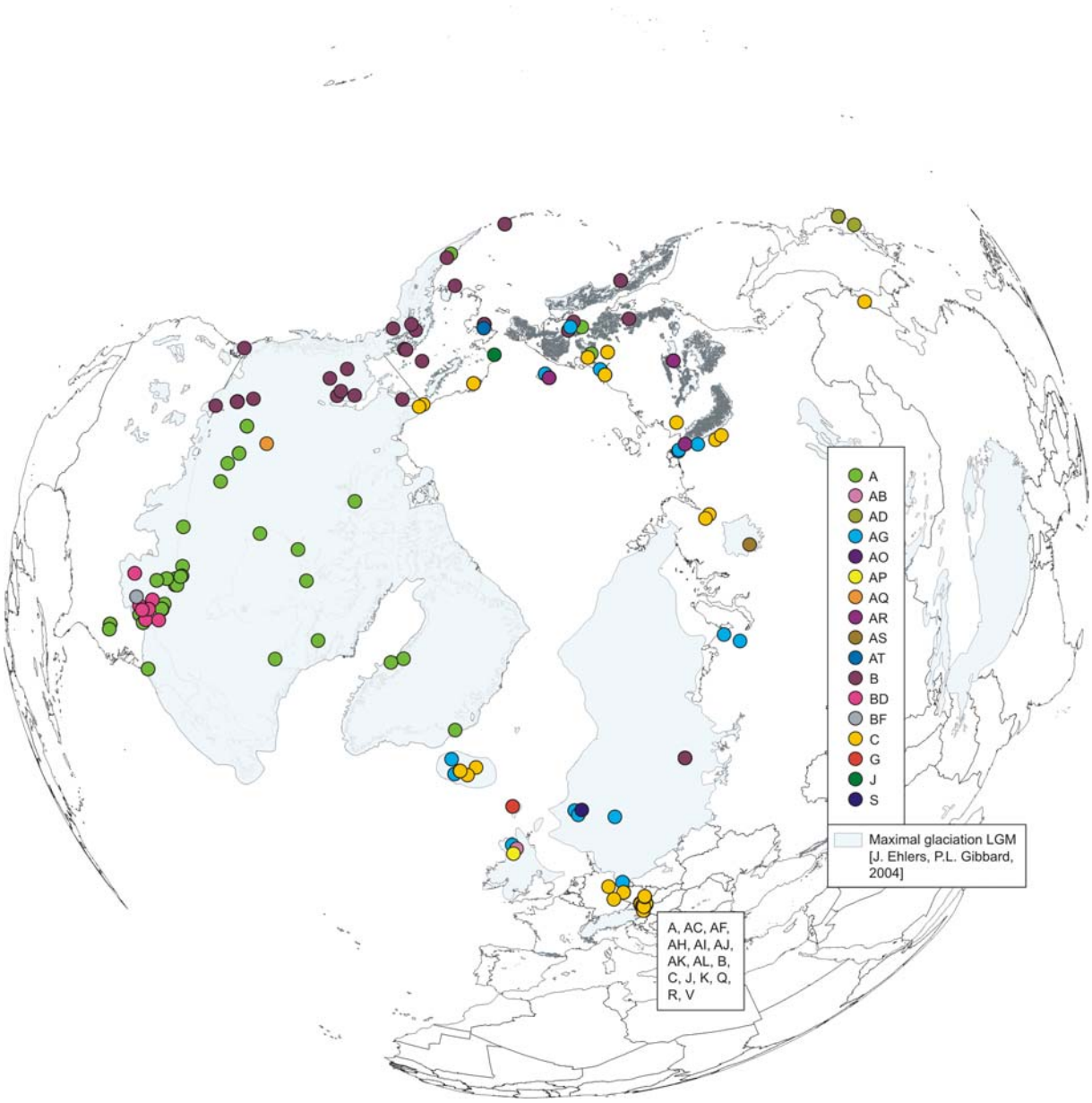


Fig. 3. Distribution of cpDNA *trn*L/F suprahaplotypes in the *Arabidopsis lyrata* complex, combining the results obtained in this study with the results of Koch and Matschinger (2007) and Schmickl et al. (2008a). Accessions from Austria are listed separately. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard (2004).

The investigated accessions, spanning the whole distribution range of *Arabidopsis lyrata* species complex, could be grouped into three genetic lineages, which were clearly separated geographically and characterised by major and widely distributed cpDNA types: cpDNA *trn*L/F suprahaplotype C in Eurasia, A in North America, and B in the amphi-pacific region (Fig. 3). The Eurasian lineage with suprahaplotype C had the largest distribution, found in unglaciated Central Europe, formally glaciated northern Europe, arctic Russia, Beringia, and Alaska (north of Brooks Range). The North American lineage, characterised by suprahaplotype A, included the United States (mainly around the Great Lakes), northeastern and central Canada (up to the Canadian Rocky Mountains in the west), and Greenland. The third and Amphi-Pacific lineage spanned from Kamtchatca via Beringia into western Canada (with the Canadian Rocky Mountains as eastern border). These three suprahaplotypes are central in the suprahaplotype network (Fig. 4), indicating the split of an ancient ancestor into these three lineages. All three genetic lineages were additionally characterised by younger, lineage-specific "tip" suprahaplotypes, which are derived from the central suprahaplotypes (Fig. 4): In the Eurasian lineage locally widespread AC was found in Central Europe, mainly Austria, AG predominantly in the Arctic (Iceland, Scandinavia, Russia), and AR in western Beringia (incl. Wrangel Island) (Fig. 3). Unique suprahaplotypes, occurring only once in the whole dataset, were observed in Scotland (AB, AS), Austria (AH, AI, AJ, AK, AL, K, R, V), Iceland (AO, AP), G (Faroe Islands), and S (Sweden). The North American lineage was additionally characterised by suprahaplotypes AQ, BD and BF, with BD more widespread, and unique AQ and BF. In the amphi-pacific lineage AD was detected exclusively in Japan. Although all three lineages were characterised both by lineage-specific central and "tip" suprahaplotypes, sharing of central suprahaplotypes was observed, e.g. in few accessions of the Eurasian lineage suprahaplotypes A and B were detected (Fig. 3). This finding is congruent with the observation of central suprahaplotype sharing between the three main species complexes of the genus, *Arabidopsis lyrata*, *Arabidopsis halleri*, and *Arabidopsis arenosa* (Koch and Matschinger, 2007). This observation has been explained by ancestral cpDNA polymorphism predating the genus´ radiation approximately two million years ago (Koch and Matschinger, 2007).

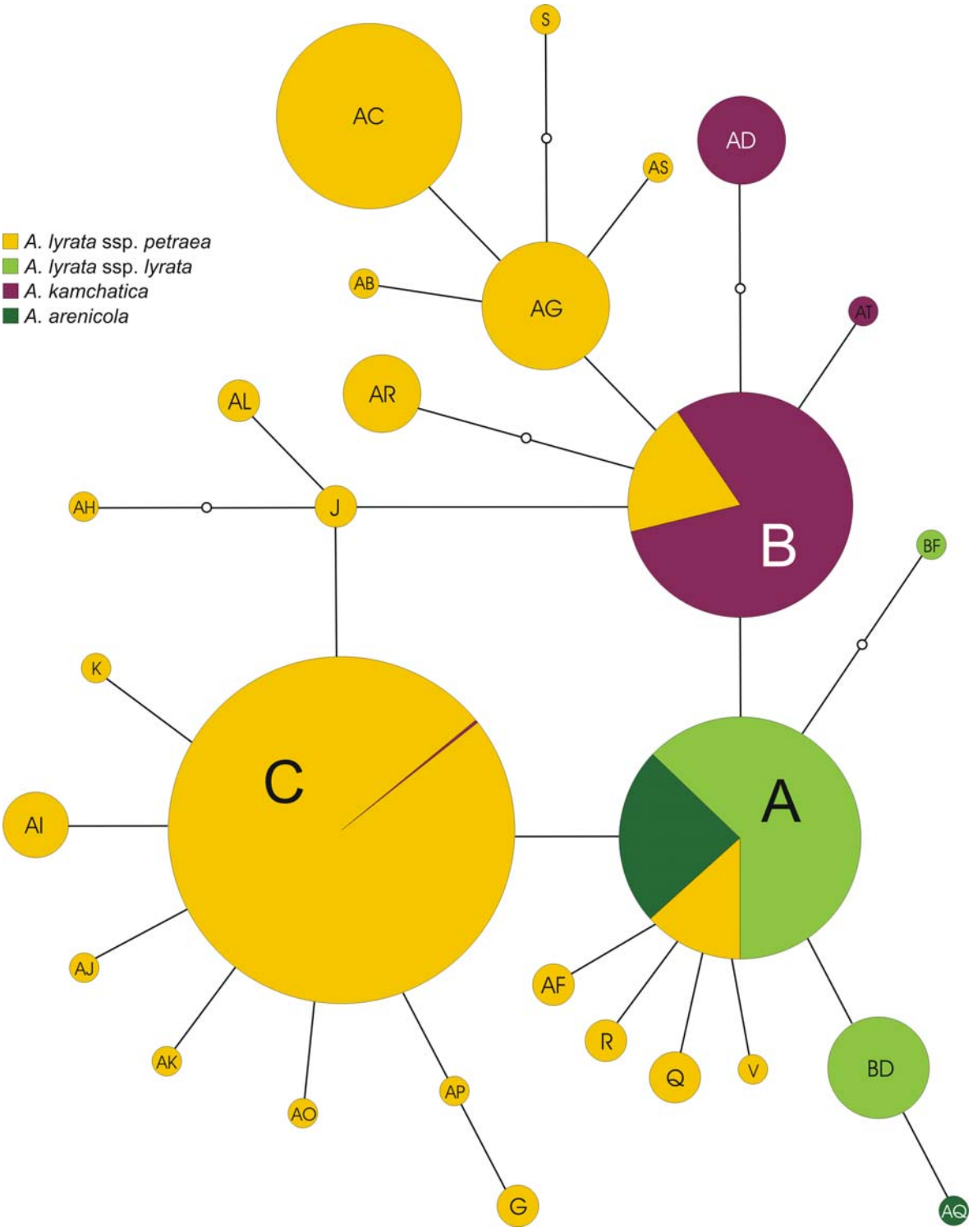


Fig. 4. CpDNA *trn*L/F suprahaplotype network of the *Arabidopsis lyrata* complex, combining the results obtained in this study with the results of Koch and Matschinger (2007) and Schmickl et al. (2008a). The sizes of the circles indicate the relative frequency of a suprahaplotype.

1.3.2. Cytosolic phosphoglucose isomerase prooves multiple hybrid origin of amphi-pacific Arabidopsis kamchatica

Arabidopsis kamchatica from Japan is a tetraploid taxon and was already proven to be of allopolyploid origin, with the paternal genome from a member of the *Arabidopsis lyrata* species complex and the maternal genome from *Arabidopsis halleri* ssp. *gemmifera*, a member of the *Arabidopsis halleri* species complex (Shimizu et al., 2005; Koch and Matschinger, 2007). In case of the herein investigated duplicated *Pgi*C gene we detected a taxon-specific group of alleles at the *Pgi*C1 locus: Alleles were only found in *Arabidopsis halleri* ssp. *gemmifera* and other subspecies of *Arabidopsis halleri*, but not in members of the *Arabidopsis lyrata* complex (Fig. 2). Appropriate selective primers have been successfully developed utilising a deletion in the forward primer binding site (Supplementary material Fig. 1). Unfortunately, it was not possible to develop a PCR-based reciprocal marker system characterising alleles from the *Arabidopsis lyrata* gene pool. Detection of this distinct *Pgi*C1 allele pool in *Arabidopsis kamchatica* must, therefore, be the result of introgression of alleles of *Arabidopsis halleri* ssp. *gemmifera* and indicate hybrid origin of tetraploid *Arabidopsis kamchatica*. Amplification of these *Pgi*C1 alleles was successful in all *Arabidopsis kamchatica* accessions, including those from China, Taiwan and Japan, as well as accessions from eastern Russia, western and eastern Beringia, Alaska, and western Canada (Fig. 5, 6). As expected, *Pgi*C1 amplification was also positive in all *Arabidopsis halleri* ssp. *gemmifera* accessions and additionally in European subspecies of *Arabidopsis halleri* (*Arabidopsis halleri* ssp. *dacica*, *Arabidopsis halleri* ssp. *halleri*, *Arabidopsis halleri* ssp. *tatrica*) (Fig. 5, 6). Hence, occurrence of *Pgi*C1 alleles without deletion in the forward primer sequence was characteristic for the whole *Arabidopsis halleri* species complex, except for *Arabidopsis halleri* ssp. *ovirensis* (data not shown), where either a secondary loss of this locus or a complementary mutation in the primer binding site might have occured. *Arabidopsis halleri* ssp. *ovirensis* is a genetically distinct, local endemite of the southeastern Austrian Alps (Koch and Matschinger, 2007; Koch et al., 2008).

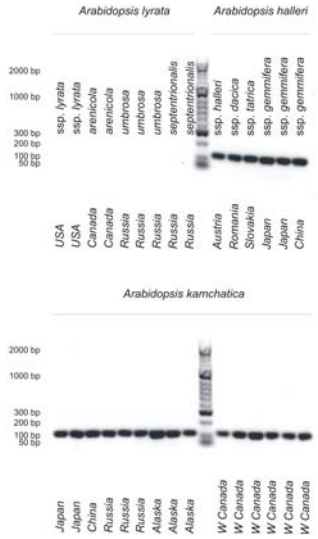


Fig. 5. Unsuccessful *Pgi*C1 amplification in members of *Arabidopsis lyrata* (*Arabidopsis lyrata* ssp. *lyrata*, *Arabidopsis arenicola*, *Arabidopsis umbrosa*, and *Arabidopsis septentrionalis*). Successful *Pgi*C1 amplification in members of *Arabidopsis halleri* (ssp. *halleri*, ssp. *dacica*, ssp. *tatrica*, and ssp. *gemmifera*), and *Arabidopsis kamchatica*.

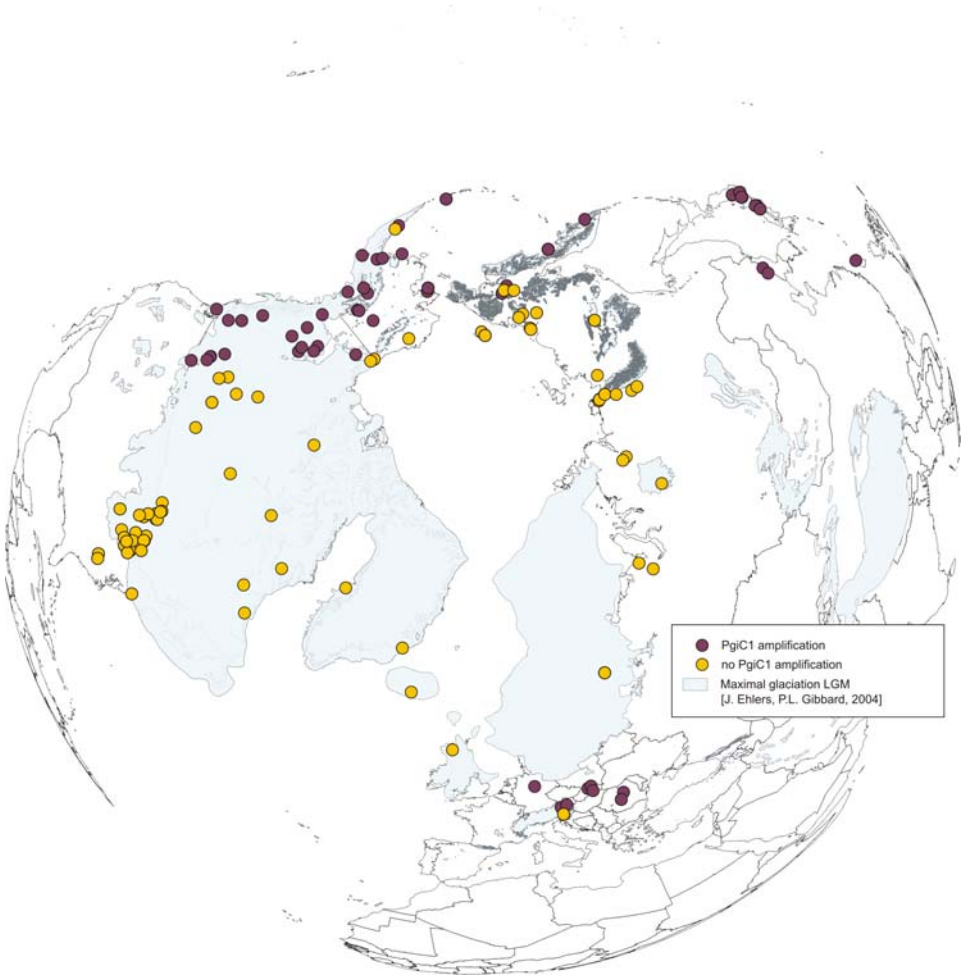


Fig. 6. Distribution of accessions with/without *Pgi*C1 amplification in the *Arabidopsis lyrata* complex, in *Arabidopsis halleri* ssp. *gemmifera* from China and Japan, and in Central European subspecies of the *Arabidopsis halleri* complex (ssp. *dacica*, ssp. *halleri*, ssp. *ovirensis*, and ssp. *tatrica*). Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard (2004).

*Pgi*C1 amplification failed in Eurasian and North American members of *Arabidopsis lyrata*. Therefore, an allopolyploid origin of *Arabidopsis kamchatica* with *Arabidopsis halleri* ssp. *gemmifera* as one parental taxon and a member of *Arabidopsis lyrata* as the second parental taxon could be confirmed not only for Japanese, but for all *Arabidopsis kamchatica* accessions. However, *Arabidopsis kamchatica* is genetically heterogenuous throughout its whole distribution range from East Asia to western Canada. According to chloroplast *trn*L/F data, three different genetic groups were found, which are geographically isolated from each other (Fig. 4): (1) accessions of a widespread distribution range from Kamtchatca, western and eastern Beringia, to pacific western Canada (*trn*L/F suprahaplotype B), (2) Japanese accessions (*trn*L/F suprahaplotype AD), and (3) accessions from pacific eastern China (*trn*L/F suprahaplotype C). In a more detailed comparison of chloroplast *trn*L/F and nuclear encoded ITS sequence data two directions of geneflow could be observed (Fig. 7): Either the paternal genome originated from a member of the *Arabidopsis lyrata* species complex (ITS supratype b) and the maternal genome from *Arabidopsis halleri* ssp. *gemmifera* (*trn*L/F suprahaplotype AD), as already reported for *Arabidopsis kamchatica* from Japan (Koch and Matschinger, 2007). Or *Arabidopsis halleri* ssp. *gemmifera* represented the paternal genome (ITS supratype – defined here as ribosomal internal transcribed spacers 1 and 2 without ambiguous sites – z, derived from ITS supratype r exclusively found in *Arabidopsis halleri* ssp. *gemmifera*; Fig. 7), and a member of the *Arabidopsis lyrata* species complex served as donor of the maternal genome (*trn*L/F suprahaplotype C), as in *Arabidopsis kamchatica* from China. The North American lineage of the *Arabidopsis lyrata* complex could be excluded as a hybridisation partner, as neither ITS supratype e nor *trn*L/F suprahaplotype A were detected in *Arabidopsis kamchatica*. According to these data we suggest at least three independent origins of *Arabidopsis kamchatica*, first with maternal *Arabidopsis halleri* ssp. *gemmifera* in Japan, second with paternal *Arabidopsis halleri* ssp. *gemmifera* in China, and third with an unknown direction of geneflow in Kamtchatca, but in all cases with a member of the Eurasian *Arabidopsis lyrata* lineage as hybridisation partner.

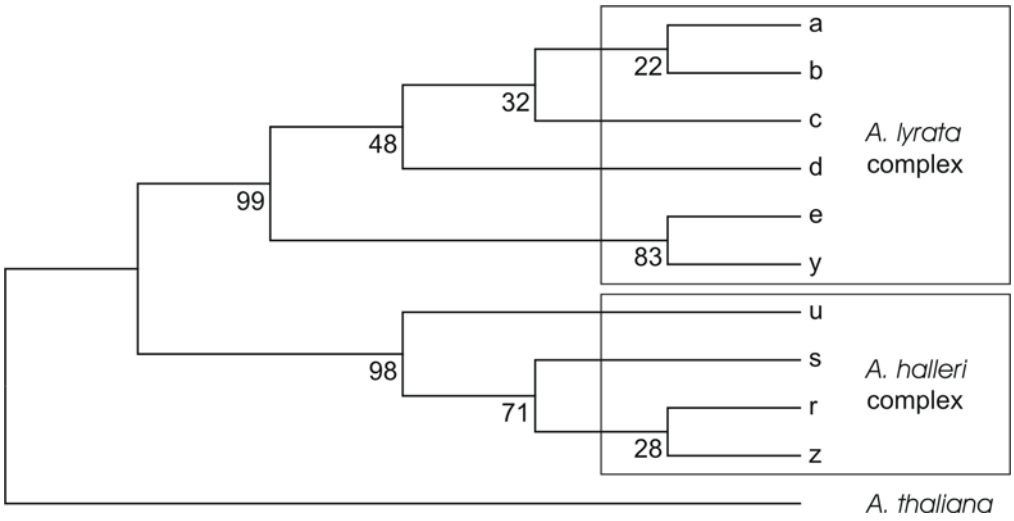


Fig. 7. Maximum Parsimony analysis (consensus tree) of ITS supratypes in the *Arabidopsis lyrata* and *Arabidopsis halleri* complex with bootstrap values indicated.

1.3.3. Gene diversity statistics shows highest genetic diversity in the Eurasian lineage, strongly reduced diversity in the North American lineage, and extremely low diversity in the allopolyploid amphi-pacific lineage

For genetic diversity statistics the distribution range of the *Arabidopsis lyrata* species complex was divided into nine different groups, according to genetic lineages and Pleistocene history (TABLE 2). The Eurasian lineage (*Arabidopsis lyrata* ssp. *petraea*) was split into four groups: unglaciated Central Europe, previously glaciated northern Europe or permafrost areas, northern Russia/western Beringia, and eastern Beringia. The North American lineage (*Arabidopsis lyrata* ssp. *lyrata*) was separated into unglaciated North America/glaciated Great Lakes region and glaciated North America/Greenland. Amphi-pacific *Arabidopsis kamchatica* was differentiated into groups from Japan, Russia Far East, and Alaska/western Canada. The Eurasian lineage of *Arabidopsis lyrata* ssp. *petraea* showed the highest cpDNA-based (*trn*L/F) nucleotide diversity (π), both in the Arctic and Central Europe. Interestingly, it was equally high in unglaciated Central Europe ($\pi = 0.0060$), glaciated northern Europe ($\pi = 0.0076$), and northern Russia/western Beringia ($\pi = 0.0077$), but significantly reduced in eastern Beringia (0.0034). Pairwise Φ_{ST} and F_{ST} calculations also showed only a slight differentiation of unglaciated Central Europe, glaciated northern Europe, and northern Russia/western Beringia – for eastern Beringia data were not statistically significant (TABLE 4). These observations indicate long-term survival of the *Arabidopsis lyrata* complex both in unglaciated Central

Europe and northern Russia/western Beringia and postglacial colonisation of formerly glaciated northern Europe from those two regions. The generally high nucleotide diversity in the Eurasian lineage was caused by a high number of unique and rare suprahaplotypes (Fig. 3, TABLE 3). In Central Europe the highest number of unique *trn*L/F suprahaplotypes (AC, AF, AH, AI, AJ, AK, AL, J, K, Q, R, V) was found in the foothills of the Eastern Austrian Alps, which remained unglaciated during Pleistocene climate oscillations. Together with the finding of both diploid and cytogenetically stabilised tetraploid populations (Schmickl and Koch, unpublished data), an old refugium of *Arabidopsis lyrata* ssp. *petraea* was indicated for this region. A high number of unique and rare suprahaplotypes was also found in formerly glaciated northern Europe (AB, AP – Scotland, AO – Iceland, G – Faeroer Islands, S – Norway). Subsequent geographic isolation of those populations during Holocene warming might have caused restriction of these suprahaplotypes to single geographic locations.

In contrast to the Eurasian lineage of *Arabidopsis lyrata* ssp. *petraea*, the North American lineage showed approximately tenfold reduced nucleotide diversity and strong differentiation according to pairwise Φ_{ST} and F_{ST} (TABLE 4). Hence, immigration from Eurasia into North America via the Bering Land Bridge, which had formed several times during Pleistocene climate oscillations, can be assumed. Although crossing this narrow land bridge was obviously correlated with strong genetic bottlenecks, Pleistocene climate fluctuations obviously led to slight differences between formerly glaciated and unglaciated regions within the North American lineage: Nucleotide diversity of accessions from predominantly unglaciated southeastern North America ($\pi = 0.0008$) was slightly higher than from North America and Greenland ($\pi = 0.0003$) (TABLE 2), which had been under the Laurentide ice sheet during the LGM, indicating genetic bottlenecks with subsequent rapid postglacial immigration.

Extremely reduced nucleotide diversity was reported from amphi-pacific *Arabidopsis kamchatica* ($\pi = 0.0000$) (TABLE 2). If all three genetic groups of *Arabidopsis kamchatica* are treated separately, only a single *trn*L/F suprahaplotype was found in each group (Japan: AD; China: C; Russia Far East/Alaska/western Canada: B) (TABLE 3). Additionally, haplotype diversity was low (haplotypes are defined here as *trn*L intron - *trn*L/F intergenic spacer *including* the 3´-end of the *trn*L/F intergenic spacer with various copies of *trn*F pseudogenes), especially in *Arabidopsis kamchatica* with suprahaplotype B. Only one haplotype (no. 84) was detected over a vast amphi-pacific area from Kamtchatca to western Canada (TABLE 5). This extremely reduced genetic diversity, particularly of the group with suprahaplotype B, can probably best be explained by rapid postglacial colonisation. The very strong differentiation

of the amphi-pacific lineage with both the Eurasian, but especially the North American lineage, according to pairwise Φ_{ST} and F_{ST} (TABLE 4), furthermore underscores the distinct genetic position of *Arabidopsis kamchatica* as a hybrid taxon, more closely related to the Eurasian than the North American lineage.

Geographic region	*n*	π
A. lyr. ssp. *petraea*: Unglac. Central Europe	196	0.006026 +/- 0.003291
A. lyr. ssp. *petraea*: Glac. N Europe	21	0.007639 +/- 0.004249
A. lyr. ssp. *petraea*: N Russia, W Beringia	33	0.007741 +/- 0.004219
A. lyr. ssp. *petraea*: E Beringia	7	0.003392 +/- 0.002363
A. lyr. ssp. *lyrata*: Unglac. N America, glac. Great Lakes region	55	0.000798 +/- 0.000716
A. arenicola: Glac. N America, Greenland	17	0.000312 +/- 0.000424
A. kamch.: Japan	9	0.000000 +/- 0.000000
A. kamch.: Russia Far East	9	0.000000 +/- 0.000000
A. kamch.: Alaska, W Canada	38	0.000070 +/- 0.000185

TABLE 2. Taxonomic and regional genetic differentiation based on cpDNA suprahaplotypes [sample size (n) and nucleotide diversity π +/- standard deviation].

TABLE 3 (next page). Taxonomic and regional genetic differentiation based on cpDNA suprahaplotypes [number of cpDNA suprahaplotypes occurring in each region].

Geographic region	A	AB	AC	AD	AF	AG	AH	AI	AJ	AK	AL	AO	AP	AQ	AR	AS	AT	B	BD	BF	C	G	J	K	Q	R	S	V
A. lyr. ssp. *petraea*: Unglac. Central Europe	7		40		2	1	1	5	1	1	2							8			120		1	1	3	2		1
A. lyr. ssp. *petraea*: Glac. N Europe		1				7						1	1					4			4	2					1	
A. lyr. ssp. *petraea*: N Russia, W Beringia	2					11									7	1		1			11							
A. lyr. ssp. *petraea*: E Beringia	1																				5		1					
A. lyr. ssp. *lyrata*: Unglac. N America, glac. Great Lakes region	42																		12	1								
A. lyr. ssp. *arenicola*: Glac. N America, Greenland	16													1														
A. kamtch.: Japan				9																								
A. kamtch.: Russia Far East																		9										
A. kamtch.: Alaska, W Canada																	1	37										

Φ_{ST}\F_{ST}	*A. lyr.* ssp. *petraea*: Unglac. Central Europe	*A. lyr.* ssp. *petraea*: Glac. N Europe	*A. lyr.* ssp. *petraea*: N Russia, W Beringia	*A. lyr.* ssp. *petraea*: E Beringia	*A. lyr.* ssp. *lyrata*: Unglac. N America, glac. Great Lakes region	*A. arenicola*: Glac. N America, Greenland	*A. kamch.*: Japan	*A. kamch.*: Russia Far East	*A. kamch.*: Alaska, W Canada
A. lyr. ssp. *petraea*: Unglac. Central Europe		0.21832	0.17315	[-0.00029]	0.47816	0.52477	0.54797	0.53255	0.56147
A. lyr. ssp. *petraea*: Glac. N Europe	0.14828		[0.03253]	0.17947	0.44407	0.50004	0.48036	0.37743	0.53503
A. lyr. ssp. *petraea*: N Russia, W Beringia	0.19685	[-0.00620]		0.12632	0.43290	0.48110	0.49746	0.48437	0.60275
A. lyr. ssp. *petraea*: E Beringia	[-0.00716]	0.23663	0.26978		0.53564	0.71126	0.76892	0.76892	0.86207
A. lyr. ssp. *lyrata*: Unglac. N America, glac. Great Lakes region	0.37925	0.54668	0.49762	0.79114		[0.08456]	0.70822	0.70822	0.76248
A. arenicola: Glac. N America, Greenland	0.33441	0.40526	0.37345	0.77234	[0.02119]		0.92283	0.92283	0.92688
A. kamch.: Japan	0.59120	0.45048	0.40209	0.88490	0.93046	0.97806		1.00000	0.95735
A. kamch.: Russia Far East	0.49758	0.25161	0.19769	0.85450	0.90536	0.96954	1.00000		[-0.05564]
A. kamch.: Alaska, W Canada	0.54130	0.44301	0.33922	0.94213	0.93015	0.97891	0.97880	[-0.05564]	

TABLE 4. Pairwise genetic differentiation among taxonomic and regional groups: F_{ST} (on diagonal) and Φ_{ST} (below diagonal) are both estimated from cpDNA suprahaplotypes. Values with $P < 0.05$ are given in brackets (permutation test with 1000 permutations).

TABLE 5 (next page). List of types and numbers *(italic)* of ITS supratypes, ITS types, *trn*L/F suprahaplotypes, and *trn*L/F haplotypes in the *Arabidopsis lyrata* complex.

Species	ITS supratype	ITS type	*trn*L/F suprahaplotype	*trn*L/F haplotype
Arabidopsis arenicola	*1x:* **ambiguous b/e** *2x:* **b** *13x:* **e**	*1x:* 3, 103, 104, 107 *12x:* 16	*1x:* **AQ** *17x:* **A**	*1x:* 48, 181, 182, 185, 187, 188, 208 *4x:* 31 *7x:* 50
Arabidopsis kamchatica **(incl. *kawasakiana*)**	*1x:* **z** *2x:* **ambiguous b/e** *8x:* **e** *51x:* **b**	*1x:* 18, 108, 109, 113 *2x:* 103 *6x:* 105 *50x:* 3	*1x:* **AT, C** *9x:* **AD** *46x:* **B**	*1x:* 191, 207, 253, 256 *8x:* 89 *45x:* 84
Arabidopsis lyrata **ssp. *lyrata***	*2x:* **y** *35x:* e	*2x:* 110 *4x:* 109 *6x:* 108 *25x:* 16	*1x:* **BF** *12x:* **BD** *41x:* **A**	*1x:* 31, 49, 249, 250 *2x:* 248, 251 *4x:* 245 *6x:* 247, 252 *10x:* 242 *20x:* 50
Arabidopsis lyrata **ssp. *petraea***	*1x:* **c, ambiguous b/e** *2x:* **d** *11x:* **e** *36x:* **a** *52x:* **b**	*1x:* 7, 9, 10, 11, 14, 15, 17, 18, 77, 97, 100, 103 *2x:* 5, 6, 8, 12, 105 *3x:* 4, 13 *5x:* 111 *7x:* 106 *9x:* 2, 16 *13x:* 1 *32x:* 3	*1x:* **AB, AH, AJ, AK, AO, AP, AS, K, S, V** *2x:* **AF, AL, G, J, R** *3x:* **Q** *5x:* **AI** *7x:* **AR** *9x:* **A** *13x:* **B** *19x:* **AG** *41x:* **AC** *142x:* **C**	*1x:* 2, 4, 17, 26, 28, 40, 43, 51, 52, 54, 78, 84, 94, 106, 132, 141, 157, 158, 159, 162, 163, 164, 165, 167, 168, 169, 170, 171, 174, 175, 176, 177, 178, 184, 186, 189, 190, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 204, 205, 206, 229, 245 *2x:* 5, 30, 48, 53, 85, 88, 151, 161, 172 *3x:* 22, 70 *4x:* 81, 87, 146, 160, 203, 255 *5x:* 166 *7x:* 152 *8x:* 11 *17x:* 16 *21x:* 13 *25x:* 1 *75x:* 29

1.3.4. Refugia as areas of secondary contact of formerly allopatric populations: Beringia as an example

Beringia, an arctic region ranging from Lena River in northeast Russia to Mackenzie River in Alaska and from the Arctic Ocean to mountains in southern Siberia and Alaska, is considered the major refugium for arctic taxa, as it remained icefree during Pleistocene climate oscillations. If we consider only the Eurasian and North American lineage of the *Arabidopsis lyrata* complex, two major ITS groups met in Beringia (Fig. 8): the mainly Eurasian group carrying ITS supratype b, comprising Europe (with additional ITS supratypes a, c, d), northern Russia and western Beringia, and the North American group carrying ITS supratype e, including eastern and central North America, Greenland, and eastern Beringia (north of Brooks Range). The main contact zone is located in eastern Beringia, where accessions with *trn*L/F suprahaplotype C, characteristic for the Eurasian lineage, showed ITS supratype e, characteristic for the North American lineage. This is obviously due to ancient and/or recent geneflow from populations of the North American lineage into populations of the Eurasian lineage. Throughout whole Beringia ambiguous sites in ITS DNA sequences, caused by multiple ITS copies within a single genome not subjected to concerted evolution (Koch et al., 2003; Schmickl et al., 2008a), were mainly found between ITS supratypes b (Eurasian lineage) and e (North American lineage), indicating geneflow between these two genetic groups. Interestingly, the allopolyploid amphi-pacific lineage (*Arabidopsis kamchatica*) also showed ITS supratype b like the Eurasian lineage, but no geneflow could be observed between this lineage and the North American lineage. Either the amphi-pacific lineage migrated only recently, during the LGM, into North America via the Bering Land Bridge, or those two lineages are reproductively isolated from each other. However, if we assume not only *Arabidopsis kamchatica* from Japan as a selfer (Sugisaka and Kudoh, 2008), but throughout its complete distribution range, the reproductive system probably had the most severe influence on the genetic isolation of the amphi-pacific from the North American lineage.

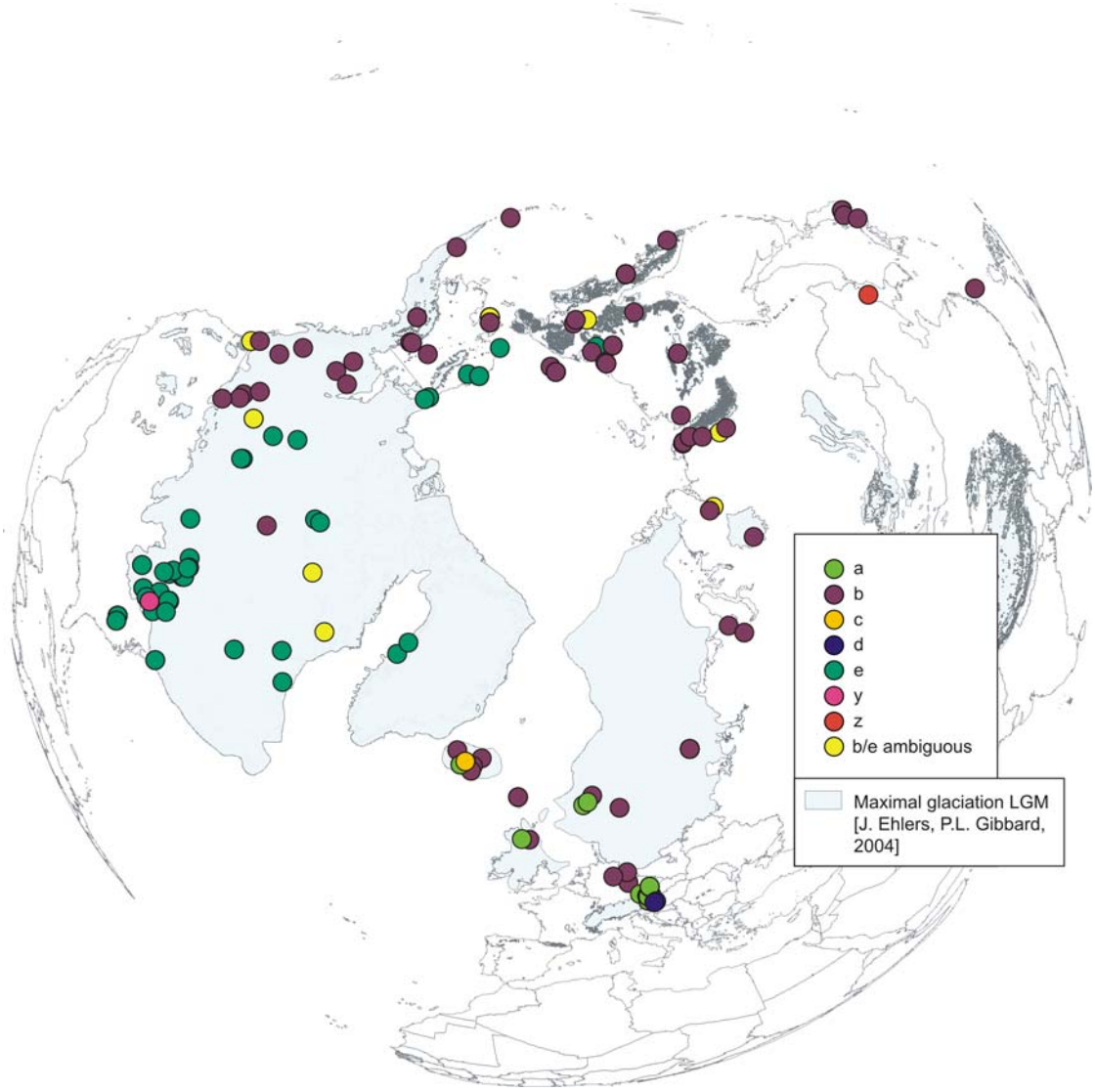


Fig. 8. Distribution of ITS supratypes in the *Arabidopsis lyrata* complex, combining the results obtained in this study with the results of Koch and Matschinger (2007) and Schmickl et al. (2008a). Accessions with sequence ambiguities between supratypes b and e are marked separately. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard (2004).

1.4. Discussion

1.4.1. Eurasia as the centre of genetic diversity of the Arabidopsis lyrata complex - postglacial migration from Central European and northern Russian refuge areas

For arctic-alpine taxa centres of species and genetic diversity are, in most cases, considered to concur with Pleistocene refugia (Abbott and Brochmann, 2003; Tribsch and Schönswetter, 2003; Hewitt, 2004). In the genus *Arabidopsis* both the number of accepted taxa, *trn*L/F suprahaplotype and ITS supratype diversity is highest in Central Europe, indicating its centre of diversity (Al-Shehbaz and O´Kane, 2002; Koch and Matschinger, 2007). However, for the *Arabidopsis lyrata* species complex Central Europe is not the only centre of diversity, as concluded in Koch and Matschinger (2007). With this study it became obvious, that the Eurasian lineage of the *Arabidopsis lyrata* complex is genetically diverse both in Central Europe and arctic Eurasia, including Beringia. Based on molecular marker studies such a pattern was not observed for any other arctic-alpine plant species up to now. Either genetic diversity was increased in the Arctic compared to the Alps, as in *Saxifraga oppositifolia*, suggesting long-term evolution in Beringia and more recent colonisation of the Alps (Abbott et al., 2000; Abbott and Comes, 2004). Or genetic diversity was highest in the Alps and decreasing towards the Arctic, as observed in *Arabis alpina* (Koch et al., 2006; Ehrich et al., 2007), indicating recent and rapid colonisation from Central Europe. Recent migration into the Arctic, associated with a loss of genetic diversity due to repeated bottlenecks, was also found in *Ranunculus glacialis* (Schönswetter, 2003). The overall high genetic diversity in the Eurasian *Arabidopsis lyrata* lineage can probably be explained by long-term glacial survival in multiple refugia (Central Europe, northern Russia, western and eastern Beringia).

So far, periglacial survival of *Arabidopsis lyrata* was only reported from unglaciated Central Europe, based on microsatellite data (Clauss and Mitchell-Olds, 2006) and cpDNA sequence polymorphism (Koch and Matschinger, 2007). Interestingly, in this study genetic diversity was also high in formerly glaciated northern Europe, caused by a high number of unique, locally distributed *trn*L/F suprahaplotypes. A possible explanation might be geographic isolation of populations, which either periglacially survived along the coastline or postglacially migrated into northern Europe. Although Brochmann et al. (2003) suggested periglacial survival for several arctic plant taxa on the Outer Hebrids of Scotland and along the Icelandic and Norwegian coastlines, we assume postglacial colonisation of formerly

glaciated northern Europe from both unglaciated Central Europe and northwestern Russia. Subsequent geographic isolation of populations during Holocene warming probably led to the fixation of local sequence types. Although in our study the major *trn*L/F suprahaplotypes C and AG were similarly frequent both in Central and northern Europe, a comparative microsatellite study between populations from Central Europe and Iceland revealed significant differences in marker polymorphism (Clauss and Mitchell-Olds, 2006). Riihimäki and Savolainen (2004) even found divergent Central and northern European physiological morphotypes, according to a latitudinal gradient from earlier and more frequent flowering in the south to later and rarer flowering in the north. Population differentiation was even strong between divergent habitats of similar latitude: According to Leinonen et al. (2009), alpine populations from Norway and coastal populations from Sweden showed a different adaptive potential.

On the one hand we observed high genetic diversity within the Eurasian *Arabidopsis lyrata* lineage, on the other hand genetic homogenisation was partly observed: The *trn*L/F suprahaplotypes C and AG were found all over its distribution range. This could be explained by repeated geneflow between populations during glacial periods, additionally facilitated by rapid long-distance dispersal of the small *Arabidopsis* seeds across the smooth snow surface of the tundra and tundra-steppe by strong winds. This mode of long-distance dispersal could even have bridged distances up to 2000 km (Savile, 1972). Meanwhile, rapid long-distance dispersal turned out to be more frequent in the Arctic than assumed years ago, responsible for the postglacial colonisation of formerly glaciated islands, such as Svalbard, from mainland founder populations (Alsos et al., 2007).

1.4.2. Ancient split of the Eurasian and North American lineage

The strong genetic differentiation we observed between the Eurasian and North American lineage is probably the result of long-term geographic isolation during Pleistocene glaciations. Our results support the finding of Muller et al. (2008), who detected genetic divergence between North America and western Europe in a comparative microsatellite study, focussing on few populations only and biased due to only one North American population. The strong reduction of genetic diversity we observed in the North American lineage in contrast to the Eurasian lineage is congruent with nuclear and plastidic marker data of Wright et al. (2003), but contradicting Balañá-Alcaide et al. (2006), who reported similar degrees of genetic diversity, based on two nuclear markers. However, both studies included only two populations

each from North America and Europe. Because of the strongly reduced genetic diversity of the North American lineage, we assume this lineage to be derived from the Eurasian lineage. Migration into North America was obviously associated with a strong genetic bottleneck, which restricted populations to one major sequence type, ITS supratype e, *trn*L/F suprahaplotype A, each. Two directions of colonisation of North America are plausible, either trans-atlantic or via the Bering Land Bridge. Both hypotheses have been discussed for various circumboreal species (Abbott and Brochmann, 2003). In the case of *Arabidopsis lyrata* colonisation of North America from Russia via the Bering Land Bridge seems most convenient, as the North American genotype was also rarely detected in western Beringia, but not at all in Iceland or neighbouring Scandinavian countries.

1.4.3. An Amphi-Beringian Arabidopsis hybrid zone – due to allopolyploid success?

The majority of arctic polyploids is found within postglacial colonisers (Brochmann et al., 2004). It has frequently been assumed that polyploids have a broader adaptive potential for recolonising formerly glaciated areas starting with early studies in *Biscutella laevigata* (Manton, 1937). This assumption is indicated by the finding that in Beringia, one of the major refugia for arctic plants during Pleistocene glaciations, a significantly higher number of diploids was reported in contrast to areas under ice cover. New adaptations in polyploids may evolve either by genome rearrangements (Soltis and Soltis, 2000) and/or epigenetic changes (Comai et al., 2000; Wendel, 2000; Liu and Wendel, 2003), even within the first generations after polyploid formation, as reported from rapid gene silencing in the allopolyploid *Arabidopsis suecica* (Comai et al., 2000). In arctic plants tetraploids are the most common polyploids (Brochmann et al., 2004). Frequently, hybridisation is involved in polyploidisation, leading to the formation of allopolyploids with one set of the parental genome each. Allopolyploidisation was reported for several arctic species complexes, such as high polyploid *Cerastium alpinum* (Brysting et al., 2007), high polyploid *Primula* sect. *Aleuritia* (Guggisberg et al., 2006; Guggisberg et al., 2008; Guggisberg et al., 2009), tetra-/hexaploid *Silene* (Popp et al., 2005; Popp and Oxelman, 2007), and tetra- to dodecaploid *Cardamine digitata* (Jørgensen, 2008). Introgression, the stable integration of genetic material from one species into another through repeated backcrossing, was observed between polyploid *Saxifraga cernua* and diploid *Saxifraga sibirica* (Kapralov et al., 2006). With tetraploid *Arabidopsis kamchatica* we detected another arctic allopolyploid and the second natural hybrid within the genus *Arabidopsis*. So far, only allopolyploid *Arabidopsis suecica* with the

maternal parent *Arabidopsis thaliana* and the paternal parent *Arabidopsis arenosa* was known, confirmed by artificial crosses (Comai et al., 2000). This species probably developed around 20000 years ago (Säll et al., 2003) or between 20000 and 300000 years ago (Jakobsson et al., 2006), with a single origin in Fennoscandinavia (Säll et al., 2003; Jakobsson et al., 2006). The distribution range of this mainly outcrossing (Säll et al., 2004) hybrid species is rather small, in contrast to the vast Amphi-Beringian distribution range of the hybrid *Arabidopsis kamchatica*. This large distribution is partly due to postglacial colonisation of formerly glaciated areas of mainly eastern Beringia and western Canada. As genetic diversity was extremely low in these regions, postglacial immigration can be assumed to have been rapid. Was this probably enhanced by de novo adaptations as a result of hybridisation and polyploidisation? The most profound change in *Arabidopsis kamchatica* in contrast to its parental species is the switch from outcrossing (with sporophytic self-incompatibility system) to selfing. Genetically, polyploids can in this way escape genome stabilisation required for meiosis. Ecologically, they are independent on pollinators and, consequently, weather conditions allowing for pollinator visits. Such a switch is already well known from *Arabidopsis thaliana*, dated from around 413000 (Bechsgaard et al., 2006) to one million years ago (Tang et al., 2007), but it is not necessarily correlated with hybridisation and polyploidisation, as *Arabidopsis thaliana* is neither a hybrid species nor a polyploid.

The success of *Arabidopsis kamchatica* as a rapid coloniser was probably enhanced by the availability of large, open landscapes, where glacial and/or permafrost activity had frequently disturbed habitats. However, such habitats were predominantly colonised by plants with low competitiveness. Regarding its ecology, *Arabidopsis kamchatica* is largely separated from both its parental species, preferring coastal arctic instead of continental arctic habitats like *Arabidopsis lyrata* ssp. *petraea* or temperate habitats of high altitudes like *Arabidopsis halleri* ssp. *gemmifera*. We can conclude, that the success of *Arabidopsis kamchatica* as a postglacial coloniser was mainly the result of a change in the mating system. Changes of the mating system are one of the major driving forces for speciation, initiating reproductive isolation of populations. However, it is still unclear, if the switch of mating systems is correlated with hybridisation and polyploidisation. First indications for a correlation could be drawn from the breakdown of self-incompatibility in an artificial cross between *Arabidopsis thaliana* and *Arabidopsis lyrata* (Nasrallah et al., 2007). Otherwise breakdown of the sporophytic self-incompatibility system was, up to now, only reported from mainly diploid individuals (Mable et al., 2005; Mable, 2008).

1.4.4. Beringia as contact zone of the Eurasian and North American lineage of the Arabidopsis lyrata complex

Beringia served as a glacial refugium for numerous arctic plant taxa such as *Dryas integrifolia* (Tremblay and Schoen, 1999), *Saxifraga hirculus* (Oliver et al., 2006), *Saxifraga oppositifolia* (Abbott et al., 2000; Abbott and Comes, 2004), and *Vaccinium uliginosum* (Alsos et al, 2005; Eidesen et al., 2007). Also for the *Arabidopsis lyrata* complex periglacial survival in Beringia can be assumed, in particular for the Eurasian lineage in arctic western and eastern Beringia north of Brooks range. As indicated by nuclear sequence data, geneflow with populations of the North American lineage (*Arabidopsis lyrata* ssp. *lyrata*, *Arabidopsis arenicola*) must have occurred inter- and postglacially. However, glacial survival of the North American lineage in Beringia can be rejected, as no plastidic sequence types of this lineage were found in Beringia.

1.5. LITERATURE CITED

ABBOTT, R. J., L. C. SMITH, R. I. MILNE, R. M. CRAWFORD, K. WOLFF, AND J. BALFOUR. 2000. Molecular analysis of plant migration and refugia in the Arctic. *Science* 289: 1343-1346.

ABBOTT, R. J., AND C. BROCHMANN. 2003. History and evolution of the arctic flora: in the footsteps of Eric Hulten. *Mol Ecol* 12: 299-313.

ABBOTT, R. J., AND H. P. COMES. 2004. Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (Purple saxifrage). *New Phytol* 161: 211-224.

AL-SHEHBAZ, I. A., S. L. O´KANE, JR., AND R. A. PRICE. 1999. Generic placement of species excluded from *Arabidopsis* (Brassicaceae). *Novon* 9: 296-307.

AL-SHEHBAZ, I. A., AND S. L. O´KANE, JR. 2002. Taxonomy and phylogeny of *Arabidopsis* (Brassicaceae), The Arabidopsis Book.

ALSOS, I. G., T. ENGELSKJØN, L. GIELLY, P. TABERLET, AND C. BROCHMANN. 2005. Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Mol Ecol* 14: 2739-2753.

ALSOS, I. G., P. B. EIDESEN, D. EHRICH, I. SKREDE, K. WESTERGAARD, G. H. JACOBSEN, J. Y. LANDVIK, P. TABERLET, AND C. BROCHMANN. 2007. Frequent long-distance plant colonization in the changing Arctic. *Science* 316: 1606-1609.

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB, AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.

BALAÑÁ-ALCAIDE, D., S. E. RAMOS-ONSINS, Q. BOONE, AND M. AGUADÉ. 2006. Highly structured nucleotide variation within and among *Arabidopsis lyrata* populations at the FAH1 and DFR gene regions. *Mol Ecol* 15: 2059-2068.

BEAULIEU, J., M. JEAN, AND F. BELZILE. 2009. The allotetraploid *Arabidopsis thaliana-Arabidopsis lyrata* subsp. *petraea* as an alternative model system for the study of polyploidy in plants. *Mol Genet Genomics* 281: 421-435.

BECHSGAARD, J. S., V. CASTRIC, D. CHARLESWORTH, X. VEKEMANS, AND M. H. SCHIERUP. 2006. The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr. *Mol Biol Evol* 23: 1741-1750.

BECK, J. B., H. SCHMUTHS, AND B. A. SCHAAL. 2008. Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Mol Ecol* 17: 902-915.

BENFEY, P. N., AND T. MITCHELL-OLDS. 2008. From genotype to phenotype: systems biology meets natural variation. *Science* 320: 495-497.

BROCHMANN, C., T. M. GABRIELSEN, I. NORDAL, J. Y. LANDVIK, AND R. ELVEN. 2003. Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon* 52: 417-450.

BROCHMANN, C., A. K. BRYSTING, I. G. ALSOS, L. BORGEN, H. H. GRUNDT, A. C. SCHEEN, AND R. ELVEN. 2004. Polyploidy in arctic plants. *Biol J Linn Soc* 82: 521-536.

BRYSTING, A. K., B. OXELMAN, K. T. HUBER, V. MOULTON, AND C. BROCHMANN. 2007. Untangling complex histories of genome mergings in high polyploids. *Syst Biol* 56: 467-476.

CHARLESWORTH, D., B. K. MABLE, M. H. SCHIERUP, C. BARTOLOMÉ, AND P. AWADALLA. 2003. Diversity and linkage of genes in the self-incompatibility gene family in *Arabidopsis lyrata*. *Genetics* 164: 1519-1535.

CHARLESWORTH, D., E. KAMAU, J. HAGENBLAD, AND C. TANG. 2006. Trans-specificity at loci near the self-incompatibility loci in *Arabidopsis*. *Genetics* 172: 2699-2704.

CLAUSS, M. J., AND M. A. KOCH. 2006. Poorly known relatives of *Arabidopsis thaliana*. *Trends Plant Sci* 11: 449-459.

CLAUSS, M. J., AND T. MITCHELL-OLDS. 2006. Population genetic structure of *Arabidopsis lyrata* in Europe. *Mol Ecol* 15: 2753-2766.

CLAUSS, M. J., S. DIETEL, G. SCHUBERT, AND T. MITCHELL-OLDS. 2006. Glucosinolate and trichome defenses in a natural *Arabidopsis lyrata* population. *J Chem Ecol* 32: 2351-2373.

CLEMENT, M., D. POSADA, AND K. A. CRANDALL. 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1659.

COMAI, L., A. P. TYAGI, K. WINTER, R. HOLMES-DAVIS, S. H. REYNOLDS, Y. STEVENS, AND B. BYERS. 2000. Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids. *Plant Cell* 12: 1551-1568.

DOBEŠ, C. H., T. MITCHELL-OLDS, AND M. A. KOCH. 2004a. Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). *Mol Ecol* 13: 349-370.

DOBEŠ, C., T. MITCHELL-OLDS, AND M. A. KOCH. 2004b. Intraspecific diversification in North American *Boechera stricta* (= *Arabis drummondii*), *Boechera x divaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers – an integrative approach. *American Journal of Botany* 91: 2087-2101.

DOBEŠ, C., C. KIEFER, M. KIEFER, AND M. A. KOCH. 2007. Plastidic *trn*FUUC pseudogenes in North American genus *Boechera* (Brassicaceae): mechanistic aspects of evolution. *Plant Biol (Stuttg)* 9: 502-515.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

EHLERS, J., AND P. L. GIBBARD. 2004. Quaternary glaciations - extent and chronology. Elsevier, Amsterdam.

EHRICH, D., M. GAUDEUL, A. ASSEFA, M. A. KOCH, K. MUMMENHOFF, S. NEMOMISSA, AND C. BROCHMANN. 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Mol Ecol* 16: 2542-2559.

EIDESEN, P. B., I. G. ALSOS, M. POPP, Ø. STENSRUD, J. SUDA, AND C. BROCHMANN. 2007. Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species. *Mol Ecol* 16: 3902-3925.

EIDESEN, P. B., T. CARLSEN, U. MOLAU, AND C. BROCHMANN. 2007. Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic. *J Biogeogr* 34: 1559-1574.

ELVEN, R., D. F. MURRAY, V. Y. RAZZHIVIN, AND B. A. YURTSEV. 2007. Panarctic flora checklist. National Centre for Biosystematics, University of Oslo, Oslo.

EXCOFFIER, L., G. LAVAL, AND S. SCHNEIDER. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.

FORD, V. S., AND L. D. GOTTLIEB. 2002. Single mutations silence PGiC2 genes in two very recent allotetraploid species of *Clarkia*. *Evolution* 56: 699-707.

GHATNEKAR, L. 1999. A polymorphic duplicated locus for cytosolic PGI segregating in sheep's fescue (*Festuca ovina* L.). *Heredity* 83: 451-459.

GOTTLIEB, L. D. 1977. Evidence for duplication and divergence of the structural gene for phosphoglucoisomerase in diploid species of *Clarkia*. *Genetics* 86: 289-308.

GUGGISBERG, A., G. MANSION, S. KELSO, AND E. CONTI. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of Primula. *New Phytol* 171: 617-632.

GUGGISBERG, A., C. BAROUX, U. GROSSNIKLAUS, AND E. CONTI. 2008. Genomic origin and organization of the allopolyploid *Primula egaliksensis* investigated by in situ hybridization. *Ann Bot (Lond)* 101: 919-927.

GUGGISBERG, A., G. MANSION, AND E. CONTI. 2009. Disentangling reticulate evolution in an arctic-alpine polyploid complex. *Syst Biol* 58: 55-73.

HAGENBLAD, J., J. BECHSGAARD, AND D. CHARLESWORTH. 2006. Linkage disequilibrium between incompatibility locus region genes in the plant *Arabidopsis lyrata*. *Genetics* 173: 1057-1073.

HEWITT, G. M. 2004. The structure of biodiversity - insights from molecular phylogeography. *Front Zool* 1: 4.

JAKOBSSON, M., J. HAGENBLAD, S. TAVARÉ, T. SÄLL, C. HALLDÉN, C. LIND-HALLDÉN, AND M. NORDBORG. 2006. A unique recent origin of the allotetraploid species *Arabidopsis suecica*: Evidence from nuclear DNA markers. *Mol Biol Evol* 23: 1217-1231.

JØRGENSEN, M. H., T. CARLSEN, I. SKREDE, AND R. ELVEN. 2008. Microsatellites resolve the taxonomy of the polyploid *Cardamine digitata* aggregate (Brassicaceae). *Taxon* 57: 882-892.

KAPRALOV, M. V., T. M. GABRIELSEN, I. E. SARAPULTSEV, AND C. BROCHMANN. 2006. Genetic enrichment of the arctic clonal plant *Saxifraga cernua* at its southern periphery via the alpine sexual *Saxifraga sibirica*. *Mol Ecol* 15: 3401-3411.

KOCH, M. A., C. DOBEŠ, AND T. MITCHELL-OLDS. 2003. Multiple hybrid formation in natural populations: Concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Mol Biol Evol* 20: 338-350.

KOCH, M. A., C. DOBEŠ, M. MATSCHINGER, W. BLEEKER, J. VOGEL, M. KIEFER, AND T. MITCHELL-OLDS. 2005. Evolution of the *trn*F(GAA) gene in *Arabidopsis* relatives and the brassicaceae family: monophyletic origin and subsequent diversification of a plastidic pseudogene. *Mol Biol Evol* 22: 1032-1043.

KOCH, M. A., C. KIEFER, D. EHRICH, J. VOGEL, C. BROCHMANN, AND K. MUMMENHOFF. 2006. Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Mol Ecol* 15: 825-839.

KOCH, M. A., AND M. MATSCHINGER. 2007. Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 104: 6272-6277.

KOCH, M. A., C. DOBEŠ, C. KIEFER, R. SCHMICKL, L. KLIMEŠ, AND M. A. LYSAK. 2007. Supernetwork identifies multiple events of plastid *trn*F(GAA) pseudogene evolution in the Brassicaceae. *Mol Biol Evol* 24: 63-73.

KOCH, M. A., M. WERNISCH, AND R. SCHMICKL. 2008. *Arabidopsis thaliana´s* wild relatives: an updated overview on systematics, taxonomy and evolution. *Taxon* 57: 933-943.

KUMAR, S., M. NEI, J. DUDLEY, AND K. TAMURA. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9: 299-306.

KUSABA, M., K. DWYER, J. HENDERSHOT, J. VREBALOV, J. B. NASRALLAH, AND M. E. NASRALLAH. 2001. Self-incompatibility in the genus *Arabidopsis*: characterization of the S locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. *Plant Cell* 13: 627-643.

LEINONEN, P. H., S. SANDRING, B. QUILOT, M. J. CLAUSS, T. MITCHELL-OLDS, J. ÅGREN, AND O. SAVOLAINEN. 2009. Local adaptation in European populations of *Arabidopsis lyrata* (Brassicaceae). *Am J Bot* 96: 1129-1137.

LIU, B., AND J. F. WENDEL. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Mol Phylogenet Evol* 29: 365-379.

MABLE, B. K., J. BELAND, AND C. DI BERARDO. 2004. Inheritance and dominance of self-incompatibility alleles in polyploid *Arabidopsis lyrata*. *Heredity* 93: 476-486.

MABLE, B. K., A. V. ROBERTSON, S. DART, C. DI BERARDO, AND L. WITHAM. 2005. Breakdown of self-incompatibility in the perennial *Arabidopsis lyrata* (Brassicaceae) and its genetic consequences. *Evolution* 59: 1437-1448.

MABLE, B. K. 2008. Genetic causes and consequences of the breakdown of self-incompatibility: case studies in the Brassicaceae. *Genet Res* 90: 47-60.

MANTON, I. 1937. The problem of *Biscutella laevigata* L. - II. The evidence from meiosis. *Ann Bot (Lond)* 1: 439-462.

MITCHELL-OLDS, T. 2001. *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends Ecol Evol* 16: 693-700.

MULLER, M. H., J. LEPPÄLÄ, AND O. SAVOLAINEN. 2008. Genome-wide effects of postglacial colonization in *Arabidopsis lyrata*. *Heredity* 100: 47-58.

NASRALLAH, J. B., P. LIU, S. SHERMAN-BROYLES, R. SCHMIDT, AND M. E. NASRALLAH. 2007. Epigenetic mechanisms for breakdown of self-incompatibility in interspecific hybrids. *Genetics* 175: 1965-1973.

OLIVER, C., P. M. HOLLINGSWORTH, AND R. J. GORNALL. 2006. Chloroplast DNA phylogeography of the arctic-montane species *Saxifraga hirculus* (Saxifragaceae). *Heredity* 96: 222-231.

O´KANE, S. L., JR., AND I. A. AL-SHEHBAZ. 1997. A synopsis of *Arabidopsis* (Brassicaceae). *Novon* 7: 323-327.

POPP, M., P. ERIXON, F. EGGENS, AND B. OXELMAN. 2005. Origin and evolution of a circumpolar polyploid species complex in *Silene* (Caryophyllaceae) inferred from low copy nuclear RNA polymerase introns, rDNA, and chloroplast DNA. *Syst Bot* 30: 302-313.

POPP, M., AND B. OXELMAN. 2007. Origin and evolution of North American polyploid *Silene* (Caryophyllaceae). *Am J Bot* 94: 330-349.

RIIHIMÄKI, M., R. PODOLSKY, H. KUITTINEN, H. KOELEWIJN, AND O. SAVOLAINEN. 2005. Studying genetics of adaptive variation in model organisms: flowering time variation in *Arabidopsis lyrata*. *Genetica* 123: 63-74.

RIIHIMÄKI, M., AND O. SAVOLAINEN. 2004. Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). *Am J Bot* 91: 1036-1045.

SAVILE, D. B. O. 1972. Arctic adaptations in plants. *Canada Department of Agriculture Monograph* 6: 1-81.

SÄLL, T., M. JAKOBSSON, C. LIND-HALLDÉN, AND C. HALLDÉN. 2003. Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*. *Journal of Evolutionary Biology* 16: 1019-1029.

SÄLL, T., C. LIND-HALLDÉN, M. JAKOBSSON, AND C. HALLDÉN. 2004. Mode of reproduction in *Arabidopsis suecica*. *Hereditas* 141: 313-317.

SCHIERUP, M. H., B. K. MABLE, P. AWADALLA, AND D. CHARLESWORTH. 2001. Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. *Genetics* 158: 387-399.

SCHIERUP, M. H., J. S. BECHSGAARD, L. H. NIELSEN, AND F. B. CHRISTIANSEN. 2006. Selection at work in self-incompatible *Arabidopsis lyrata*: mating patterns in a natural population. *Genetics* 172: 477-484.

SCHIERUP, M. H., J. S. BECHSGAARD, AND F. B. CHRISTIANSEN. 2008. Selection at work in self-incompatible *Arabidopsis lyrata*. II. Spatial distribution of S haplotypes in Iceland. *Genetics* 180: 1051-1059.

SCHMICKL, R., M. H. JØRGENSEN, A. K. BRYSTING, AND M. A. KOCH. 2008a. Phylogeographic implications for the North American boreal-arctic *Arabidopsis lyrata* complex. *Plant Ecology and Diversity* 1: 245-254.

SCHMICKL, R., C. KIEFER, C. DOBEŠ, AND M. A. KOCH. 2008b. Evolution of *trn*F(GAA) pseudogenes in cruciferous plants. *Plant Systematics and Evolution*.

SCHÖNSWETTER, P., O. PAUN, A. TRIBSCH, AND H. NIKLFELD. 2003. Out of the Alps: colonization of northern Europe by east alpine populations of the glacier buttercup *Ranunculus glacialis* L. (Ranunculaceae). *Molecular Ecology* 12: 3373-3381.

SHARBEL, T. F., B. HAUBOLD, AND T. MITCHELL-OLDS. 2000. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Mol Ecol* 9: 2109-2118.

SHIMIZU, K. K., S. FUJII, K. MARHOLD, K. WATANABE, AND H. KUDOH. 2005. *Arabidopsis kamchatica* (Fisch. ex DC.) K. Shimizu & Kudoh and *A. kamchatica* subsp. *kawasakiana* (Makino) K. Shimizu & Kudoh, new combinations. *Acta Phytotax Geobot* 56: 163-172.

SOLTIS, P. S., AND D. E. SOLTIS. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proc Natl Acad Sci U S A* 97: 7051-7057.

SUGISAKA, J., AND H. KUDOH. 2008. Breeding system of the annual Cruciferae, *Arabidopsis kamchatica* subsp. *kawasakiana*. *J Plant Res* 121: 65-68.

TANG, C., C. TOOMAJIAN, S. SHERMAN-BROYLES, V. PLAGNOL, Y. L. GUO, T. T. HU, R. M. CLARK, J. B. NASRALLAH, D. WEIGEL, AND M. NORDBORG. 2007. The evolution of selfing in *Arabidopsis thaliana*. *Science* 317: 1070-1072.

TREMBLAY, N. O., AND D. J. SCHOEN. 1999. Molecular phylogeography of *Dryas integrifolia*: glacial refugia and postglacial recolonization. *Mol Ecol* 8: 1187-1198.

TRIBSCH, A., AND P. SCHÖNSWETTER. 2003. Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon* 52: 477-497.

TURNER, T. L., E. J. VON WETTBERG, AND S. V. NUZHDIN. 2008. Genomic analysis of differentiation between soil types reveals candidate genes for local adaptation in *Arabidopsis lyrata*. *PLoS One* 3: e3183.

WARWICK, S. I., I. A. AL-SHEHBAZ, AND C. A. SAUDER. 2006. Phylogenetic position of *Arabis arenicola* and generic limits of *Aphragmus* and *Eutrema* (Brassicaceae) based on sequences of nuclear ribosomal DNA. *Can J Bot* 84: 269-281.

WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Mol Biol* 42: 225-249.

WRIGHT, S. I., B. LAUGA, AND D. CHARLESWORTH. 2003. Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. *Mol Ecol* 12: 1247-1263.

2. Autopolyploid formation in *Arabidopsis* and the evolutionary history of the *Arabidopsis arenosa* complex

Abstract

Im zweiten Kapitel „Autopolyploidisierung in *Arabidopsis* und die evolutionäre Historie des *Arabidopsis arenosa* Komplexes“ fanden wir das Zentrum der Artbildung innerhalb des *Arabidopsis arenosa* Komplexes auf der Balkanhalbinsel und in den Karpaten mit vorwiegend diploiden neben vereinzelten tetraploiden Populationen. Polyploidisierung erfolgte durch Selbstverdopplung des Genoms. Sowohl die unvergletscherten ostösterreichischen Alpen als auch die Westkarpaten dienten als pleistozäne Refugialgebiete, was anhand des molekularen Markers cpDNA *trn*L/F ermittelt wurde. Diese beiden Gebirgsstöcke wurden von *Arabidopsis arenosa* einst unabhängig voneinander von der Balkanhalbinsel aus kolonisiert.

In the second chapter “Autopolyploid formation in *Arabidopsis* and the evolutionary history of the *Arabidopsis arenosa* complex” we unravelled the cradle of speciation within the *Arabidopsis arenosa* complex on Balkan Peninsula and in the Carpathians with predominantly diploid, but also tetraploid populations. Self-doubling of the genome was found as the common mode of polyploidisation. Both the unglaciated Eastern Austrian Alps and the Western Carpathians served as Pleistocene refugia, revealed with the molecular marker cpDNA *trn*L/F. These two mountain ranges were formerly colonised independently by *Arabidopsis arenosa* from Balkan Peninsula.

2.1. Introduction

Sympatric speciation within the plant kingdom is frequently correlated with polyploidisation (Soltis and Soltis, 2009), as it can largely contribute to rapid establishment of reproductive isolation. Newly formed polyploids are often stabilised by inbreeding or asexual reproduction, e.g. apomixis, and their environmental establishment is often guaranteed by de novo adaptations (Rieseberg et al., 2003). Approximately 50-70% of all plant species are polyploids (Wendel, 2000), and in higher plants ploidy levels may even vary between (Soltis et al., 2007) and within taxa (Suda et al., 2007). Autopolyploidisation, as seen as the result of self-doubling of the whole genome, is a common mechanism of polyploidisation in higher plants. Although once considered to be rare in contrast to allopolyploidisation, which is the result of interspecific hybridisation and subsequent genome doubling, an increasing number of studies indicates, that autopolyploidisation is a frequent mode of plant speciation (Ramsey and Schemske, 1998; Soltis et al., 2007). Therefore, potential difficulties in meiosis regarding chromosome pairing are probably not that obstructive for the long-term establishment of polyploids, especially as strict autopolyploidy does not reflect the natural situation. Frequently, intermediate stages of auto- and allopolyploid pairing behaviour are reported (Ramsey and Schemske, 2002), or diploidisation of the genome, even shortly after polyploid formation, can be observed. The study of the geographic distribution of cytotypes within a species complex like *Arabidopsis arenosa* is the basic step in understanding the complex evolutionary history of polyploids and can be a prerequisite for phylogeographic studies (Baack, 2004, 2005; Baack and Stanton, 2005).

According to Avise definition of phylogeography (1987), phylogeography interprets genetic variation of taxa in space and time. On the northern hemisphere late quaternary climatic oscillations, especially the last glaciation with its maximum extent, the Last glacial maximum (LGM), about 11500 years ago, had the most severe influence on the distribution and diversity of plant taxa. During the last decade phylogeographic studies were mainly focussing on high alpine taxa, especially from the Alps. Meanwhile there is an increasing number of studies concerning other mountain ranges of the European Mountain System (Schmitt, 2009) like e.g. the Pyrenees (Segarra-Moragues and Catalán, 2008) and the Spanish Sierra Nevada (Gutiérrez Larena et al., 2002; Vargas, 2003; Gutiérrez Larena et al., 2006; Kropf et al., 2006). However, mostly high alpine taxa were studied. In contrast, extensive phylogeographic investigations of the Carpathian flora are still rare (Mráz et al., 2007). In studies of *Dryas octopetala* (Skrede, 2006), *Juncus biglumis* (Schönswetter et al., 2007), and

Vaccinium uliginosum (Alsos, 2005) only few populations from the Carpathians were included, also focussing on alpine and subalpine taxa. However, the IntraBioDiv Consortium recently brought the Carpathian Mountains back into the focus of biodiversity assessment.

The *Arabidopsis arenosa* complex is one of the three major species complexes within the genus *Arabidopsis* (O´Kane and Al-Shehbaz, 1997; Al-Shehbaz et al., 1999; Clauss and Koch, 2006; Koch and Matschinger, 2007; Koch et al., 2008), next to *Arabidopsis lyrata* (L.) Lawalrée and *Arabidopsis halleri* (L.) O'Kane & Al-Shehbaz, formerly treated as *Cardaminopsis* (O´Kane and Al-Shehbaz, 1997; Al-Shehbaz et al., 1999). *Arabidopsis arenosa* is a colline, montane, and subalpine species complex with a mainly Central European distribution range including the Alps and Carpathians. Only few studies have been attempted to unravel the evolutionary history of the *Arabidopsis arenosa* complex (Koch and Matschinger, 2007; Koch et al., 2008). Several studies focussed on the natural hybrid *Arabidopsis suecica*, which is of allopolyploid origin with the maternal parent *Arabidopsis thaliana* and the paternal parent *Arabidopsis arenosa*, confirmed by artificial crosses (Comai et al., 2000). This species probably developed around 20000 years ago (Säll et al., 2003) or between 20000 and 300000 years ago (Jakobsson et al., 2006), with a single origin in Fennoscandinavia (Säll et al., 2003; Jakobsson et al., 2006). Polyploidisation, mainly tetraploidisation, is reported to be frequent within several taxa of the *Arabidopsis arenosa* complex (Mesicek, 1970), indicating repeated independent polyploidisation events. Introgression, the stable integration of genetic material from one species into another through repeated backcrossing, was observed between members of the *Arabidopsis arenosa* and *Arabidopsis lyrata* complex (Schmickl and Koch, unpublished data).

According to different authors, the *Arabidopsis arenosa* complex comprises several taxa on various taxonomic levels. Al-Shehbaz and O'Kane (2002) treat the complex as one species (*Arabidopsis arenosa* (L.) O'Kane and Al-Shehbaz) with two subspecies of partly overlapping distribution ranges in Central Europe: the tetraploid ssp. *arenosa* ($2n = 32$), also occuring in northern Europe, growing mainly on silicious bedrock, and the tetraploid ssp. *borbasii* (Zapal) O'Kane and Al-Shehbaz ($2n = 32$), growing predominantly on calcarous bedrock, additionally found in the Carpathians. Diploid *Arabidopsis neglecta* (Schultes) O´Kane and Al-Shehbaz ($2n = 16$) was described mainly from the Carpathians and rarely from the Alps. Based on morphological and caryological data Mesicek (1970) proposed several additional, mainly diploid taxa of the Carpathians on the species rank, which at that time still were attributed to the genus *Cardaminopsis*. In the Checklist of Non-Vascular and Vascular Plants of Slovakia (http://ibot.sav.sk/page/intro.htm) these species were transferred

to the genus *Arabidopsis*: *Arabidopsis carpatica* nom. prov. (2*n* = 16), *Arabidopsis nitida* nom. prov. (2*n* = 16), *Arabidopsis petrogena* (A. Kern) V.I.Dorof. ssp. *petrogena* (2*n* = 16), and Arabidopsis *petrogena* ssp. *exoleta* nom. prov. (2*n* = 32).

In this study we present the first comprehensive evolutionary history of the *Arabidopsis arenosa* species complex by using a maternally inherited chloroplast genome marker (*trn*L intron (*trn*L) and *trn*L/F intergenic spacer (*trn*L/F-IGS) of $tRNA^{Ser}$ and $tRNA^{Thr}$, respectively). Additionally, cytological data were used to further interpret the phylogeographic pattern observed with molecular markers. The following five aspects are in the focus of our research: (1) Unravelling the main cytogeographic patterns of diploids and tetraploids, and determining the geographic origins of the main speciation events in the *Arabidopsis arenosa* complex. We are particularly interested in contact zones of populations with different or mixed ploidy levels, as they can indicate hotspots of speciation. (2) Evaluating, if auto- or allopolyploidisation led to tetraploid formation, and considering the adaptive potential of polyploids in contrast to diploids. (3) Detecting Pleistocene refugia and postglacial colonisation routes by using genetic diversity statistics: *Arabidopsis arenosa* is both distributed in regions, which remained largely unglaciated during Pleistocene climate oscillations, and areas formerly under ice cover, making it a well suited system to study the effects of glacial cycles. (4) Unravelling, if populations from the Alps and Carpathians are genetically isolated: Genetic isolation may indicate independent colonisation of these two mountain ranges and/or subsequent geographic isolation. (5) Discussing taxonomic treatment with respect to molecular data.

2.2. Material and methods

2.2.1. Plant material

Analyses of 480 accessions of the *Arabidopsis arenosa* complex comprised 264 accessions from Austria, Bosnia, Bulgaria, Croatia, Czech Republic, Denmark, Finland, France, Germany, Hungary, Lithuania, Luxembourg, Norway, Poland, Romania, Slovakia, Slovenia, Sweden, Switzerland, and Ukraine, sequenced by Koch and Matschinger (2007), and 216 accessions from Austria, Belgium, Bosnia, Croatia, Czech Republic, Denmark, England, Finland, France, Germany, Greenland, Hungary, Luxembourg, Romania, Slovakia, Slovenia, and Switzerland, sequenced by Schmickl, Klein and Matschinger. Plant material was mainly collected from herbarium vouchers from BM (Natural History Museum, London), LI (Upper Austrian Provincial Museum, Linz), MO (Missouri Botanical Garden), SAV (Academy of Sciences, Bratislava), W (Natural History Museum, Vienna), WU (Herbarium of Vienna University Botanical Institute), ZT (University of Zurich), and partly collected in the field, documented at HEID (Herbarium Heidelberg). Accessions were taxonomically determined according to the voucher labels, Slovakian accessions were partly verified in collaboration with Martin Kolnic (unpublished data). However, taxon determination was difficult, as no morphological studies based on a large number of accessions, especially of the various Carpathian taxa (*Arabidopsis carpatica* nom. prov., *Arabidopsis nitida* nom. prov., *Arabidopsis petrogena* (A. Kern) V.I.Dorof. ssp. *petrogena*, and *Arabidopsis petrogena* ssp. *exoleta* nom. prov.), are available. The distribution of the investigated accessions is shown in Fig. 1. The accession list is provided as Supplementary material TABLE 2.

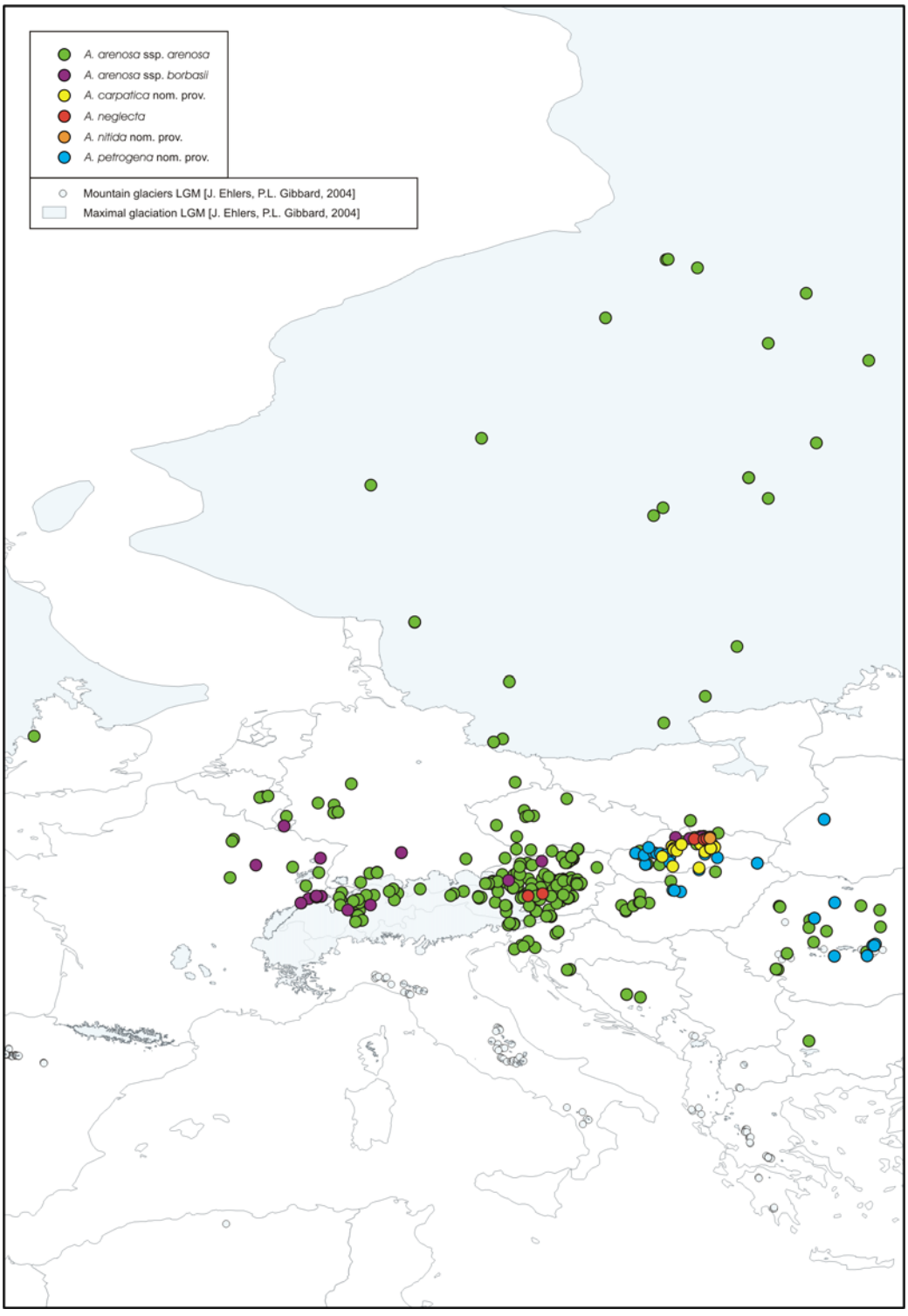


Fig. 1. Distribution of accessions investigated. Taxonomy according to Al-Shehbaz and O´Kane (2002) and the Checklist of Non-Vascular and Vascular Plants of Slovakia (http://ibot.sav.sk/page/intro.htm). Maximal glaciation of the LGM and local mountain glaciers are drawn according to Ehlers and Gibbard (2004).

2.2.2. Mitotic chromosome preparations

Chromosome numbers were determined with classic cytological methods and light microscopy. Flower buds were collected from plants from the field during a collection trip (20.04.-02.05.2005), immediately put in freshly prepared fixative of ethanol – glacial acetic acid (3:1) in 2 ml centrifuge tubes, and incubated at room temperature for 24 h. Long-term storage at -80 °C followed for about 6 months until preparation, according to the protocol of Baldauf (2006). Tubes were first kept at room temperature to warm up. Flower buds were then removed from the fixative and washed with deionised water. Tissue was hydrolysed in 5.0 mol/l hydrochloric acid at room temperature for 30 min as a prerequisite for Schiff staining and to reduce cytoplasmic staining. After washing with deionised water, DNA was stained in a first step with Schiff´s reagent at room temperature for 30 min and once more washed with deionised water. Flower buds were placed on an object slide in a drop of 45% acetic acid to make sure that the buds did not fall dry. The flower buds were preparated with dissecting needles, leaving only the ovary or the petals for further investigations. Remaining tissue debris was removed from the slide. Ovary and petals were crushed between object slide and cover slip using a dissection pin, and immediately frozen at -80 °C on an AlCuMg plate until the acetic acid turned white. The cover slip was then quickly removed and the object dried at room temperature. Giemsa stain was used for a second staining of the chromosomes as 1:10 dilution (Giemsa stain : deionised water) for 4-5 min. Objects were finally washed with deionised water, dried at room temperature, and sealed with Euparal for permanent use. Slides were observed using a Zeiss Axioskop microscope powered by Axiovision Software Package. Pictures were taken with a Zeiss Axiocam Mrc5 digital camera. Chromosome counts were made from the microscopic pictures. 1-5 individuals/population were counted and the counts confirmed by five squashes per slide for each individual. Ploidy level was assigned due to the chromosome base number of eight.

2.2.3. DNA isolation, amplification and sequencing

Total DNA was obtained from dried leaf material and extracted according to the CTAB protocol of Doyle and Doyle (1987) with the following modifications: 50–75 mg of dry leaf tissue were ground in 2 ml tubes using a Retsch swing mill (MM 200), 2 units of ribonuclease per extraction were added to the isolation buffer, and the DNA pellets were washed twice with 70% ethanol. DNA was dissolved in 50 µl TE-buffer for storage and diluted 1:3 in TE-buffer before use. For the cpDNA markers *trn*L intron and *trn*L/F intergenic spacer (*trn*L/F IGS) primer design and PCR cycling scheme followed the protocol of Dobeš et al. (2004), using a PTC200 (MJ Research) thermal cycler. The PCR reaction volume of 50 µl contained 1x PCR buffer (10 mM TRIS/50 mM KCl buffer, pH 8.0), 3 mM $MgCl_2$, 0.4 µM of each primer, 0.2 mM of each dNTP, 1 U Taq DNA polymerase (Amersham), and approximately 1 ng of template DNA. Amplified sequences of *trn*L/F IGS included the complete *trn*L/F IGS and the first 18 bases of the *trn*F gene.

Before sequencing PCR products were checked for length and concentrations on 1.5% agarose gels and purified with the NucleoFast Kit (Macherey & Nagel, Germany). Cycle sequencing was performed using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences) and the original primers. Both strands were amplified in order to gain the complete sequence. PCR products were resolved in 10 µl Loading Solution and run on a MegaBace 500 sequencer.

2.2.4. Plastidic trnL/F sequence definition and map reconstruction

Plastidic *trn*L/F sequences were defined as haplotypes and suprahaplotypes, according to our previous studies (Koch and Matschinger, 2007; Koch et al., 2008; Schmickl et al., 2008a). As the 3´-region of the *trn*L/F IGS close to the functional *trn*F gene is characterised by multiple *trn*F pseudogenes (Koch et al., 2005; Dobeš et al., 2007; Koch and Matschinger, 2007; Koch et al., 2007; Koch et al., 2008; Schmickl et al., 2008b), which evolve with a higher mutation rate than single nucleotide polymorphisms, we excluded the pseudogene region when defining *trn*L/F suprahaplotypes. Hence, haplotypes are defined as *trn*L intron - *trn*L/F-IGS *including* the 3´-end of the *trn*L/F IGS with various copies of *trn*F pseudogenes, suprahaplotypes as *trn*L intron - *trn*L/F-IGS *excluding* the 3´-end of the *trn*L/F-IGS with various copies of *trn*F pseudogenes. Consequently, each *trn*L/F suprahaplotype comprises a varying number of haplotypes, which vary both in length and base content of their pseudogene-rich region.

Maps were constructed using BioOffice version 2.0.6 to create shapefiles and drawn with ArcView version 8.2. Shapefiles for visualising the maximum extent of the ice sheets during the LGM were kindly provided by Jürgen Ehlers and Phil Gibbard (2004).

2.2.5. Network analyses and genetic diversity statistics

Network analyses and genetic diversity statistics were exclusively performed with suprahaplotypes. The *trn*L/F network was constructed using TCS Version (Clement et al., 2000), indels (except polyT stretches) were coded as additional binary characters. Genetic diversity statistics were performed with Arlequin version 3.11 (Excoffier and Schneider, 2005). Pairwise genetic differentiation was calculated among the following seven regional groups: (1) Balkan Peninsula, (2) Southeastern and Western Carpathians, (3) unglaciated Eastern and Southeastern Alps, (4) glaciated Eastern Alps, (5) glaciated Western Alps, (6) unglaciated Central Europe, and (7) glaciated northern Europe. Genetic diversity was estimated as nucleotide diversity π and Nei´s genetic diversity He.

2.3. Results

2.3.1. Chromosome counts identify diploids exclusively on Balkan Peninsula and in the Carpathian Mountains

Two ploidy levels, diploids and tetraploids, were observed throughout the whole dataset (Fig. 2A, 2B, 3). Diploids were exclusively found in Southeastern and Eastern Europe on Balkan Peninsula (Bosnia, Croatia; *Arabidopsis arenosa* ssp. *arenosa*), in northern Hungary (*Arabidopsis arenosa* ssp. *arenosa*), and the Southeastern and Western Carpathian Mountains (Romania, Slovakia; *Arabidopsis arenosa* ssp. *arenosa*; *Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis petrogena* nom. prov.). In those regions the percentage of diploids exceeded 80% (Fig. 2A). In contrast, tetraploids had a large distribution range, spanning from the Julian Alps (Slovenia) in the south and the Western Carpathians (Slovakia) in the east to France and Belgium in the west and Scandinavia in the north (Fig. 3). However, several regions were reported by Meusel et al. (1965) as areas of recent, mainly anthropogenically influenced colonisation after 1890 (Belgium, Finland, France, Great Britain, Greenland), frequently along railway tracks. The Balkan Peninsula and the Western Carpathians were the only regions where both diploid and tetraploid populations were found. Populations of mixed ploidy levels were not observed, but can not be completely excluded, as only a limited number of populations from the Eastern Alps (28) and the Western Carpathians (9) were analysed with more than one individual/population (in general 20 individuals/population).

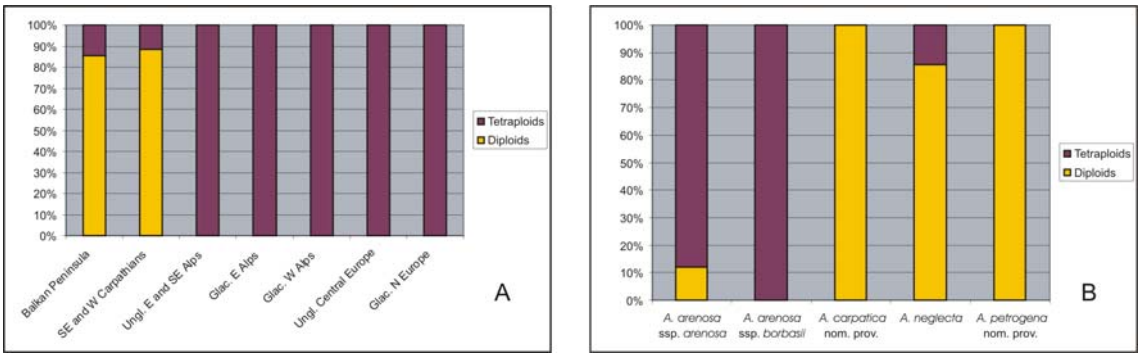
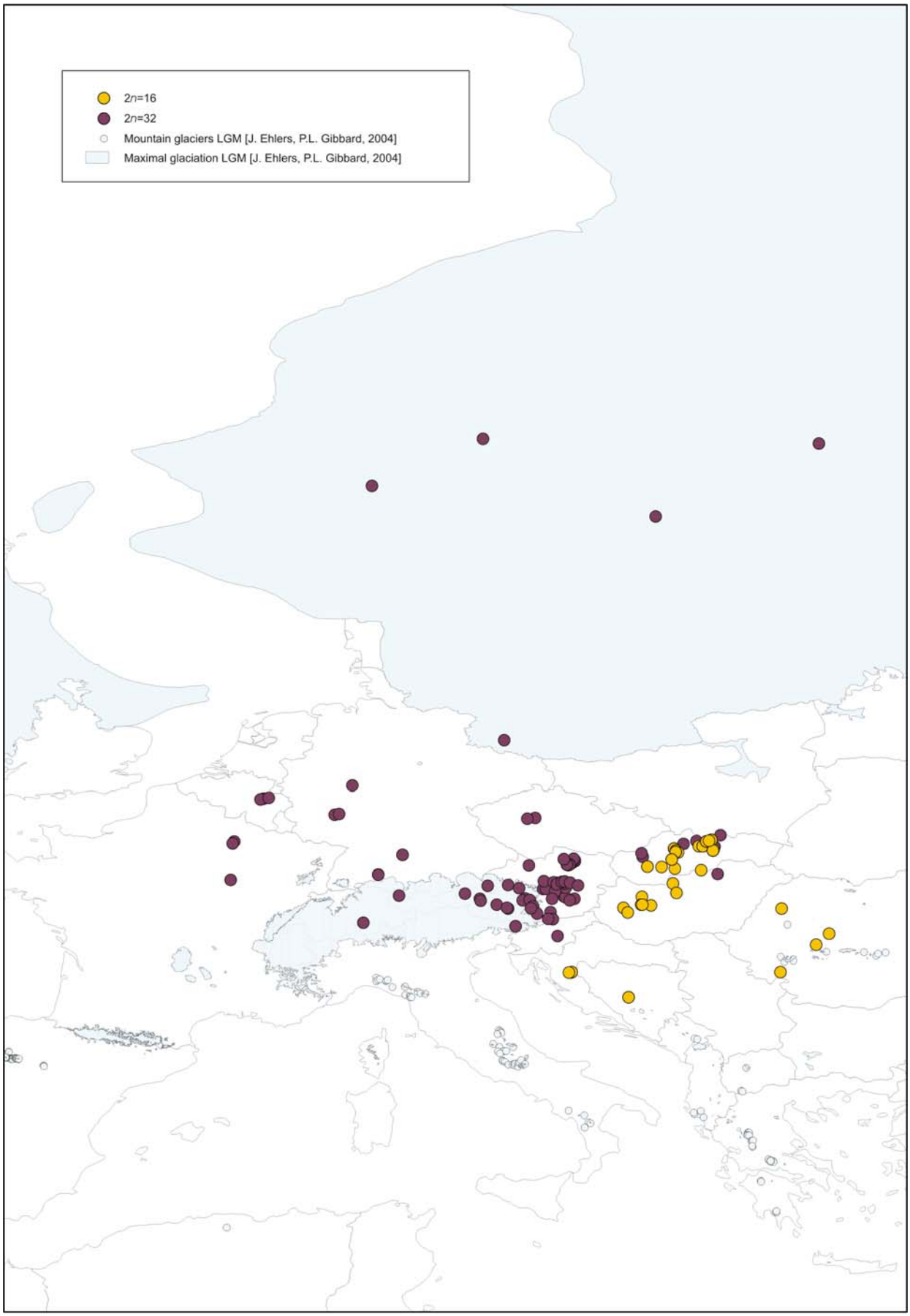


Fig. 2. A. Frequencies of diploid and tetraploid cytotypes in each of the geographic regions, spanning the whole distribution range of the *Arabidopsis arenosa* complex. B. Frequencies of diploid and tetraploid cytotypes in each of the taxonomic units.

Fig. 3 (next page). Distribution of diploid and tetraploid cytotypes within the *Arabidopsis arenosa* complex. Each dot represents 1-20 analysed individuals. Maximal glaciation of the LGM and local mountain glaciers are drawn according to Ehlers and Gibbard (2004).


2n=16
2n=32
Mountain glaciers LGM [J. Ehlers, P.L. Gibbard, 2004]
Maximal glaciation LGM [J. Ehlers, P.L. Gibbard, 2004]

2.3.2. Chloroplast sequence data reveal the genetic relationships between diploids and tetraploids

Diploid accessions mainly showed suprahaplotype A (35%) (TABLE 1), which is a central and the most common suprahaplotype of the *Arabidopsis arenosa* species complex (Fig. 5). Suprahaplotypes are defined here as *trn*L intron - *trn*L/F intergenic spacer *excluding* the 3´-end of the *trn*L/F intergenic spacer with various copies of *trn*F pseudogenes Additionally, central suprahaplotype B (2%) and "tip" suprahaplotype AY (2%, derived from the central suprahaplotype C) were found in diploids from Balkan Peninsula, AV (2%, derived from A) in diploids from the Southeastern Carpathians, and suprahaplotype L (29%, derived from A) in diploids from northern Hungary, the Southeastern and Western Carpathians. Diploids from the Western Carpathians displayed numerous additional suprahaplotypes, both central (C (2%), E (15%)) and "tip" ones (BA (2%), BE (2%), P (2%), U (5%), Y (2%)). Central suprahaplotype A (40%) was also the most frequent one in tetraploid accessions (TABLE 1). Additional suprahaplotypes of tetraploids were the following: B (26%), C (3%), D (3%), E (9%) (central), AU (2%), AV (1%), AW (1%), AX (1%), BB (1%), BC (1%), L (1%), Q (6%), U (4%) ("tip", all derived from A), and AZ (1%) ("tip", derived from E). In a comparison between diploids and tetraploids it became obvious, that approximately 50% of all suprahaplotypes were shared, especially the central (A, B, C, E), but also several "tip" ones (AV, L, U), including the majority of accessions (91% of diploids, 85% of tetraploids). Suprahaplotypes exclusively found in either diploids (AY, BA, BE, P, Y) or tetraploids (D, AU, AW, AX, AZ, BB, BC, Q) were predominantly "tip" suprahaplotypes, usually derived from the *arenosa*-specific suprahaplotype A (Fig. 5), and normally occuring only once. Based on these results we suggest autopolyploidisation within the *Arabidopsis arenosa* species complex as the common mode of polyploidisation. Genetic introgression by taxa of other main genetic lineages of the genus *Arabidopsis*, e.g. *Arabidopsis lyrata*, is probably rare, and can not be detected with *trn*L/F sequence data, as the central suprahaplotypes are shared between the main species complexes of the genus *Arabidopsis* (Koch and Matschinger, 2007). This observation has been explained by ancestral cpDNA polymorphism predating the genus´ radiation approximately two million years ago (Koch and Matschinger, 2007).

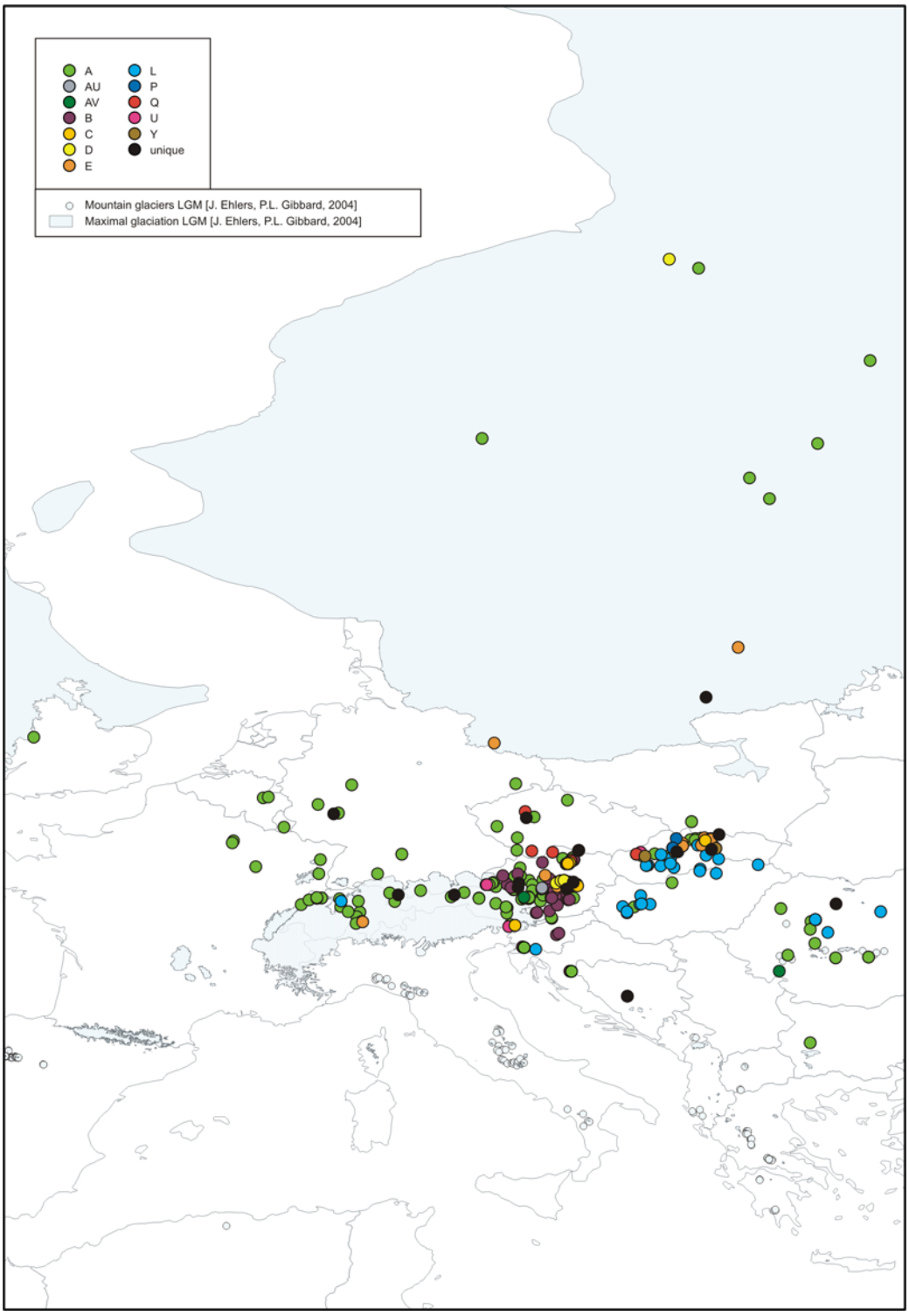


Fig. 4. Distribution of cpDNA *trn*L/F suprahaplotypes in the *Arabidopsis arenosa* complex, combining the results obtained in this study with the results of Koch and Matschinger (2007). Maximal glaciation of the LGM and local mountain glaciers are drawn according to Ehlers and Gibbard (2004).

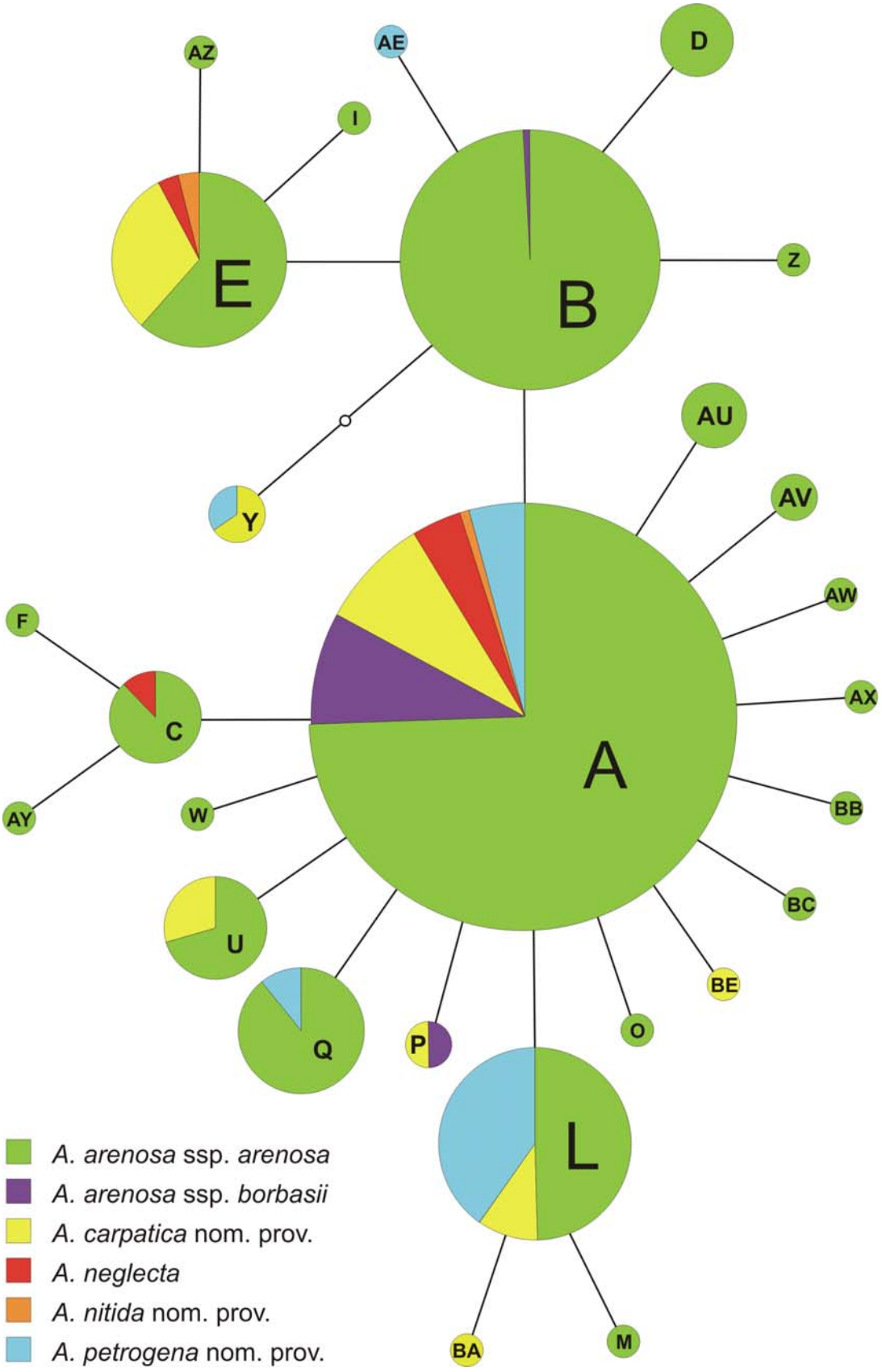


Fig. 5. CpDNA *trn*L/F suprahaplotype network of the *Arabidopsis arenosa* complex, combining the results obtained in this study with the results of Koch and Matschinger (2007). The sizes of the circles indicate the relative frequency of a suprahaplotype.

Ploidy level	*trn*L/F suprahaplotype
Diploids	*1x:* **AV, AY, B, BA, BE, C, P, Y** *3x:* **U** *8x:* **E** *16x:* **L** *19x:* **A**
Tetraploids	*1x:* **AV, AW, AX, AZ, BB, BC** *2x:* **L** *3x:* **AU** *4x:* **D** *5x:* **C** *6x:* **U** *8x:* **Q** *12x:* **E** *35x:* **B** *55x:* **A**

TABLE 1. List of types and numbers *(italic)* of *trn*L/F suprahaplotypes in diploid and tetraploid cytotypes of the *Arabidopsis arenosa* complex.

2.3.3. Balkan Peninsula, the Southeastern and Western Carpathians and the unglaciated Eastern and Southeastern Alps are centres of chloroplast sequence diversity of the Arabidopsis arenosa species complex

Based on *trn*L/F sequence data we detected three centres of genetic diversity within the *Arabidopsis arenosa* species complex (TABLE 2). The first centre of genetic diversity was the Balkan Peninsula with a nucleotide diversity of $\pi = 0.0040$. Within 10 accessions studied, 4 different suprahaplotypes and 8 haplotypes (defined here as *trn*L intron - *trn*L/F intergenic spacer *including* the 3´-end of the *trn*L/F intergenic spacer with various copies of *trn*F pseudogenes) were found (TABLE 3). As the number of mutational steps was extremely low within the *Arabidopsis arenosa* complex, and the majority of suprahaplotypes was derived from the central suprahaplotype A (Fig. 5), we also used the measurement of genetic diversity after Nei (1987), using suprahaplotype frequencies only. Nei´s genetic diversity was He = 0.7333 on Balkan Peninsula (TABLE 2). As the second centre of genetic diversity we detected the Southeastern and Western Carpathians. Nucleotide and genetic diversity after Nei were similarly high as for Balkan Peninsula ($\pi = 0.0035$; He = 0.7191) (TABLE 2), with 13 suprahaplotypes and 42 haplotypes within 107 investigated accessions (TABLE 3). In the Carpathians, especially the Western Carpathians, high genetic diversity was additionally

correlated with species´ diversity. Accessions of numerous taxonomic ranks, although mainly provisional after Mesicek (1970) and the Checklist of Non-Vascular and Vascular Plants of Slovakia (http://ibot.sav.sk/page/intro.htm), were described: *Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis nitida* nom. prov., and *Arabidopsis petrogena* nom. prov.. *Arabidopsis neglecta* is characterised by a different ecological adaptation in comparison with the other members of the *Arabidopsis arenosa* complex. It grows in subalpine-alpine habitats on mountain slopes and along creeks at the forest margin. *Arabidopsis petrogena* nom. prov. underwent a switch of life cycle from biennial to winter annual (personal observation). The third centre of genetic diversity, the unglaciated Eastern and Southeastern Alps, displayed both the highest nucleotide diversity ($\pi = 0.0048$) and Nei´s genetic diversity (He = 0.7643) (TABLE 2). Within 131 accessions 12 different suprahaplotypes and 52 haplotypes were found (TABLE 3). In contrast to these three regions of high genetic diversity, which remained largely unglaciated during Pleistocene climate oscillations except for few local mountain glaciers in the Carpathians, formerly glaciated regions showed reduced levels of genetic diversity. The part of the Eastern Alps formerly under ice cover was characterised by lower nucleotide diversity ($\pi = 0.0031$), but relatively high genetic diversity after Nei (He = 0.5608) (TABLE 2), also indicated by 10 suprahaplotypes and 22 haplotypes within 51 studied accessions (TABLE 3). Consequently, postglacial colonisation can be assumed from the unglaciated Eastern and Southeastern Alps, one of the centres of genetic diversity of the *Arabidopsis arenosa* complex. The formerly glaciated Western Alps showed even more reduced genetic diversity ($\pi = 0.0011$; He = 0.2047) (TABLE 2). Here only 3 suprahaplotypes and 8 haplotypes were detected within 19 analysed individuals (TABLE 3). Glaciated northern Europe displayed a lower value of nucleotide diversity ($\pi = 0.0031$) (TABLE 2). However, several different *trn*L/F sequence types (4 suprahaplotypes and 6 haplotypes within 14 investigated accessions, TABLE 3) were found, resulting in increased Nei´s genetic diversity (He = 0.4909) (TABLE 2). Although Central Europe remained largely unglaciated during Pleistocene climate oscillations, genetic diversity was reduced ($\pi = 0.0024$, He = 0.4034, 6 suprahaplotypes; TABLE 2, 3), in contrast to Balkan Peninsula, the Southeastern/Western Carpathians and the unglaciated Eastern/Southeastern Alps. However, the number of haplotypes was high: 19 haplotypes within 35 studied accessions (TABLE 3) could indicate, that there was glacial survival in unglaciated Central Europe.

Geographic region	*n*	π	Nei´s genetic diversity He
Balkan Peninsula	10	0.003952 +/- 0.002536	0.7333 +/- 0.1005
SE and W Carpathians	107	0.003592 +/- 0.002125	0.7191 +/- 0.0298
Unglac. E and SE Alps	132	0.004831 +/- 0.002713	0.7643 +/- 0.0239
Glac. E Alps	51	0.003101 +/- 0.001898	0.5608 +/- 0.0759
Glac. W Alps	19	0.001085 +/- 0.000901	0.2047 +/- 0.1191
Unglac. Central Europe	35	0.002414 +/- 0.001572	0.4034 +/- 0.1020
Glac. N Europe	11	0.003147 +/- 0.002092	0.4909 +/- 0.1754

TABLE 2. Regional genetic differentiation based on cpDNA suprahaplotypes [sample size (n), nucleotide diversity π +/- standard deviation, and genetic diversity after Nei (1987) He +/- standard deviation].

Taxon	*trn*L/F suprahaplotype	*trn*L/F haplotype
Balkan Peninsula	*1x:* **AY, L** *4x:* **A, B**	*1x:* 83, 86, 98, 111, 219, 237, 246 *3x:* 33
SE and W Carpathians	*1x:* **AE, AV, AZ, BA, BE, C** *2x:* **P, Q** *3x:* **Y** *4x:* **U** *13x:* **E** *31x:* **L** *46x:* **A**	*1x:* 11, 34, 39, 43, 45, 108, 145, 213, 216, 217, 218, 221, 224, 226, 231, 232, 233, 234, 238, 239 *2x:* 15, 19, 36, 44, 62, 72, 113, 222, 225 *3x:* 20, 33, 65, 95, 107, *4x:* 23, 57 *5x:* 18, 59 *7x:* 8, 17, 110 *15x:* 109
Ungl. E and SE Alps	*1x:* **F, O, Z** *2x:* **L** *3x:* **AU** *4x:* **D, U** *5x:* **C** *11x:* **Q** *13x:* **E** *38x:* **A** *49x:* **B**	*1x:* 9, 11, 13, 21, 23, 27, 29, 34, 40, 41, 64, 65, 69, 79, 91, 92, 99, 100, 104, 105, 133, 135, 138, 153, 155, 156, 223, 228, 229 *2x:* 10, 33, 61, 76, 78, 83, 102, 109,146, 150, 154, 210, 230 *3x:* 67 *4x:* 57, 116 *6x:* 59, 77 *7x:* 17 *9x:* 70, 71 *14x:* 96 *15x:* 8
Glac. E Alps	*1x:* **AU, AV, AW, AX, I, M** *2x:* **C, U** *8x:* **B** *33x:* **A**	*1x:* 18, 19, 41, 42, 56, 91, 112, 136, 137, 210, 227, 234, 235, 236 *2x:* 57, 77, 96, 101, 211 *6x:* 33 *10x:* 17 *11x:* 8
Glac. W Alps	*1x:* **E, L** *17x:* **A**	*1x:* 17, 32, 33, 35, 65, 107 *2x:* 18 *11x:* 8
Ungl. Central Europe	*1x:* **BB, BC, E** *2x:* **Q** *3x:* **B** *27x:* **A**	*1x:* 33, 37, 38, 65, 70, 73, 75, 76, 78, 82, 96, 136, 209, 215, 240, 241 *2x:* 23 *3x:* 18 *14x:* 8
Glac. N Europe	*1x:* **D, E, W** *8x:* **A**	*1x:* 66, 114, 140, 214, 254 *6x:* 8

TABLE 3. List of types and numbers *(italic)* of *trn*L/F suprahaplotypes and *trn*L/F haplotypes in each of the geographic regions of the *Arabidopsis arenosa* complex.

2.3.4. Chloroplast sequence data indicate a strong genetic differentiation between the Southeastern/Western Carpathians and the unglaciated Eastern/Southeastern Alps

Populations from the Southeastern/Western Carpathians and the unglaciated Eastern and Southeastern Alps showed significant genetic differentiation, based on chloroplast sequence data (TABLE 3, Fig. 4). The two regions could be characterised by the occurence of two regional suprahaplotypes: Suprahaplotype L was nearly exclusively found in the Southeastern/Western Carpathians (31x). In the unglaciated Eastern and Southeastern Alps this suprahaplotype occurred only twice. In contrast, suprahaplotype B was frequently observed in the unglaciated Eastern and Southeastern Alps (49x), but no occurrence was reported from the Carpathians. Other rarer suprahaplotypes were also unique to either the Southeastern/Western Carpathians (AE, AV, AZ, BA, BE, P, Y) or the unglaciated Eastern/Southeastern Alps (AU, D, F, O, Z). The only frequently shared *trn*L/F suprahaplotypes were: A (SE/W Carpathians: 46x, Ungl. E/SE Alps: 38x), E (SE/W Carpathians: 13x, Ungl. E/SE Alps: 13x), and U (SE/W Carpathians: 4x, Ungl. E/SE Alps: 4x).

2.3.5. Comment on taxonomic units with respect to chloroplast sequence data

According to molecular data, taxonomic assignment was impossible, as most central suprahaplotypes were shared between various taxa (TABLE 4, Fig. 5): A (*Arabidopsis arenosa* ssp. *arenosa*, *Arabidopsis arenosa* ssp. *borbasii*, *Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis nitida* nom. prov., *Arabidopsis petrogena* nom. prov.), B (*Arabidopsis arenosa* ssp. *arenosa*, *Arabidopsis arenosa* ssp. *borbasii*), C (*Arabidopsis arenosa* ssp. *arenosa*, *Arabidopsis neglecta*), and E (*Arabidopsis arenosa* ssp. *arenosa*, *Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis nitida* nom. prov.). Several "tip" suprahaplotypes (L, P, Q, U, Y) were also shared. "Tip" suprahaplotypes AE (*Arabidopsis petrogena* nom. prov.), AU (*Arabidopsis arenosa* ssp. *arenosa*), AV (*Arabidopsis arenosa* ssp. *arenosa*), AW (*Arabidopsis arenosa* ssp. *arenosa*), AX (*Arabidopsis arenosa* ssp. *arenosa*), AY (*Arabidopsis arenosa* ssp. *arenosa*), AZ (*Arabidopsis arenosa* ssp. *arenosa*), BA (*Arabidopsis carpatica* nom. prov.), BB (*Arabidopsis arenosa* ssp. *arenosa*), BC (*Arabidopsis arenosa* ssp. *arenosa*), BE (*Arabidopsis carpatica* nom. prov.), M (*Arabidopsis arenosa* ssp. *arenosa*), O (*Arabidopsis arenosa* ssp.

arenosa), W (*Arabidopsis arenosa* ssp. *arenosa*), and Z (*Arabidopsis arenosa* ssp. *arenosa*) were taxon-specific, mainly found only in one accession. However, additional analyses with eight microsatellite markers revealed a distinct genetic cluster of diploid *Arabidopsis carpatica* nom. prov. in contrast to tetraploid *Arabidopsis arenosa* ssp. *arenosa* from Eastern Austria. Based on the microsatellite markers, *Arabidopsis neglecta* and *Arabidopsis petrogena* nom. prov. also showed up as distinct genetic clusters (this PhD thesis, chapter 3), indicating, that at least for the latter two taxa species´ rank is justified.

Taxon	*trn*L/F suprahaplotype	*trn*L/F haplotype
Arabidopsis arenosa	*1x:* **AW, AX, AY, AZ, BB, BC, F, I, M, O, W, Z** *2x:* **AV** *4x:* **AU** *5x:* **D** *7x:* **C, U** *13x:* **Q** *17x:* **L** *18x:* **E** *63x:* **B** *131x:* **A**	*1x:* 9, 11, 13, 15, 21, 27, 29, 37, 39, 40, 42, 56, 62, 64, 66, 69, 73, 75, 79, 82, 86, 92, 98, 99, 100, 104, 105, 111, 112, 114, 133, 135, 137, 138, 140, 153, 155, 156, 209, 214, 215, 216, 217, 218, 219, 223, 227, 228, 229, 232, 233, 235, 236, 237, 238, 240, 241, 246 *2x:* 10, 34, 41, 61, 76, 78, 91, 101, 102, 107, 110, 136, 146, 150, 154, 211, 230, 234, 254 *3x:* 65, 67, 83, 210 *4x:* 116 *5x:* 23 *7x:* 57, 59 *8x:* 77 *9x:* 18, 71 *10x:* 70, 109 *13x:* 33 *17x:* 17, 96 *55x:* 8
Arabidopsis arenosa **ssp. *borbasii***	*1x:* **B, P** *12x:* **A**	*1x:* 15, 32, 33, 35, 36, 38, 76, 78 *6x:* 8
Arabidopsis carpatica **nom. prov.**	*1x:* **BA, BE, P** *2x:* **Y** *3x:* **U** *4x:* **L** *9x:* **E** *16x:* **A**	*1x:* 36, 44, 62, 107, 108, 221, 224, 226, 239 *2x:* 8, 19, 23, 65, 95, 109, 222, 225 *3x:* 57, 59 *6x:* 17
Arabidopsis neglecta	*1x:* **C, E** *7x:* **A**	*1x:* 8, 11, 19, 65, 213 *2x:* 17, 33
Arabidopsis nitida **nom. prov.**	*1x:* **E** *2x:* **A**	*1x:* 44, 45, 59
Arabidopsis petrogena **nom. prov.**	*1x:* **AE, Y** *2x:* **Q** *6x:* **A** *14x:* **L**	*1x:* 43, 95, 107, 145, 231 *2x:* 18, 72, 113 *3x:* 20 *5x:* 109, 110

TABLE 4. List of types and numbers *(italic)* of *trn*L/F suprahaplotypes and *trn*L/F haplotypes in each of the taxonomic units of the *Arabidopsis arenosa* complex.

2.4. Discussion

2.4.1. Balkan Peninsula and the Carpathian Mountains as the cradle of speciation within the Arabidopsis arenosa complex

Within the *Arabidopsis arenosa* complex two ploidy levels were observed, diploids and tetraploids. As contact zones of diploids and tetraploids were localised in the northern part of Balkan Peninsula and in the Western Carpathians, at least two independent polyploidisation events in each of these regions can be assumed. We observed a clear boundary of the distribution of diploids in the eastern part of the Western Carpathians and tetraploids in the western part of the Western Carpathians. This stands in contrast to the mosaic pattern of diploids and tetraploids in the Western Carpathians, which was found by Mesicek (1970), based on chromosome counts (see also the Karyological database of ferns and flowering plants of Slovakia, http://www.chromosomes.sav.sk). Hence, this discrepancy is either due to a sampling bias in our study or single, incorrect measurements by Mesicek (1970).

In the contact zones of diploids and tetraploids on Balkan Peninsula and in the Western Carpathians we found no populations with mixed ploidy levels, which is rather the exception in plant species of various ploidy levels. In populations of *Potentilla argentea* from Sweden diploids and hexaploids were found (Holm, 1995), and in populations of *Potentilla alpicola* from Switzerland even tetra-, penta-, hexa- and heptaploids (Dobeš and Paule, personal communication). The South American grass *Paspalum simplex* displayed a mixture of tetra- and hexaploid populations (Urbani et al., 2002). However, all studies mentioned so far concentrated on taxa with asexual reproduction in terms of apomixis. Apomictic taxa are normally independent on sexual processes, and, therefore, pollen flow between individuals of divergent ploidy levels is not involved in reproduction at all. Once originated within a population or dispersed from another locality, the apomictic polyploid can reproduce asexually within this population. In studies of sexually reproducing, outcrossing plants populations of mixed ploidy levels are not that frequent. Within the genus *Melampodium* intrapopulational cytotype mixture was reported in particular for the white-rayed species complex with mainly diploids and tetraploids, but was emphasised by the authors as rare cases (Stuessy et al, 2004). In the *Knautia arvensis* agg. both diploids and tetraploids or tetraploids and hexaploids occured within several populations (Kolář et al., 2009). These few studies reflect either populations, that recently underwent polyploidisation, as in *Melampodium*,

where the authors assumed postglacial formation of polyploids via autopolyploidisation. Or populations of mixed ploidy levels are the result of secondary contact of formerly allopatric populations with different ploidy levels, as in the *Knautia arvensis* agg.. Therefore, we could conclude, that the lack of mixed ploidy populations and the lack of uneven ploidy and aneuploidy in the *Arabidopsis arenosa* complex excludes recent polyploidisation events as well as secondary contact zones. Hence, we assume ancient polyploidisation, probably dated several glacial cycles ago. However, slightly different ecological adaptations of diploids and tetraploids might also have caused such a distinct pattern of diploid and tetraploid populations, similar to *Senecio carniolicus* with an altitudinal, ecological gradient of mainly diploid and hexaploid populations in the Eastern Alps (Schönswetter et al., 2007). But as no ecological studies are currently available, this hypothesis remains to be tested.

2.4.2. Tetraploidisation via autopolyploidisation

We assume, that within the originally diploid *Arabidopsis arenosa* species complex autopolyploidisation led to the formation of tetraploids, which nowadays account for the majority of populations. Our assumption is based on the finding of similar frequences of *trn*L/F sequence data between diploids and tetraploids (approximately 88% congruence). Single divergent sequence types, characteristic either for diploids or tetraploids, obviously arose by one mutational step from the mainly *arenosa*-specific suprahaplotype A. Hence, we do not see introgression of members of other *Arabidopsis* lineages (*Arabidopsis halleri*, *Arabidopsis lyrata*) into *Arabidopsis arenosa*. Nevertheless, partly introgressed *Arabidopsis arenosa* populations may exist, but the *trn*L/F sequence marker is certainly not suited to detect this, as the most common central suprahaplotypes are shared between *Arabidopsis arenosa*, *Arabidopsis halleri* and *Arabidopsis lyrata* (Koch and Matschinger, 2007). Additional evidence for autopolyploidisation in contrast to allopolyploidisation arose from the lack of intermediate cytotypes, especially triploids, which are frequently reported in allopolyploid formation (Husband, 2000; Ramsey and Schemske, 2002).

2.4.3. Long-term evolution in three glacial refugia: Balkan Peninsula, the Southeastern/Western Carpathians and the unglaciated Eastern/Southeastern Alps

In numerous studies of both plant and animal species three classical LGM refugia were reported, based on fossil record (Willis, 1996) and species´ and genetic diversity (Petit et al., 2002): the Balkan Peninsula, the Iberian Peninsula, and the Appenin. Out of these three Pleistocene refugia the Balkan Peninsula was emphasized as the most important refugium, especially for tree species (Heuertz et al., 2004), but also for upper and lower montane taxa of especially eastern European distribution (Hewitt, 1999, 2000). The Balkan Peninsula probably also served as the major refuge area for the *Arabidopsis arenosa* complex. Except for one accession in the northern part (Slovenia), only diploids were found there, which is the original cytotype within the *Arabidopsis arenosa* complex as well as the whole genus *Arabidopsis*. Based on *trn*L/F sequence data we detected both suprahaplotype B, characteristic for the unglaciated Eastern/Southeastern Alps, and suprahaplotype L, characteristic for the Southeastern/Western Carpathians, on Balkan Peninsula, indicating colonisation from this refuge area. As genetic diversity is similarly high in the Southeastern/Western Carpathians and the unglaciated Eastern/Southeastern Alps, immigration into these regions probably occurred several glacial cycles ago.

Although the Carpathians remained mainly unglaciated during Pleistocene climate oscillations, except for the southern part, they were reported for only few plant species as refugium so far, including temperate trees (Petit, 2002; Magri, 2006) and herbs, e.g. *Campanula alpina* (Ronikier et al., 2008). In contrast, the unglaciated eastern Austrian Alps have long been assumed as glacial refugium. Along the eastern border of the Austrian Alps a cryptic refugium for tree species was already suggested by Frenzel (1964), according to palynological data, and supported by phylogeographic analyses of Birks and Willis (2008). Additionally, the northeastern Calcarous Alps were already desribed by Pawłowski (1970) to be rich in subalpine endemics, Tribsch and Schönswetter (2003) supported that finding. Based on species´ diversity and the number of endemics they suggested refugia for numerous calcicolous and silicolous plants. As *Arabidopsis arenosa* is both a colline, montane, and subalpine species complex, only parts of the subalpine refugia of the northeastern Calcarous Alps, suggested by Tribsch and Schönswetter (2003), overlapped with the *Arabidopsis arenosa* refuge area of the unglaciated Eastern/Southeastern Alps: Ybbstaler Alpen, Türnitzer Alpen, Gutensteiner Alpen, Rax-Schneeberg, Mürzsteger Alpen, Hochschwab, Ennstaler Alpen, eastern Totes Gebirge, calcareous parts of Grazer Bergland. We describe additional

colline and montane regions of high genetic diversity for the *Arabidopsis arenosa* complex in the Wachau (Danube Valley) and the Krems and Kamp Valley, suggesting additional glacial refugia for colline and montane taxa. This is in congruence with Niklfeld (1972) and Zimmermann (1972), who also considered the eastern edge of the northeastern Calcarous Alps as areas of periglacial survival for montane plant species.

Postglacial colonisation of regions formerly under ice cover must have occurred from one or several of these glacial refugia. However, as accessions from formerly glaciated areas were exclusively tetraploid, colonisation probably started from tetraploid populations of either the Southeastern/Western Carpathians or the unglaciated Eastern/Southeastern Alps. As genetic diversity was still quite high in the formerly glaciated Eastern Alps and numerous chloroplast sequence types from the unglaciated part were found, multiple postglacial immigration from the unglaciated Eastern/Southeastern Alps can be assumed for the formerly glaciated Eastern Alps. Local periglacial survival in Central Europe can not be excluded, as genetic diversity is not extremely low in contrast to the three glacial refugia. Numerous *Arabidopsis arenosa* populations in Central Europe are restricted to relict habitats on exposed rocks in low mountain ranges of the Black Forest, the Eifel, the Elbe Sandstone Mountains, the Harz Mountains, and the Swabian Mountains, often cooccurring with other Pleistocene relict species such as *Dianthus gratianapolitanus* (Marcus Koch, personal observation).

2.4.4. Long-term isolation of the Carpathians and the Alps

According to *trn*L/F sequence data the unglaciated Eastern/Southeastern Alps and the Southeastern/Western Carpathians formed two distinct genetic groups, the Alps characterised by suprahaplotype B, and the Carpathians characterised by suprahaplotype L. Although other suprahaplotypes, e.g. A and E, were shared between these two mountain ranges, we assume strong barriers to seed dispersal and geneflow between the Alps and Carpathians, obviously due to the Pannonian Basin, which constitutes a lowland barrier for mainly montane to subalpine taxa since Holocene warming. But is range fragmentation really the reason for genetic differentiation between *Arabidopsis arenosa* populations from the Alps and Carpathians? Based on our karyological data, we already suggested colonisation of both mountain ranges from the Balkan Peninsula. Based on our chloroplast sequence data, this assumption could be confirmed, as both suprahaplotypes B and L were found in this region. A closer look at the geographic distribution of these two *trn*L/F sequence types unraveled, that suprahaplotype B was located in the northwest and suprahaplotype L in the northeast of

Balkan Peninsula. Therefore, the ancient split into a western (Alpine) and eastern (Carpathian) genetic lineage probably already occurred on northern Balkan Peninsula, and subsequent geneflow between these two lineages, e.g. via long-distance dispersal, was prevented by the large geographic and climatic barrier of the Pannonian Bassin. Long-term genetic isolation between the Alps and Carpathians was also assumed for several other plant species, like *Campanula alpina*, based on AFLP and chloroplast sequence data (Ronikier et al., 2008). Additionally, *Ranunculus glacialis* (Schönswetter et al., 2003) and *Rosa pendulina* (Fér et al., 2007) showed a similar pattern.

2.5. LITERATURE CITED

AL-SHEHBAZ, I. A., AND S. L. O´KANE, JR. 2002. Taxonomy and phylogeny of *Arabidopsis* (Brassicaceae), The Arabidopsis Book.

AL-SHEHBAZ, I. A., S. L. O´KANE, JR., AND R. A. PRICE. 1999. Generic placement of species excluded from *Arabidopsis* (Brassicaceae). *Novon* 9: 296-307.

ALSOS, I. G., T. ENGELSKJØN, L. GIELLY, P. TABERLET, AND C. BROCHMANN. 2005. Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Mol Ecol* 14: 2739-2753.

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB, AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Ann Rev Ecol Syst* 18: 489-522.

BAACK, E. J. 2004. Cytotype segregation on regional and microgeographic scales in snow buttercups (*Ranunculus adoneus*: Ranunculaceae). *Am J Bot* 91: 1783-1788.

BAACK, E. J. 2005. Ecological factors influencing tetraploid establishment in snow buttercups (*Ranunculus adoneus*, Ranunculaceae): Minority cytotype exclusion and barriers to triploid formation. *Am J Bot* 92: 1827-1835.

BAACK, E. J., AND M. L. STANTON. 2005. Ecological factors influencing tetraploid speciation in snow buttercups (*Ranunculus adoneus*): niche differentiation and tetraploid establishment. *Evolution* 59: 1936-1944.

BALDAUF, H. 2006. Retikulate Evolution in glazialen Refugialgebieten: Zytologische Untersuchungen in der Gattung *Arabidopsis* (Brassicaceae). Universität Heidelberg.

BIRKS, H. J. B., AND K. J. WILLIS. 2008. Alpines, trees, and refugia in Europe. *Plant Ecol Div* 1: 147-160.

CLAUSS, M. J., AND M. A. KOCH. 2006. Poorly known relatives of *Arabidopsis thaliana*. *Trends Plant Sci* 11: 449-459.

CLEMENT, M., D. POSADA, AND K. A. CRANDALL. 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1659.

COMAI, L., A. P. TYAGI, K. WINTER, R. HOLMES-DAVIS, S. H. REYNOLDS, Y. STEVENS, AND B. BYERS. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12: 1551-1568.

DOBEŠ, C., C. KIEFER, M. KIEFER, AND M. A. KOCH. 2007. Plastidic *trn*FUUC pseudogenes in North American genus *Boechera* (Brassicaceae): mechanistic aspects of evolution. *Plant Biol (Stuttg)* 9: 502-515.

DOBEŠ, C. H., T. MITCHELL-OLDS, AND M. A. KOCH. 2004. Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). *Mol Ecol* 13: 349-370.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

EHLERS, J., AND P. L. GIBBARD. 2004. Quaternary glaciations - extent and chronology. Elsevier, Amsterdam.

EXCOFFIER, L., G. LAVAL, AND S. SCHNEIDER. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.

FÉR, T., P. VAŠÁK, J. VOJTA, AND K. MARHOLD. 2007. Out of the Alps or Carpathians? Origin of Central European populations of *Rosa pendulina*. *Preslia* 79: 367-376.

FRENZEL, B. 1964. Über die offene Vegetation der letzten Eiszeit am Ostrande der Alpen. *Vhdlg. Zool.-Botan. Ges. Wien* Band 103/104: 110-143.

GUTIÉRREZ LARENA, B., J. F. AGUILAR, AND G. N. FELINER. 2002. Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. *Mol Ecol* 11: 1965-1974.

GUTIÉRREZ LARENA, B., J. F. AGUILAR, AND G. N. FELINER. 2006. Dispersal across southern Iberian refugia? Integrating RAPDs, sequence data and morphometrics in *Armeria* (Plumbaginaceae). *Folia Geobot* 41: 305-322.

HEUERTZ, M., J. F. HAUSMAN, O. J. HARDY, G. G. VENDRAMIN, N. FRASCARIA-LACOSTE, AND X. VEKEMANS. 2004. Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution* 58: 976-988.

HEWITT, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907-913.

HEWITT, G. M. 1999. Post-glacial re-colonization of European biota. *Biol J Linn Soc* 68: 87-112.

HOLM, S. 1995. Unexpectedly high levels of genetic variation in *Potentilla argentea* L. (s. l.) in southern Sweden. *Hereditas* 123: 127-139.

HUSBAND, B. C. 2000. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proc Biol Sci* 267: 217-223.

JAKOBSSON, M., J. HAGENBLAD, S. TAVARÉ, T. SÄLL, C. HALLDÉN, C. LIND-HALLDÉN, AND M. NORDBORG. 2006. A unique recent origin of the allotetraploid species *Arabidopsis suecica*: Evidence from nuclear DNA markers. *Mol Biol Evol* 23: 1217-1231.

KOCH, M. A., C. DOBEŠ, C. KIEFER, R. SCHMICKL, L. KLIMEŠ, AND M. A. LYSAK. 2007. Supernetwork identifies multiple events of plastid *trn*F(GAA) pseudogene evolution in the Brassicaceae. *Mol Biol Evol* 24: 63-73.

KOCH, M. A., C. DOBEŠ, M. MATSCHINGER, W. BLEEKER, J. VOGEL, M. KIEFER, AND T. MITCHELL-OLDS. 2005. Evolution of the *trn*F(GAA) gene in *Arabidopsis* relatives and the Brassicaceae family: monophyletic origin and subsequent diversification of a plastidic pseudogene. *Mol Biol Evol* 22: 1032-1043.

KOCH, M. A., AND M. MATSCHINGER. 2007. Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 104: 6272-6277.

KOCH, M. A., M. WERNISCH, AND R. SCHMICKL. 2008. *Arabidopsis thaliana´s* wild relatives: an updated overview on systematics, taxonomy and evolution. *Taxon* 57: 933-943.

KOLÁŘ, F., M. ŠTECH, P. TRÁVNÍČEK, J. RAUCHOVÁ, T. URFUS, P. VÍT, M. KUBEŠOVÁ, AND J. SUDA. 2009. Towards resolving the *Knautia arvensis* agg. (Dipsacaceae) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales. *Ann Bot (Lond)* 103: 963-974.

KROPF, M., H. P. COMES, AND J. W. KADEREIT. 2006. Long-distance dispersal vs vicariance: the origin and genetic diversity of alpine plants in the Spanish Sierra Nevada. *New Phytol* 172: 169-184.

MAGRI, D., G. G. VENDRAMIN, B. COMPS, I. DUPANLOUP, T. GEBUREK, D. GÖMÖRY, M. LATAŁOWA, T. LITT, L. PAULE, J. M. ROURE, I. TANTAU, W. O. VAN DER KNAAP, R. J. PETIT, AND J. L. DE BEAULIEU. 2006. A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytol* 171: 199-221.

MESICEK, J. 1970. Chromosome counts in *Cardaminopsis arenosa* agg. (Cruciferae). *Preslia* 42: 225-248.

MEUSEL, H., E. JÄGER, AND E. WEINERT. 1965. Vergleichende Chorologie der zentraleuropäischen Flora. Gustav Fischer Verlag, Jena.

MRÁZ, P., M. GAUDEUL, D. RIOUX, L. GIELLY, P. CHOLER, P. TABERLET, AND THE INTRABIODIV CONSORTIUM. 2007. Genetic structure of *Hypochaeris uniflora* (Asteraceae) suggests vicariance in the Carpathians and rapid post-glacial colonization of the Alps from an eastern Alpine refugium. *J Biogeogr* 34: 2100-2114.

NIKLFELD, H. 1972. Der niederösterreichische Alpenostrand – ein Glazialrefugium montaner Pflanzensippen. *Jahrbuch des Vereins zum Schutze der Alpenpflanzen und -tiere* 37: 1-54.

O´KANE, S. L., JR., AND I. A. AL-SHEHBAZ. 1997. A synopsis of *Arabidopsis* (Brassicaceae). *Novon* 7: 323-327.

PAWŁOWSKI, B. 1970. Remarques sur l´endemisme dans la flore des Alpes et des Carpates. *Vegetatio* 21: 181-243.

PETIT, R. J., S. BREWER, S. BORDÁCS, K. BURG, R. CHEDDADI, E. COART, J. COTTRELL, U. M. CSAIKL, B. VAN DAM, J. D. DEANS, S. ESPINEL, S. FINESCHI, R. FINKELDEY, I. GLAZ, P. G. GOICOECHEA, J. S. JENSEN, A. O. KÖNIG, A. J. LOWE, S. F. MADSEN, G. MÁTYÁS, R. C. MUNRO, F. POPESCU, D. SLADE, H. TABBENER, S. G. M. DE VRIES, B. ZIEGENHAGEN, J. L. DE BEAULIEU, AND A. KREMER. 2002. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *For Ecol Manage* 156: 49-74.

RAMSEY, J., AND D. W. SCHEMSKE. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467-501.

RAMSEY, J., AND D. W. SCHEMSKE. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589-639.

RIESEBERG, L. H., A. WIDMER, A. M. ARNTZ, AND J. M. BURKE. 2003. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos Trans R Soc Lond B Biol Sci* 358: 1141-1147.

RONIKIER, M., E. CIEŚLAK, AND G. KORBECKA. 2008. High genetic differentiation in the alpine plant *Campanula alpina* Jacq. (Campanulaceae): evidence for glacial survival in several Carpathian regions and long-term isolation between the Carpathians and the Alps. *Mol Ecol* 17: 1763-1775.

SÄLL, T., M. JAKOBSSON, C. LIND-HALLDÉN, AND C. HALLDÉN. 2003. Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*. *J Evol Biol* 16: 1019-1029.

SCHMICKL, R., M. H. JØRGENSEN, A. K. BRYSTING, AND M. A. KOCH. 2008a. Phylogeographic implications for the North American boreal-arctic *Arabidopsis lyrata* complex. *Plant Ecology and Diversity* 1: 245-254.

SCHMICKL, R., C. KIEFER, C. DOBEŠ, AND M. A. KOCH. 2008b. Evolution of *trn*F(GAA) pseudogenes in cruciferous plants. *Plant Systematics and Evolution*.

SCHMITT, T. 2009. Biogeographical and evolutionary importance of the European high mountain systems. *Front Zool* 6: 9.

SCHÖNSWETTER, P., M. LACHMAYER, C. LETTNER, D. PREHSLER, S. RECHNITZER, D. S. REICH, M. SONNLEITNER, I. WAGNER, K. HULBER, G. M. SCHNEEWEISS, P. TRÁVNÍČEK, AND J. SUDA. 2007. Sympatric diploid and hexaploid cytotypes of *Senecio carniolicus* (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. *J Plant Res* 120: 721-725.

SCHÖNSWETTER, P., O. PAUN, A. TRIBSCH, AND H. NIKLFELD. 2003. Out of the Alps: colonization of northern Europe by east alpine populations of the glacier buttercup *Ranunculus glacialis* L. (Ranunculaceae). *Mol Ecol* 12: 3373-3381.

SCHÖNSWETTER, P., J. SUDA, M. POPP, H. WEISS-SCHNEEWEISS, AND C. BROCHMANN. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Mol Phylogenet Evol* 42: 92-103.

SEGARRA-MORAGUES, J. G., AND P. CATALÁN. 2008. Glacial survival, phylogeography, and a comparison of microsatellite evolution models for resolving population structure in two species of dwarf yams (*Borderea*, Dioscoreaceae) endemic to the central Pyrenees. *Plant Ecol Div* 1: 229-243.

SKREDE, I., P. B. EIDESEN, R. P. PORTELA, AND C. BROCHMANN. 2006. Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Mol Ecol* 15: 1827-1840.

SOLTIS, D. E., P. S. SOLTIS, D. W. SCHEMSKE, J. F. HANCOCK, J. N. THOMPSON, B. C. HUSBAND, AND W. S. JUDD. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56: 13-30.

SOLTIS, P. S., AND D. E. SOLTIS. 2009. The role of hybridization in plant speciation. *Annu Rev Plant Biol* 60: 561-588.

STUESSY, T. F., H. WEISS-SCHNEEWEISS, AND D. J. KEIL. 2004. Diploid and polyploid cytotype distribution in *Melampodium cinereum* and *M. leucanthum* (Asteraceae, Heliantheae). *Am J Bot* 91: 889-898.

SUDA, J., H. WEISS-SCHNEEWEISS, A. TRIBSCH, G. M. SCHNEEWEISS, P. TRÁVNÍČEK, AND P. SCHÖNSWETTER. 2007. Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high mountain plant *Senecio carniolicus* (Asteraceae). *Am J Bot* 94: 1391-1401.

TRIBSCH, A., AND P. SCHÖNSWETTER. 2003. Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon* 52: 477-497.

URBANI, M. H., C. L. QUARIN, F. ESPINOZA, M. I. O. PENTEADO, AND I. F. RODRIGUES. 2002. Cytogeography and reproduction of the *Paspalum simplex* polyploid complex. *Plant Syst Evol* 236: 99-105.

VARGAS, P. 2003. Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon* 52.

WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Mol Biol* 42: 225-249.

WILLIS, K. J. 1996. Where did all the flowers go? The fate of temperate European flora during glacial periods. *Endeavour* 20: 110-114.

ZIMMERMANN, A. 1972. Pflanzenareale am niederösterreichischen Alpenostrand und ihre florengeschichtliche Deutung. *Dissertationes Botanicae* Band 18: 1-146 (excl. Beilagen).

3. Reticulate evolution in glacial refuge areas – the genus *Arabidopsis* in the eastern Austrian Danube Valley (Wachau)

Abstract

Im dritten Kapitel „Retikulate Evolution in eiszeitlichen Refugialgebieten – die Gattung *Arabidopsis* im ostösterreichischen Donautal (Wachau)“ fanden wir rezente Introgression von *Arabidopsis arenosa* in *Arabidopsis lyrata* ssp. *petraea* in zwei Hybridzonen, eine in der nördlichen Wachau, die andere am Fuße der Ostalpen. In diesen beiden Gebieten liegen die Populationen beider Arten nahe beieinander, was zur Annahme von aktuellem Genfluss zwischen ihnen führte. Die Hybridzone in der nördlichen Wachau wurde sowohl mit molekularen Markern (cpDNA *trn*L/F Sequenzdaten und sieben Mikrosatelliten) als auch morphologischen Daten charakterisiert. Die Hybridzone am Fuße der Ostalpen wurde mit Hilfe von Mikrosatelliten entdeckt. Tetraploide, besonders von *Arabidopsis lyrata* ssp. *petraea*, zeigten stark erhöhte geno- und phänotypische Plastizität im Gegensatz zu Diploiden. Polyploidisierung innerhalb von *Arabidopsis arenosa* und *Arabidopsis lyrata* ssp. *petraea* fand vermutlich durch Selbstverdopplung des Genoms statt.

In the third chapter “Reticulate evolution in glacial refuge areas – the genus *Arabidopsis* in the eastern Austrian Danube Valley (Wachau)” we found ongoing introgression from *Arabidopsis arenosa* into *Arabidopsis lyrata* ssp. *petraea* in two hybrid zones, one in the northern Wachau, the other in the foothills of the Eastern Alps. In both areas the two species are in close geographic contact, which led to the assumption of recent geneflow between them. The hybrid zone in the northern Wachau was supported both by cpDNA *trn*L/F sequence data, seven ntDNA microsatellite markers, and morphological data. The hybrid zone in the foothills of the Eastern Alps was detected with ntDNA microsatellite markers. Tetraploids, especially of *Arabidopsis lyrata* ssp. *petraea*, showed strongly increased geno- and phenotypic plasticity over diploids. Autopolyploidisation probably was the common mode of polyploidisation within both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*.

3.1. Introduction

Evolutionary theory is based on the fundamental principle, that evolutionary change is a populational process in a temporal and spatial framework. Genetic variation within populations arises by random mutation and recombination. However, mutation and recombination are only prerequisites for two evolutionary phenomena, which contribute to changes of allele and genotype frequencies and even replacement of genotypes over generations within populations: random genetic drift and nonrandom natural selection. Therefore, genetic drift and natural selection are the main driving forces for the initiation of speciation. Meanwhile it is widely accepted, that hybridisation and/or polyploidisation do play a major role in plant and animal speciation processes (Cronn and Wendel, 2004; Mallet, 2007; Rieseberg and Willis, 2007; Abbott et al., 2008; Doyle et al., 2008; Flagel and Wendel, 2009; Soltis and Soltis, 2009), where genetic drift and natural selection may even be increased, caused by increased genetic variability within species and transfer of adaptations between species (Arnold, 2004). Hybridisation and polyploidisation are extremely common in angiosperms (> 70%) and pteridophytes (> 95%) (Rieseberg, 1997; Cronn and Wendel, 2004; Hegarty and Hiscock, 2005).

According to Anderson (1948), the most common mode of hybridisation is introgression. Introgressive hybridisation, the repeated backcrossing of a hybrid with one of the parental species, is proven to be an important factor in evolution, as it largely contributes to increased genotypic and phenotypic variation, even transferring adaptive traits and characters from one taxon to another. Introgressive hybridisation is only one possibility for hybrid speciation, however, a very common one. Although main barriers to the formation of hybrids are prezygotic (e.g. differences in flowering time, divergent pollinator spectrum, genetic incompatibilities), the establishment of a new hybrid taxon in its environment depends on successfully overcoming the postzygotic incompatibilities (e.g. embryo lethality, hybrid mortality, hybrid sterility). Introgressed taxa are rapidly stabilised by backcrossing of the hybrid F1 generation with one of its parental populations over generations. With only little rapid genomic changes involved, introgressed hybrids overcome the postzygotic barriers to hybrid formation. In contrast, rapid genomic changes play a major role in homo- and heteroploid hybrid formation. In homoploid hybridisation gametes of the same ploidy level from taxonomically different populations fuse and constitute a new hybrid individual. Gametes of different ploidy levels are involved in heteroploid hybrid formation. Rapid

genomic changes include alterations in genome size (e.g. gene loss), chromosomal structure, and gene expression (e.g. gene silencing) (Baack and Rieseberg, 2007). In the course of rearrangement of the hybrid genome de novo adaptations may originate (Rieseberg et al., 2003), selection could act on advantageous adaptations, contribute to ecological and geographical separation of the hybrid from its parental taxa and, therefore, to hybrid speciation. Homoploid hybrids are frequently stabilised by chromosome doubling, which leads to the formation of allopolyploids with one set of the parental genome each (Rieseberg, 1997; Ramsey and Schemske, 1998). Rapid genomic changes in allopolyploids include organ-specific duplicate gene silencing and activation of dormant transposable elements (Adams and Wendel, 2005a, 2005b). The combination of two genomes results in a novel, but coordinated combination of the transcriptional networks, as studied in the genus *Gossypium* (Rapp et al., 2009). Even during the different developmental stages of a single cotton fibre cell, the duplicate genes showed altered ratios of contribution to transcription (Hovav et al., 2008). Hybridisation is not involved in the second common mechanism of polyploidisation, autopolyploidisation, which is the result of self-doubling of a whole genome through the fusion of unreduced gametes (Bretagnolle and Thompson, 1995). A first study on synthetic autotetraploid *Arabidopsis thaliana* indicated, that no rapid genomic changes in form of gene loss occurred during the early stages of autopolyploid formation (Ozkan et al., 2006). Frequently, intermediate stages of auto- and allopolyploid pairing behaviour are reported, so-called segmental allopolyploids (Stebbins, 1947; Ramsey and Schemske, 2002). Intermediate inheritance patterns can be found throughout several generations during a shift of inheritance from, e.g., tetrasomic to disomic, which is frequent in polyploids (Wolfe, 2001; Ramsey and Schemske, 2002).

Both hybridisation and polyploidisation are observed within and between the three main species complexes of the genus *Arabidopsis* (former genus *Cardaminopsis*), *Arabidopsis halleri* ($2n = 16$), *Arabidopsis arenosa* ($2n = 16, 32$), and *Arabidopsis lyrata* ($2n = 16, 32$) (Clauss and Koch, 2006; Koch and Matschinger, 2007; Koch et al., 2008), as well as the sister species *Arabidopsis thaliana* ($2n = 10$), probably fascilitated by partly overlapping distribution ranges in Central and northern Europe. The first natural hybrid, *Arabidopsis suecica* ($2n = 4x = 26$ (20-28)), was described between diploid *Arabidopsis thaliana* ($n = 5$) and diploid *Arabidopsis arenosa* ($n = 8$) in Fennoscandinavia (O´Kane et al., 1996), with unidirectional geneflow from *Arabidopsis arenosa* to *Arabidopsis thaliana*. Second, *Arabidopsis kamchatica* ($2n = 32$) was recently reported to be a hybrid between Siberian *Arabidopsis lyrata* ssp. *petraea* ($2n = 16$) and a member of the *Arabidopsis halleri* complex,

probably *Arabidopsis halleri* ssp. *gemmifera* (2*n* = 16) (Koch and Matschinger, 2007), also indicated by nuclear *Pgi*C, nuclear ITS, and plastidic *trn*L/F marker data (this PhD thesis, first chapter; Schmickl et al., in prep.). In contrast, no hybridisation and/or polyploidisation events were reported from the two endemic species *Arabidopsis cebennensis* (southern France) and *Arabidopsis pedemontana* (northern Italy).

Here we focus on members of the *Arabidopsis arenosa* and *Arabidopsis lyrata* species complex, in particular on *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* (Fig. 1A-D). The main distribution range of *Arabidopsis arenosa* is Central and northern Europe, with diploids exclusively found in southeastern Europe and the Carpathians (this PhD thesis, second chapter). Diploid *Arabidopsis arenosa* is restricted to calcarous bedrock, whereas tetraploid *Arabidopsis arenosa* also grows on siliceous bedrock. *Arabidopsis lyrata* ssp. *petraea* is distributed from Central to northern Europe and arctic Russia. Central European *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* show significant differences in life cycle, seed dispersal, and bedrock preference: *Arabidopsis arenosa* has a mainly biennial life cycle, in contrast to the perennial life cycle of *Arabidopsis lyrata* ssp. *petraea*. *Arabidopsis arenosa* is characterised by an increased average plant height over *Arabidopsis lyrata* ssp. *petraea*, facilitating seed dispersal across larger distances in *Arabidopsis arenosa*, as siliques are more exposed to wind, and seeds are stronger expelled from the siliques (personal measurements and observations, unpublished data). *Arabidopsis lyrata* ssp. *petraea* predominantly colonises limestone and dolomite, whereas *Arabidopsis arenosa* also grows on siliceous bedrock. Furthermore, *Arabidopsis arenosa* is characterised by a broader ecological amplitude than *Arabidopsis lyrata* ssp. *petraea*, which made it a successful postglacial coloniser (Meusel et al., 1965). *Arabidopsis lyrata* ssp. *petraea* is restricted to few cryptic, warm-stage refugia on exposed rocks and rocky slopes with low competition (Jonsell et al., 1995; Clauss and Mitchell-Olds, 2003), resulting in a disjunct distribution.

Regarding the *Arabidopsis arenosa* species complex no studies on the population level have been published so far, and, therefore, this study will provide first insights into neutral variation by applying nuclear microsatellites and plastidic sequences. In contrast, several population studies of the *Arabidopsis lyrata* complex have been conducted, including isozymes, microsatellites, and nuclear as well as plastidic sequence polymorphism. These studies mainly focussed on northern European populations of *Arabidopsis lyrata* ssp. *petraea* and *Arabidopsis lyrata* ssp. *lyrata* populations from North America. Most populations from Europe revealed a higher genetic diversity than North American populations, indicating population bottlenecks during Pleistocene climate oscillations (Van Treuren et al., 1997;

Savolainen et al., 2000; Wright et al., 2003; Ramos-Onsins et al., 2004). Low genetic diversity of North American populations in comparison to populations from Europe was recently supported in a phylogeographic study based on plastidic sequence polymorphism (this PhD thesis, first chapter; Schmickl et al., in prep.). Central European populations of *Arabidopsis lyrata* ssp. *petraea* were characterised by high within population diversity and a high degree of genome-wide heterozygosity, according to microsatellite analyses (Clauss and Mitchell-Olds, 2006). These findings supported periglacial survival, at least during the LGM, of this cold-adapted taxon in tundra and tundra-steppe vegetation. However, all population studies of the *Arabidopsis lyrata* species complex have up to now concentrated on diploid populations. Hence, this study will be the first one including tetraploids as well.

Our sampling was focussed on the northeastern Austrian Alps, spanning an approximately 900 km^2 large area (Fig. 1). The whole region remained unglaciated during Pleistocene climate oscillations, even enabling periglacial survival of *Homo sapiens*, as reported from the Danube Valley (Szombathy, 1884): Two cultic, probably Venus statuettes, were found in the Wachau, one approximately 32,000 years old near Krems, the other, "Venus of Willendorf" (Szombathy, 1910), with an age of about 24,000 years.

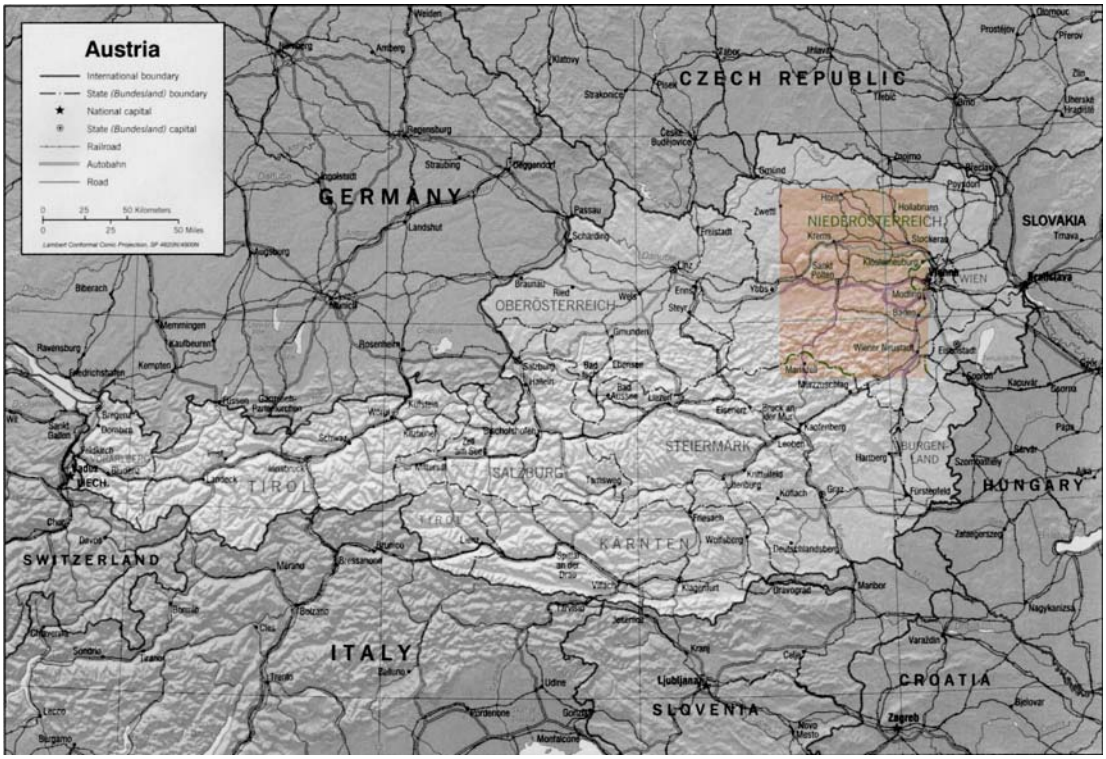


Fig. 1. Localisation of the approximately 900 km^2 large research area in Lower Austria, marked in light pink. Map taken from http://www.2cairo.com (free global information).

The study area can be subdivided into three major regions, according to bedrock and relief: (1) Mariazeller Land (Fig. 2), a limestone area, which is located in the foothills of the northeastern Austrian Alps. For a better geographic orientation of the reader this region will frequently be called “Eastern Alps”, although the Mariazeller Land is only a small part of the Eastern Alps. In the western part of the Mariazeller Land tetraploid *Arabidopsis arenosa* populations expand from colline and montane habitats into subalpine boulder of the northeastern Limestone Alps, up to an altitude of approximately 1000 mNN. Diploid *Arabidopsis lyrata* ssp. *petraea* populations are restricted to the warmer, eastern margin of the Mariazeller Land, adjoining the Viennese Basin with mild, pannonical climate. The small populations are found in cryptic warm-stage refugia like single, exposed rocks or rocky slopes, surrounded by temperate forest. In the central and northern Mariazeller Land tetraploid *Arabidopsis lyrata* ssp. *petraea* populations mainly colonised oak-pine forests with considerable population sizes, expanding into the adjacent Wachau.

Fig. 2. Localisation of investigated *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations in the Eastern Alps (Mariazeller Land). Each coloured circle represents a population. The population numbers are indicated in the circles. The four diploid *Arabidopsis lyrata* ssp. *petraea* populations are marked with an asterisk.

(2) The Wachau (Fig. 3, 5G), part of the Bohemian Massif, was intersected by the Danube River approximately five million years ago. This area is characterised by siliceous bedrock of mainly Gföhler Gneiss and amphibolite, with few heavy metal polluted sites on serpentine bedrock. The Wachau is rich in landscape structure, with steep rocks and numerous side valleys. In contrast to the southern part, the northern and northeastern parts are under pannonical climate influence. Both tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations are mainly found in oak-pine forests with gravel or scattered exposed rocks (Fig. 5H-J). These habitats are located on the slopes of hills facing the Danube River. The populations are clearly separated from each other by beech forest, which covers the shady side valleys. Interestingly, *Arabidopsis lyrata* ssp. *petraea* colonised nearly the whole Wachau on both riversides except for the northern part. Here large *Arabidopsis arenosa* populations are distributed, which expand into the nearby Krems and Kamp Valley.

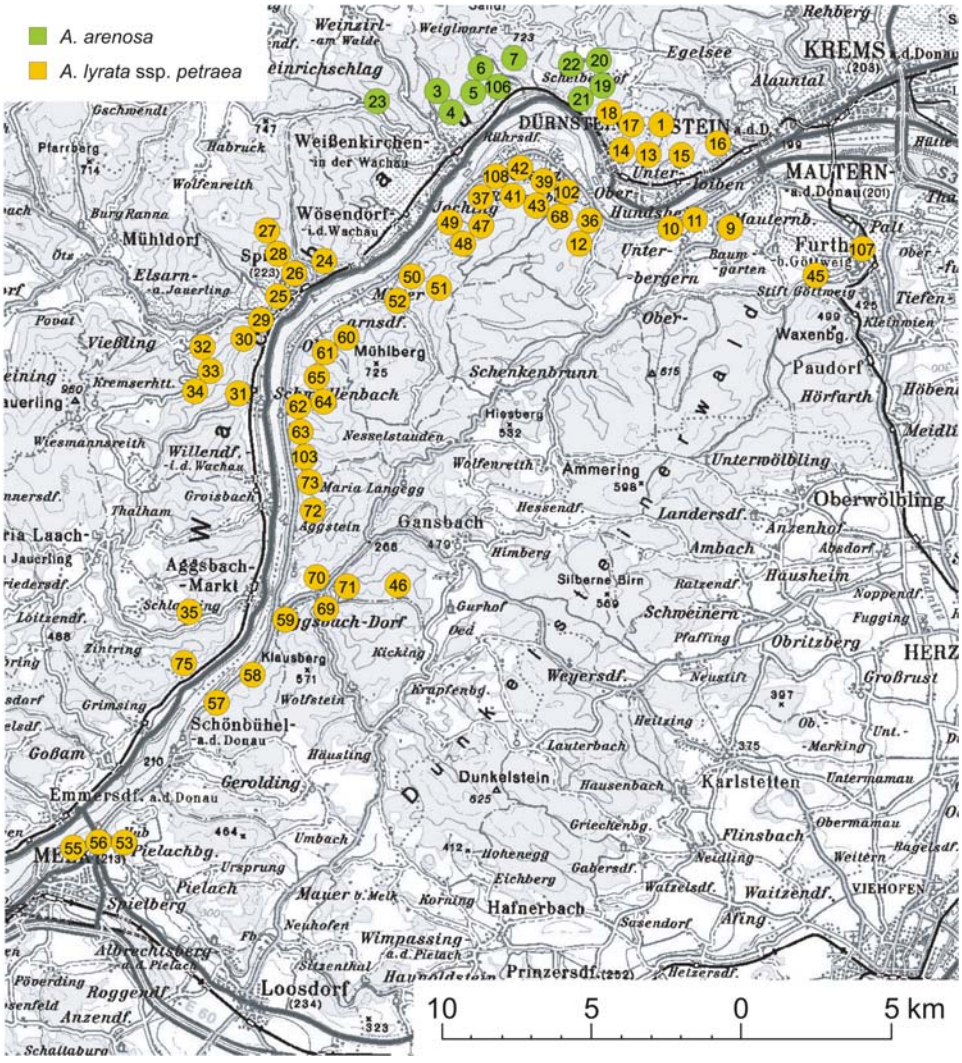


Fig. 3. Localisation of investigated *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations in the Wachau. Each coloured circle represents a population. The population numbers are indicated in the circles.

(3) In the Krems and Kamp Valley (Fig. 4), also part of the Bohemian Massif, similar habitats are colonised. However, several oak-pine forests are currently invaded by *Robinia pseudoacacia*, leading to nutrient-rich, shady forests with a competitive herbal layer (Fig. 5K). In these forests *Arabidopsis arenosa* shows decreased population densities and overall fitness (personal observation).

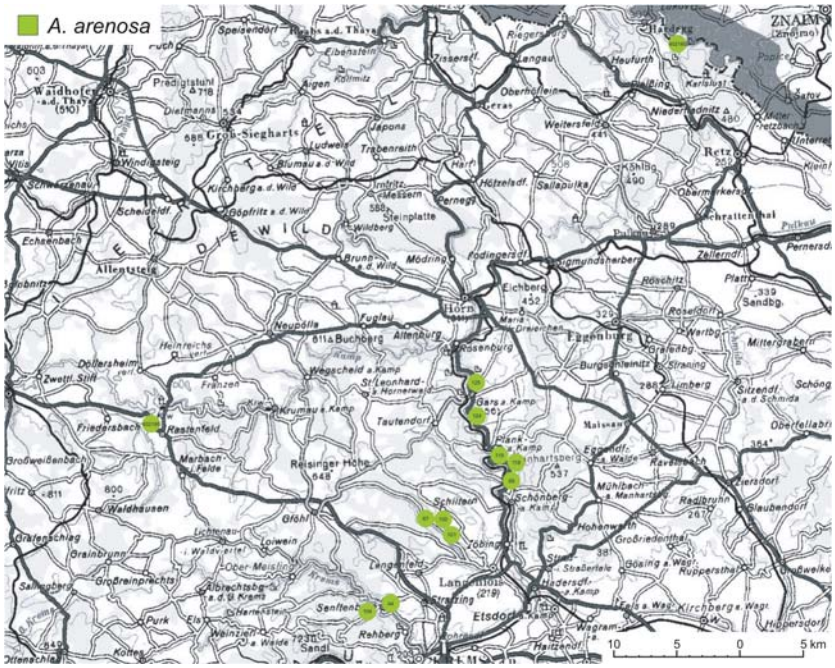


Fig. 4. Localisation of investigated *Arabidopsis arenosa* populations in the Wachau. Each coloured circle represents a population. The population numbers are indicated in the circles.

In all three major regions anthropogenic influence played an important role for nowadays distribution patterns. Castle ruins serve as additional habitats for both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* (Fig. 5L), road construction results in population expansion along rocks and rocky batter (Fig. 5M, N). In contrast, forestry and viticulture, especially in the Wachau, cause restriction of populations. As no diploid *Arabidopsis arenosa* populations were found in Eastern Austria, populations from the Western Carpathians were analysed, being geographically closest to the three major regions of our study area.

Previous work on morphological and molecular data indicated hybridisation between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* in the Wachau (Mall, 2005; Matschinger, 2007). Additionally, the unglaciated part of the Eastern Alps was described as the centre of genetic diversity for both the *Arabidopsis arenosa* and *Arabidopsis lyrata* species complex (Koch and Matschinger, 2007). Hence, our study area is an evolutionary hotspot within the genus *Arabidopsis*.

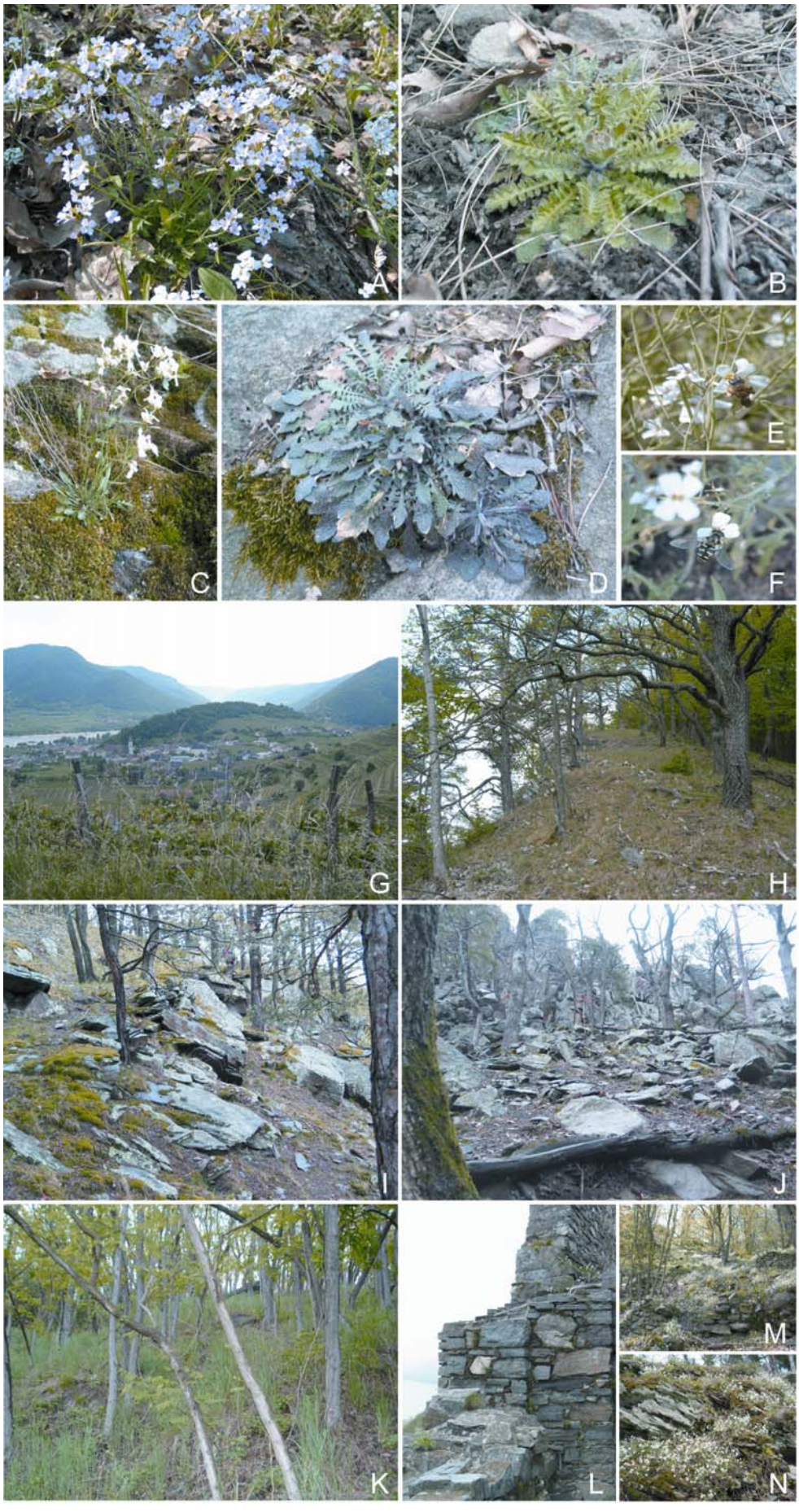

A
B
C
D
E
F
G
H
I
J
K
L
M
N

Fig. 5 (previous page). A. *Arabidopsis arenosa* (2*n* = 32) single individual in full bloom at the end of May, northern Wachau (Pop5). B. *Arabidopsis arenosa* (2*n* = 32) ground rosettes of three individuals at the beginning of March, northern Wachau (Pop22). C. *Arabidopsis lyrata* ssp. *petraea* (2*n* = 32) single individual in full bloom at the end of May, southern Wachau (Pop56). D. *Arabidopsis lyrata* ssp. *petraea* (2*n* = 32) ground rosettes of three individuals at the beginning of March, northern Wachau (Pop14). E. Indetermined member of the *Apiformes* pollinating *Arabidopsis petrogena* nom. prov. in the Botanical Garden Heidelberg. F. Indetermined member of the *Syrphidae* pollinating *Arabidopsis arenosa* ssp. *borbasii* in the eastern Swabian Mountains. G. View into the southern part of the Wachau. In the foreground the "Tausendeimerberg" in Spitz. H. Oak-pine forest with *Arabidopsis lyrata* ssp. *petraea* in the herbal layer, Eastern Alps (Pop116). I. Exposed rocks in an oak-pine forest with *Arabidopsis arenosa* in the herbal layer, northern Wachau (Pop3). J. Exposed rocks in an oak-pine forest with *Arabidopsis lyrata* ssp. *petraea* in the herbal layer, southern Wachau (Pop50). K. Oak-pine forest invaded by *Robinia pseudacacia* with *Arabidopsis arenosa* in the herbal layer, Kamp Valley (Pop118). L. *Arabidopsis lyrata* ssp. *petraea* growing on a wall of Castle Ruin Aggstein, southern Wachau (Pop72). M.-N. *Arabidopsis lyrata* ssp. *petraea* on exposed rocks and rocky batter along a street (Pop34).

In this study we want to clarify the following aspects, using the maternally inherited, plastidic markers *trn*L intron (*trn*L) and *trn*L/F intergenic spacer (*trn*L/F-IGS) of $tRNA^{Ser}$ and $tRNA^{Thr}$, respectively, seven genome-wide nuclear microsatellite markers, and classical morphometric measurements: (1) Evaluating, if hybrid formation between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* is common and occurred via uni- or bidirectional introgression or homoploid hybridisation. (2) Unravelling, if hybrids were subsequently stabilised by polyploidisation, leading to allopolyploid formation, or if autopolyploidisation was the mode of polyploidisation. (3) Detecting, if hybrids preferentially occurred in hybrid zones, and localising those hybrid zones. (4) Unravelling the evolutionary history of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* in Eastern Austria, and detecting colonisation routes of the two species.

3.2. Material and Methods

3.2.1. Plant material

CpDNA. The majority of specimen was collected in the field during the following main collection trips: Eastern Austria, 12.04.-16.05.2003, leg. Michaela Matschinger; Eastern Austria, 17.04.-04.05.2005, leg. Roswitha Schmickl and Michaela Matschinger; Eastern Austria, 13.06.-24.06.2005, leg. Roswitha Schmickl and Marcus Koch; Eastern Austria, 19.03.-01.04.2006, leg. Roswitha Schmickl and Tilman Ellwanger; Eastern Austria and Western Carpathians, 08.07.-20.07.2007, leg. Roswitha Schmickl and Marcus Koch. In addition, several populations from Eastern Austria, Germany, and northern Italy were collected from 2006 to 2009 during short field trips. Chloroplast sequence polymorphism was analysed in ten individuals per population on average. Altogether 4 diploid *Arabidopsis arenosa* populations with 23 individuals from the Western Carpathians were investigated (Pop126-128, Pop131). Twenty-six tetraploid *Arabidopsis arenosa* populations with 227 individuals were included in the analyses: 22 populations from the Wachau and neighbouring regions (Pop3-7, Pop19-23, Pop106, Pop44, Pop67, Pop89, Pop100, Pop101, Pop104, Pop118, Pop119, Pop124, Pop125, Pop402195), and 4 populations from the Eastern Alps (Pop402140, Pop402159, Pop402168, Pop402192). The following 60 populations with 641 individuals in total of tetraploid *Arabidopsis lyrata* ssp. *petraea* were studied: 58 populations from the Wachau (Pop1, Pop9-18, Pop24-37, Pop39, Pop41-43, Pop45-53, Pop55-65, Pop68-73, Pop75, Pop102, Pop103), and 2 populations from the Eastern Alps (Pop116, Pop402196). Two diploid *Arabidopsis lyrata* ssp. *petraea* populations with 14 individuals from the Eastern Alps (Pop115, Pop120) were analysed. Additionally, the following populations were cultivated from seeds in the Botanical Garden Heidelberg and included in the analyses: 2 diploid *Arabidopsis arenosa* populations with 20 individuals (Pop915140, Pop915141), 1 tetraploid *Arabidopsis arenosa* population with 6 individuals (Pop915142), 2 diploid *Arabidopsis lyrata* ssp. *petraea* populations with 11 individuals (Pop915143, Pop915145), and 1 tetraploid *Arabidopsis lyrata* ssp. *petraea* population with 4 individuals (Pop915144). In total 946 individuals from 98 populations were analysed.

Microsatellites. The same populations were included in the analyses as in the study with plastidic *trn*L/F, however, with several modifications: Numerous populations from the Wachau were not investigated, as their habitat was anthropogenically influenced very recently, e.g. through road construction or gardening (Pop19, Pop33, Pop43, Pop47, Pop49, Pop52,

Pop63, Pop73). In addition, 4 tetraploid *Arabidopsis lyrata* ssp. *petraea* populations were studied (Pop85, Pop107, Pop108, Pop402160). Furthermore, several diploid populations of various taxonomic ranking were included in the analyses: *Arabidopsis neglecta* from the Western Carpathians (Pop130), *Arabidopsis petrogena* nom. prov. from the Western Carpathians (Pop915138, Pop915139 – both populations cultivated from seeds in the Botanical Garden Heidelberg), diploid *Arabidopsis lyrata* ssp. *petraea* from Czech Republic (Pop96), diploid *Arabidopsis lyrata* ssp. *petraea* from Franconian Switzerland/Germany (Pop112, Pop133), *Arabidopsis halleri* ssp. *halleri* from Sauerland/Germany (Pop132), *Arabidopsis halleri* ssp. *ovirensis* from Carinthia/Austria (PopOvir), and *Arabidopsis pedemontana* from northern Italy (Pop114). On average 20 individuals per population were analysed.

Morphology. Morphological investigations of tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* were concentrated on populations from the Wachau. Additionally, diploid populations of both species were included in the analyses. The majority of specimen was collected in the field during a collection trip from 12.04.-16.05.2003 by Michaela Matschinger, with 5-20 individuals per population. Altogether 21 tetraploid *Arabidopsis arenosa* populations with 185 individuals were analysed: 15 populations from the Wachau and adjacent regions (Pop3-Pop7, Pop19-Pop23, Pop67, Pop89-91, Pop93), 3 populations from the Eastern Alps (Pop81, Ex04, Ex05), and 3 populations from Upper Austria (Pop54) and Slovenia (Ex07, Ex09). Fifty-one populations with 435 individuals from tetraploid *Arabidopsis lyrata* ssp. *petraea* were under study: 46 populations from the Wachau (Pop1, Pop9, Pop10, Pop12, Pop14-18, Pop24, Pop26-37, Pop39, Pop41, Pop43, Pop45, Pop46, Pop51-53, Pop55-58, Pop60-62, Pop64, Pop65, Pop68-73, Pop75), and 5 populations from the Eastern Alps (Pop66, Pop76-78, Pop84). Six diploid *Arabidopsis lyrata* ssp. *petraea* populations with 71 individuals were included in the analyses: 4 populations from the Eastern Alps (Pop74, Pop87, Pop115, Pop915143), and 2 populations from Czech Republic (Pop94, Pop95). Additionally, the following populations were cultivated from seeds in the Botanical Garden Heidelberg and collected in their first flowering year: 2 diploid *Arabidopsis arenosa* populations with 25 individuals (Pop915140, Pop915141), and 1 diploid *Arabidopsis lyrata* ssp. *petraea* population with 15 individuals (Pop915145). In total 731 individuals from 81 populations were analysed.

The accession list is provided as Supplementary material TABLE 3.

3.2.2. DNA isolation, amplification and sequencing

Total DNA was obtained from dried leaf material and extracted according to the CTAB protocol of Doyle and Doyle (1987) with the following modifications: 50–75 mg of dry leaf tissue were ground in 2 ml tubes using a Retsch swing mill (MM 200), 2 units of ribonuclease per extraction were added to the isolation buffer, and the DNA pellets were washed twice with 70% ethanol. DNA was dissolved in 50 µl TE-buffer for storage and diluted 1:3 in TE-buffer before use. For the cpDNA markers *trn*L intron and *trn*L/F intergenic spacer (*trn*L/F IGS) primer design and PCR cycling scheme followed the protocol of Dobeš et al. (2004), using a PTC200 (MJ Research) thermal cycler. The PCR reaction volume of 50 µl contained 1x PCR buffer (10 mM TRIS/50 mM KCl buffer, pH 8.0), 3 mM $MgCl_2$, 0.4 µM of each primer, 0.2 mM of each dNTP, 1 U Taq DNA polymerase (Amersham), and approximately 1 ng of template DNA. Amplified sequences of *trn*L/F IGS included the complete *trn*L/F IGS and the first 18 bases of the *trn*F gene.

Before sequencing PCR products were checked for length and concentrations on 1.5% agarose gels and purified with the NucleoFast Kit (Macherey & Nagel, Germany). Cycle sequencing was performed using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences) and the original primers. Both strands were amplified in order to gain the complete sequence. PCR products were resolved in 10 µl Loading Solution and run on a MegaBace 500 sequencer.

3.2.3. Genotyping

23 microsatellites were chosen from population studies of *Arabidopsis lyrata* ssp. *lyrata* and *Arabidopsis lyrata* ssp. *petraea* (Clauss et al., 2002) and, additionally, from The Arabidopsis Information Resource (TAIR) database (http://arabidopsis.org/home.html). According to the comparative genetic map of *Arabidopsis thaliana* ($x = 5$) and *Arabidopsis lyrata* ssp. *lyrata* ($x = 8$) (Kuittinen et al., 2004; Koch and Kiefer, 2005), microsatellites were selected spanning the whole *Arabidopsis lyrata* ssp. *lyrata* genome, assuming, that a similar SSR distribution can be expected in *Arabidopsis lyrata* ssp. *petraea* and *Arabidopsis arenosa* (TABLE 1). The microsatellites were tested for their quality of both PCR and genotyping run, marker polymorphism and species-specificity of alleles was also considered. Seven SSRs were finally chosen for the analyses (TABLE 2), with the following localisation in the *Arabidopsis lyrata* ssp. *lyrata* genome: SLL2 (“CA” dinucleotide repeat units), ICE13 (“ATC” trinucleotide repeat units), ATTS0392 (“AAG” trinucleotide repeat units), AthZFPG (“CT” dinucleotide

repeat units), all on chromosome 1; NGA162 ("GA" dinucleotide repeat units) on chromosome 3; ICE14 ("GAT" trinucleotide repeat units) on chromosome 4, and ICE7 ("GAA" trinucleotide repeat units) on chromosome 6. Chromosome 1 is overrepresentated, due to the localisation of microsatellites with either high species specificity (SLL2) or high polymorphism (ICE13, ATTS0392, AthZFPG). As no SSRs fulfilling these criteria were found on other chromosomes, these four markers were selected, next to the other three markers located on other chromosomes. Primers flanking the repeats were taken from Clauss et al. (2002).

PCR reactions were performed in a total volume of 12 µl containing 1x PCR buffer (10 mM TRIS / 50 mM KCl buffer / 1.5 mM $MgCl_2$, pH 8.5), 1 mM $MgCl_2$, 0.13 µM of each primer, 0.05 mM of each dNTP, 0.5 U GoTaq DNA polymerase (Amersham), and approximately 5 ng of template DNA, using a PTC200 (MJ Research) thermal cycler. Fragment amplification followed a protocol of denaturation for 3 min at 94 °C, 35 cycles of 20 s at 94 °C, 30 s at 50 °C, and 20 s at 72 °C, final elongation for 20 s at 72 °C, and indefinite hold at 4 °C. Fragments were visualised with FAM, TET or HEX end-labeled forward primers (Metabion), and fragment analyses were conducted as multiplex setup on a MegaBACE 500 capillary sequencer. Each reaction contained 25 ng PCR product, 0.2 µl 400 bp ET-ROX standard and deionised water to bring the final volume up to 6 µl. Samples were denatured at 95 °C for 2 min, placed on ice for 1 min, and then run for 1.5 h on the sequencer at 44 °C through 40 cm capillaries with linear polyacrylamide (GE Healthcare). Scoring of fragment sizes and fluorescence intensity (in tetraploids) was manually performed from the raw data displayed with Genetic Profiler (GE Healthcare).

TABLE 1 (next page). List of the 23 microsatellite primer pairs tested for quality of PCR and genotyping run, marker polymorphism, and species-specificity of alleles. Localisation on the eight *Arabidopsis lyrata* chromosomes, marker names, types of repeat units, range of allele sizes in *Arabidopsis lyrata*, according to Clauss et al. (2002), and the forward and reverse primer sequences are given. SSR markers chosen from The Arabidopsis Information Resource (TAIR) database (http://arabidopsis.org/home.html) are underlined.

Chromosome number *A. lyrata*	SSR marker	Repeat unit	Range of allele sizes in *A. lyrata*, according to Clauss et al. (2002)	Primer forward	Primer reverse
1	F20D22	GTTT	170-210	AACAAAATGAGTTTCTCTGCATG	CCCAAGTGACGTCTGGTTTC
1	F19K23-438	TTC	170-220	GGTCTAATTGCCGTTGTTGC	GAATTCTGTAACATCCCATTTCC
1	SLL2	CA	250-320	CATGTACTGGGATTCAGTGTCC	CGTCCTTTGTGTGGTTACACG
1	ICE13	ATC	210-260	GATCCTTCACCGGGTCTTG	GTGGTGGAGACTCTTCGAGC
1	ADH1	GGT	250-340	ACCACCGGACAGATTATTCG	CCCAGAAGTAAACATCGGTGTG
1	ATTS0392	AAG	130-180	TTTGGAGTTAGACACGGATCTG	GTTGATCGCAGCTTGATAAGC
1	AthZFPG	CT	120-170	TTGCGTTTCCACATTTGTTT	TGGGTCAATTCACATGTAGAGA
2	NGA692	GA	-	TTTAGAGAGAGAGAGCGCGG	AGCGTTTAGCTCAACCCTAGG
3	AthGAPAb	CTT	120-150	CACCATGGCTTCGGTTACTT	TCCTGAGAATTCAGTGAAACCC
3	MDC16	GA	110-140	GAGTGGCCTCGTGTAGAGAAAG	TGTCACTCTTTTCCTCTGGTTTG
3	NGA162	GA	70-110	CATGCAATTTGCATCTGAGG	CTCTGTCACTCTTTTCCTCTGG
4	ICE14	GAT	190-240	TCGAGGTGCTTTCTGAGGTT	TACCTCACCCTTTTGACCCA
4	NGA168	TC	-	TCGTCTACTGCACTGCCG	GAGGACATGTATAGGAGCCTCG
5	NGA112	CT	220-260	TAATCACGTGTATGCAGCTGC	CTCTCCACCTCCTCCAGTACC
5	NGA707	CT	-	CTCTCTGCCTCTCGCTGG	TGAATGCGTCCAGTGAGAAG
6	ICE4	CT	180-220	CACGAGGAATCTGGCATGGTCG	AGCGATTGCAAGCGGCTCAAG
6	ICE7	GAA	80-160	TTCAAGGGCAGGATCAAAAC	GTCTCACTGCTATCGTCACAGG
7	ICE3	CT	70-150	GACTAATCATCACCGACTCAGCCAC	ATTCTTCTTCACTTTTCTTGATCCCG
7	NGA1107	GA	-	CGACGAATCGACAGAATTAGG	GCGAAAAAACAAAAAAATCCA
8	NGA151	CT	90-170	GTTTTGGGAAGTTTTGCTGG	CAGTCTAAAAGCGAGAGTATGATG
8	NGA249	TC	70-140	TACCGTCAATTTCATCGCC	GGATCCCTAACTGTAAAATCCC
8	CA72	CT	190-250	AATCCCAGTAACCAAACACACA	CCCAGTCTAACCACGACCAC
8	ICE1	GA	40-100	GAAGAAACGAAGACGAAGAAGTCG	CCCTTTTGTTCTTCTCCTTTCTC

Chromosome number *A. lyrata*	SSR marker	Repeat unit	Range of allele sizes in diploid *A. arenosa*	Range of allele sizes in tetraploid *A. arenosa*	Range of allele sizes in diploid *A. lyr.* ssp. *petraea*	Range of allele sizes in tetraploid *A. lyr.* ssp. *petraea*
1	SLL2	CA	303-313	263-313	289-313	263-313
1	ICE13	ATC	221-272	197-269	239-254	209-269
1	ATTS0392	AAG	136-180	136-198	136-163	126-180
1	AthZFPG	CT	120-172	128-202	124-168	126-206
3	NGA162	GA	79-99	75-93	79-85	75-93
4	ICE14	GAT	218-239	206-248	224-242	206-260
6	ICE7	GAA	97-142	94-139	94-100	94-139

TABLE 2. List of the 7 microsatellite primer pairs chosen. Ranges of allele sizes in both diploids and tetraploids of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* are given, evaluated within this study.

Data analyses

3.2.4. Plastidic trnL/F sequence definition and network analysis

Plastidic *trn*L/F sequences were defined as haplotypes and suprahaplotypes, according to our previous studies (Koch and Matschinger, 2007; Koch et al., 2008; Schmickl et al., 2008a). As the 3´-region of the *trn*L/F IGS close to the functional *trn*F gene is characterised by multiple *trn*F pseudogenes (Koch et al., 2005; Dobeš et al., 2007; Koch and Matschinger, 2007; Koch et al., 2007; Koch et al., 2008; Schmickl et al., 2008b), which evolve with a higher mutation rate than single nucleotide polymorphisms, we excluded the pseudogene region when defining *trn*L/F suprahaplotypes. Hence, haplotypes are defined as *trn*L intron - *trn*L/F IGS *including* the 3´-end of the *trn*L/F IGS with various copies of pseudogenes, suprahaplotypes as *trn*L intron - *trn*L/F IGS *excluding* the 3´-end of the *trn*L/F IGS with various copies of pseudogenes. Consequently, each *trn*L/F suprahaplotype comprises a varying number of haplotypes, which vary both in length and base content of their pseudogene-rich region.

Network analysis was exclusively performed with suprahaplotypes. The *trn*L/F network was constructed using TCS version 1.21 (Clement et al., 2000), indels (except polyT stretches) were coded as additional binary characters.

3.2.5. Principal Coordinates Analysis and genetic diversity statistics of cpDNA data

Pairwise F_{ST} values were calculated from suprahaplotypes, defined here as *trn*L intron - *trn*L/F intergenic spacer *excluding* the 3´-end of the *trn*L/F intergenic spacer with various copies of pseudogenes (Clauss and Koch, 2006; Koch and Matschinger, 2007; Koch et al., 2008). Calculation of pairwise F_{ST} values was based on a permutation test with 1000 permutations with Arlequin version 3.11 (Excoffier and Schneider, 2005). Based on the resulting distance matrix, Principal Coordinates Analysis (PCoA) was performed to see overall structuring of the dataset, and to detect hybrids. PCoA was carried out with SYN-TAX 2000 (Podani, 2001) and graphically displayed with SPSS version 16.0. Genetic diversity after Nei (1987) was calculated for each population with Arlequin version 3.11 (Excoffier and Schneider, 2005).

3.2.6. Coding of microsatellite alleles

Each allele of each of the seven SSR markers was annotated with a unique number. Automatic coding was performed with the software "webPopCoder", developed by Markus Kiefer (http://ephedra.hip.uni-heidelberg.de/popcoder.php). The outputfile was very similar to GenePop format and could be modified for subsequent analyses. All analyses, except for calculations with Tetraploide version 1.0 (Decarli and Leinemann, 2003), were carried out including missing data, as the results were consistent with the results of analyses without missing data. However, individuals with more than four missing SSR markers were excluded from the calculations.

3.2.7. Genetic assignment tests and population cluster analyses of microsatellite data

Two Bayesian Analyses were applied to identify population structure. BAPS version 5.1 (Corander et al., 2003, 2004, 2008) was used to detect overall patterns of genetic clustering. Group clustering of a mixed dataset of diploid and tetraploid populations, altogether 1769 individuals, including missing data, was performed with K = 2, 3, 4, 5, 9, as calculation of the correct K value remained ambiguous. Populations with a mixed cytotype pattern of both diploids and tetraploids were not observed at all. Admixture analysis of the same dataset, based on individual clustering, was carried out with K = 9. The minimum size of a population constituted five individuals, and calculation was performed with 200 iterations and 100 reference individuals with 20 iterations each. In addition, group clustering of a dataset of

exclusively diploids across the genus *Arabidopsis*, altogether 228 individuals, including missing data, was carried out. With Structure version 2.1 (Pritchard et al., 2000; Falush et al., 2003) only tetraploids of both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, altogether 1541 individuals, including missing data, were analysed. The admixture model was applied, assuming individuals to have inherited a genome fraction from ancestors in population K. This model is certainly correct for the majority of the dataset, due to a common colonisation history of each of the two species in Eastern Austria. The correlated allele frequency model was used, as allele frequencies in different populations of each species are likely to be similar, due to past migration events and/or shared ancestry. Analysis was run for 200000 generations, of which the first 100000 were discarded as burn-in. K value was estimated with R version 2.4.1 and Structure2.1-sum.R script (Ehrich et al., 2007) from seven runs with K values ranging from K = 1-7. Structure analysis was finally performed with K = 4 as inflexion point of the mean L (K) graph (Fig. 6). Proportions of each genetic cluster of individuals from the same population were summed up to the total proportions of all genetic clusters within this population, using the software "Rosis kleine Konditorei", developed by Markus Kiefer (http://emboss.bot.uni-heidelberg.de/konditorei/index.php). Resulting pie charts were plotted on a map. The hybrid index of each tetraploid *Arabidopsis lyrata* ssp. *petraea* population corresponds to the fraction (in percentage) of the genetic cluster characteristic for tetraploid *Arabidopsis arenosa* in tetraploid *Arabidopsis lyrata* ssp. *petraea*.

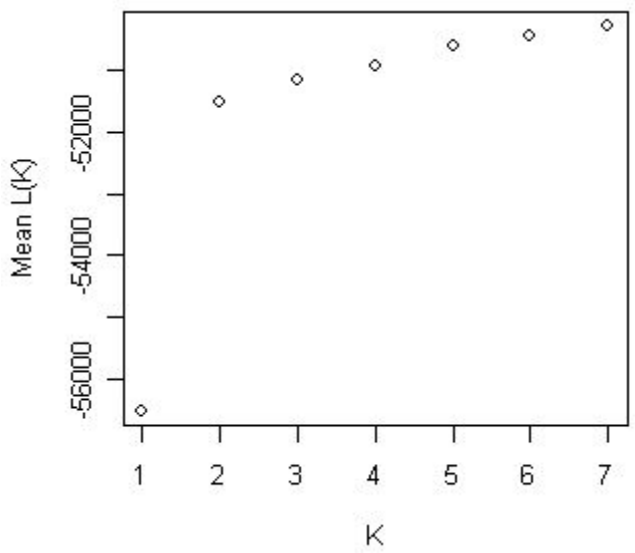


Fig. 6. Distribution of likelihood values across seven independent Structure runs (K = 1-7), assuming admixture and correlated allele frequencies.

3.2.8. Bayesian Test of the mode of inheritance in tetraploids as exemplified by microsatellite data

For the test of tetrasomic versus disomic inheritance after Catalán et al. (2006), the alleles of an SSR locus have to be assigned to the potential parental genomes. In our dataset this could only be achieved for one out of seven microsatellite loci, SLL2, as most alleles were frequently shared between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. In general, several loci across the whole genome should be tested, as not all genomic regions are similarly sensitive to genetic introgression. With the SLL2 locus we tested only a small region on chromosome 1 for tetra- versus disomic inheritance, assuming, that the other six loci would behave in a similar manner. As the SLL2 locus was extremely monomorphic in tetraploid *Arabidopsis arenosa*, only tetraploid *Arabidopsis lyrata* ssp. *petraea* was tested. The assignment of alleles to genotypes was based on their amplification dosage (Fig. 7). Only few populations were excluded from the analysis because of overlapping segregation patterns. Marginal probabilities of the data were computed for the two competing hypotheses for each population.

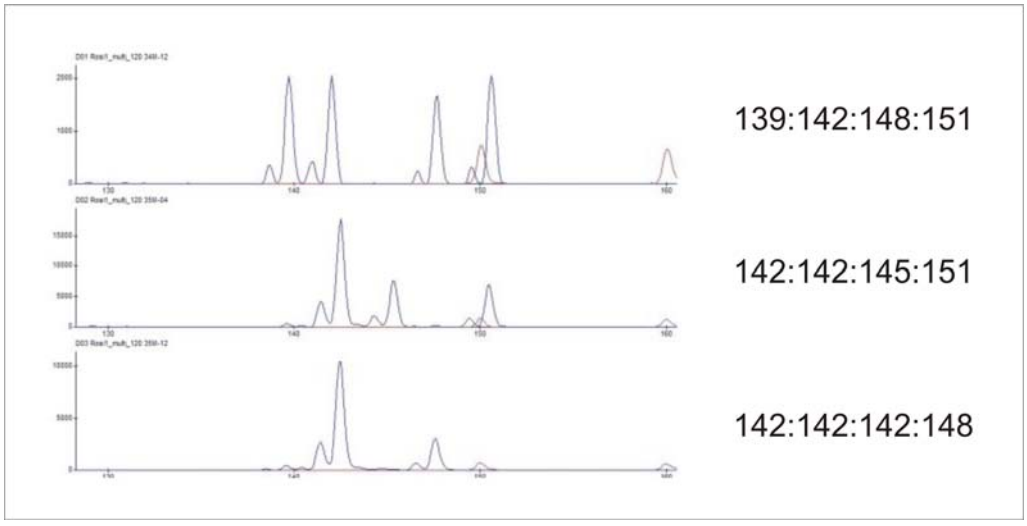


Fig. 7. Electropherograms of genotyping three different individuals of tetraploid *Arabidopsis lyrata* ssp. *petraea* at the SSR locus ATTS0392. Fragment sizes are displayed along the x-axis, fluorescence intensity along the y-axis. The amplification dosage of each allele is measured as total size of the area confined by the allelic peak (in blue). The four alleles of each individual are noted to the right. Peaks in red resemble the 400 bp ET-ROX size standard.

3.2.9. Basic population genetics of diploids and tetraploids based on microsatellite data

Diploids were analysed with FSTAT version 2.9.3.2 (Goudet, 2000). Mean gene diversity He was calculated according to Nei (1987). As sample sizes of the populations were similar, mean allelic richness A was estimated. Inbreeding coefficient F_{IS} was evaluated according to Weir and Cockerham (1984). Observed and expected heterozygosity (Ho/He) were estimated. Tetraploids were analysed with Tetraploide version 1.0 (Decarli and Leinemann, 2003), based on tetrasomic inheritance. Three microsatellites (SLL2, ICE13, ATTS0392) with two-digit allele code numbers were excluded from the analyses, due to inputfile problems with decimal allele coding. Genetic distance d_0 was calculated according to Gregorius (1974) and Prevosti et al. (1975): $d_0\,(k,l) = \frac{1}{2}\sum_i \left| p_{i(k)} - p_{i(l)} \right|$. K and l are two populations, and p_i is the frequency of the ith allele in those populations at a given locus. Mean d_0 was estimated across all loci.

3.2.10. Morphometric analyses

The following fourteen quantitative characters were measured (stem characters were measured on the main stem) (TABLE 3): number of upper stem leaves above the midpoint of the stem (PLHTP), plant height (VR), number of flowers (PK), number of stem leaves (PL), length of leafy stem part (DBPS), length of second stem leaf (D2BL), width of second stem leaf (S2BL), length of uppermost stem leaf (DHL), width of uppermost stem leaf (SHL), length of biggest rosette leaf (MDLR), length of biggest rosette leaf from the first leaf tooth to the leaf tip (DC), basic number of leaf teeth of the biggest rosette leaf (PZ), length of stem from the ground rosette to the biggest stem leaf (VNDL), length of stem from the ground rosette to the smallest stem leaf (VNKL). From these characters six ratios were calculated, as the size of vegetative organs is largely influenced by environmental factors (Dale and Elkington, 1974): VR/DBPS, D2BL/S2BL, DHL/SHL, MDLR/DC, MDLR/PZ, VNKL/VNDL. In addition, fifteen qualitative characters were measured (again stem characters were measured on the main stem) (TABLE 3): hairs on leafy stem part (TRB), hairs on stem from the ground rosette to the first leaf (TRB), hairs on stem from the ground rosette to the first leaf simple, bifurcated, trifurcated (TRB), hairs on leafless stem part (TRS), simple hairs on flower buds (OLJT), bifurcated hairs on flower buds (OLVT), trifurcated hairs on flower buds (OLTT), simple hairs on second stem leaf (2BLJT), bifurcated hairs on second stem leaf (2BLVT), trifurcated hairs on second stem leaf (2BLTT), white petal colour (FLB), light pink petal colour (FKR), dark pink petal colour (FKRT). Selection of these morphological characters was based on previous morphological investigations of the

Arabidopsis arenosa complex (Kolnik and Marhold, unpublished) and further field observations by Marcus Koch. Species-specific characters were chosen by the majority. All characters were measured after drying from herbarium specimen. Raw data of the measurements were mainly obtained from Mall (2006), additional measurements were carried out by Klein (2008) and Schmickl.

Morphological characters	Abbreviation	Character type
Number of upper stem leaves above the midpoint of the stem	PLHTP	quantitative
Plant height	VR	quantitative
Number of flowers	PK	quantitative
Number of stem leaves	PL	quantitative
Length of leafy stem part	DBPS	quantitative
Length of second stem leaf	D2BL	quantitative
Width of second stem leaf	S2BL	quantitative
Length of uppermost stem leaf	DHL	quantitative
Width of uppermost stem leaf	SHL	quantitative
Length of biggest rosette leaf	MDLR	quantitative
Length of biggest rosette leaf from the first leaf tooth to the leaf tip	DC	quantitative
Basic number of leaf teeth of the biggest rosette leaf	PZ	quantitative
Length of stem from the ground rosette to the biggest stem leaf	VNDL	quantitative
Length of stem from the ground rosette to the smallest stem leaf	VNKL	quantitative
Ratios: VR/DBPS, D2BL/S2BL, DHL/SHL, MDLR/DC, MDLR/PZ, VNKL/VNDL		
Hairs on leafy stem part	TRB	qualitative
Hairs on stem from the ground rosette to the first leaf	TRB	qualitative
Simple hairs on stem from the ground rosette to the first leaf	TRB	qualitative
Bifurcated hairs on stem from the ground rosette to the first leaf	TRB	qualitative
Trifurcated hairs on stem from the ground rosette to the first leaf	TRB	qualitative
Hairs on leafless stem part	TRS	qualitative
Simple hairs on flower buds	OLJT	qualitative
Bifurcated hairs on flower buds	OLVT	qualitative
Trifurcated hairs on flower buds	OLTT	qualitative
Simple hairs on second stem leaf	2BLJT	qualitative
Bifurcated hairs on second stem leaf	2BLVT	qualitative
Trifurcated hairs on second stem leaf	2BLTT	qualitative
White petal colour	FLB	qualitative
Light pink petal colour	FKR	qualitative
Dark pink petal colour	FKRT	qualitative

TABLE 3. List of the 29 morphological characters chosen for morphometrics. From the 14 quantitative characters 6 ratios were calculated to reduce the influence of environmental factors on the analyses.

According to the biological questions addressed, two types of analyses were performed: (1) The whole dataset of diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, in total 731 individuals from 81 populations, was examined in order to see overall morphological plasticity both within and between species and cytotypes and to detect hybrid individuals. Principal Component Analysis (PCA) was performed, with standardisation by zero mean and unit standard deviation. Euclidian distance was used for computing pairwise similarities. (2) Tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from the Wachau were analysed separately in more detail, focussing on hybridisation between these two species. Each population was screened for the percentage of affiliation to putative "pure" ancestral populations of both species. However, in contrast to *Arabidopsis lyrata* ssp. *petraea*, no diploid populations of *Arabidopsis arenosa* could be found in Eastern Austria. A study with seven genome-wide microsatellite markers had already shown, that diploid *Arabidopsis arenosa* populations from the Carpathians were genetically too distinct from the Austrian populations to serve as putative ancestral populations. Therefore, tetraploid *Arabidopsis arenosa* populations from Eastern Austria, including the Wachau, were selected as putative "pure" ancestral populations, although this approach did not fulfil criteria for Canonical Discriminant Analysis (CDA). Instead, Principal Component Analysis (PCA) was performed, with standardisation by zero mean, unit standard deviation, and Euclidian distance for computing pairwise similarities. PCA was carried out with SYN-TAX 2000 (Podani, 2001) and graphically displayed with SPSS version 16.0. The hybrid index of each tetraploid *Arabidopsis lyrata* ssp. *petraea* population was calculated from the two-dimensional PCA bar plot as deviation of the mean value of the investigated tetraploid *Arabidopsis lyrata* ssp. *petraea* population from the mean value of diploid *Arabidopsis lyrata* ssp. *petraea* (in percentage). In the pie charts plotted on a map morphological proportion of diploid *Arabidopsis lyrata* ssp. *petraea* was coloured in yellow, of tetraploid *Arabidopsis arenosa* in green. Populations with less than five individuals were excluded from the analyses.

3.3. Results

CpDNA

3.3.1. Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea form two genetically mainly distinct groups, but interspecies chloroplast capture is indicated in geographic contact zones

According to the Principle Coordinates Analysis, based on pairwise F_{ST} values, populations formed two major groups, *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* (Fig. 8). The latter mainly clustered along axis one, which comprised 42% total variance. *Arabidopsis arenosa* predominantly grouped along axis two, which covered 14% total variance. Altogether the two axes expressed 56% total variance. Populations of the diploid cytotype of each species were found in the same group as populations of the tetraploid cytotype.

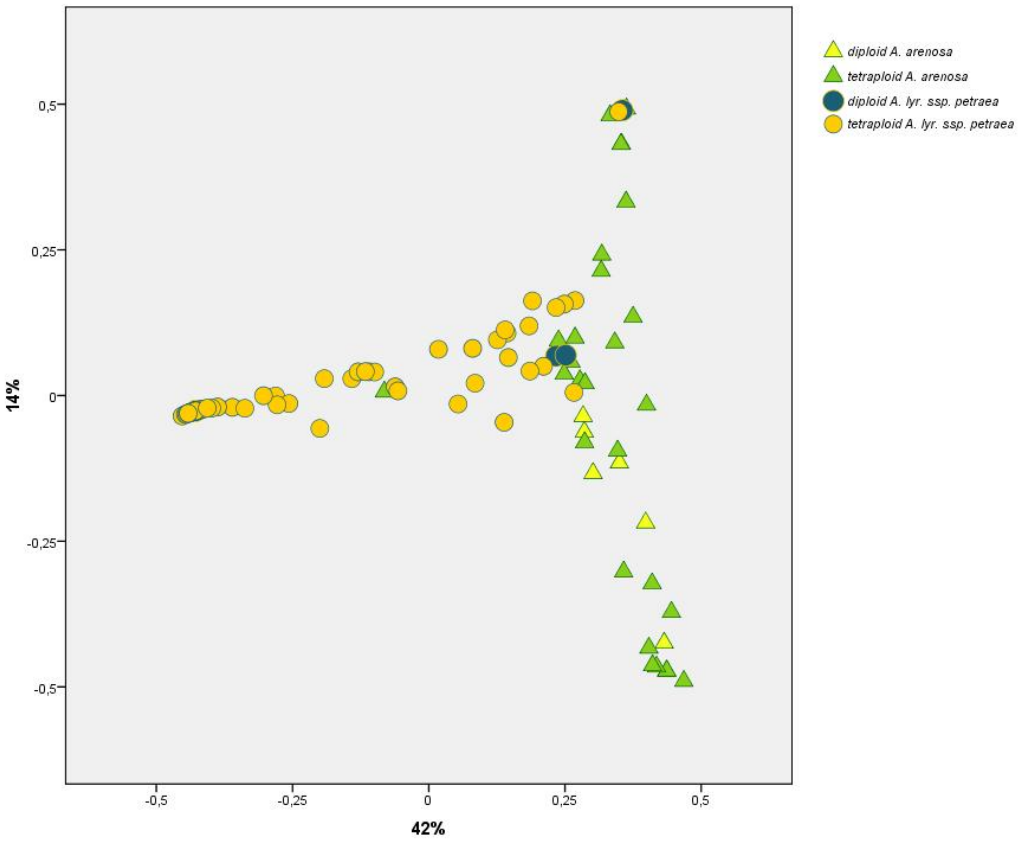


Fig. 8. Principal Coordinates Analysis based on pairwise F_{ST} values of cpDNA *trn*L/F suprahaplotypes. Each symbol (circle, triangle) represents a population of either diploid or tetraploid *Arabidopsis arenosa* or *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*).

This main differentiation of the dataset was supported by TCS network analysis of suprahaplotypes (Fig. 9). Although weak suprahaplotype sharing was observed in the central suprahaplotypes A, B, and C, these sequence types mainly characterised *Arabidopsis arenosa* (A, B) or *Arabidopsis lyrata* ssp. *petraea* (C). Except for Q, all derived "tip" suprahaplotypes were exclusively found in either *Arabidopsis arenosa* (AU, BE, E, L, P, U, Y) or *Arabidopsis lyrata* ssp. *petraea* (AC, AH, AI, AJ, AK, AL, R, V). The observed suprahaplotype sharing could either be the result of incomplete lineage sorting or hybridisation between the two taxa.

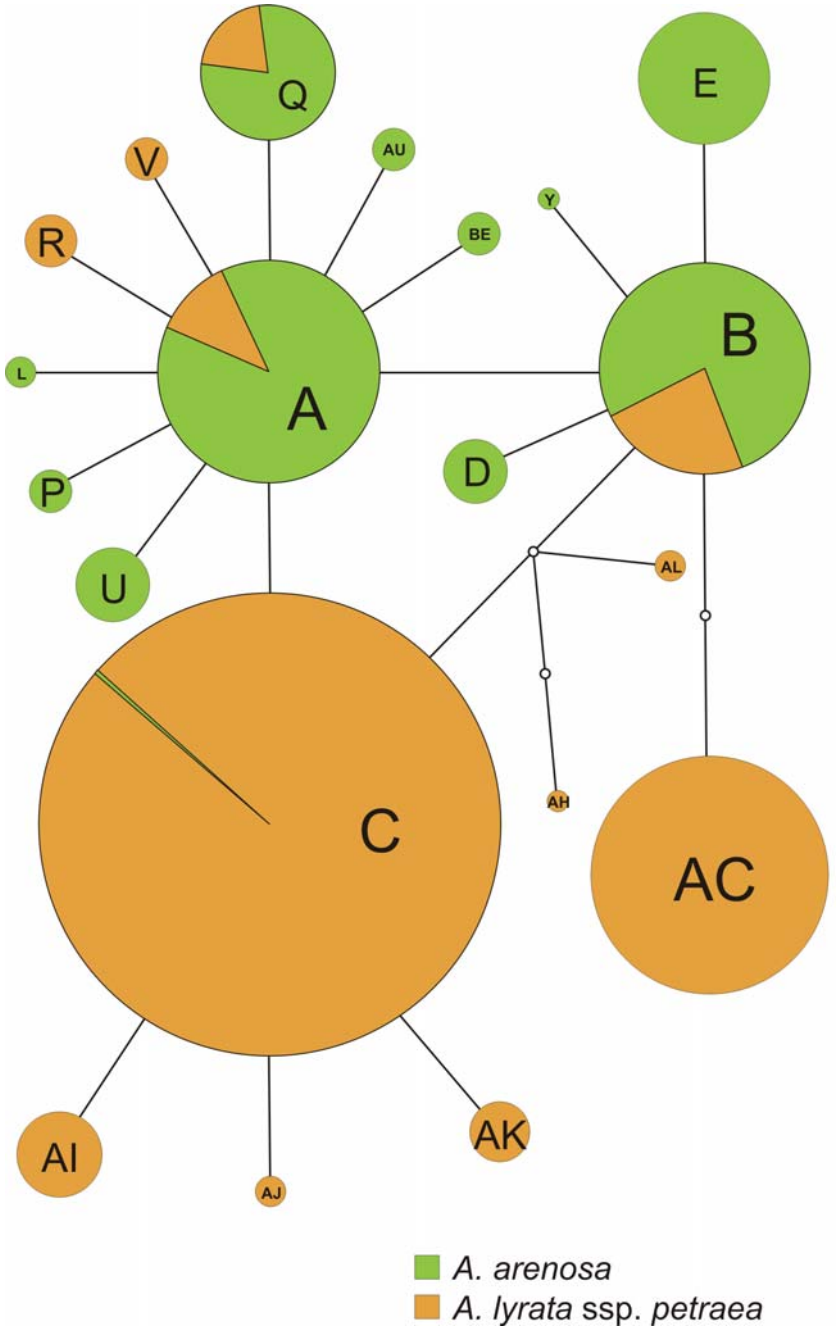


Fig. 9. CpDNA *trn*L/F suprahaplotype network of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*). The sizes of the circles indicate the relative frequency of a suprahaplotype.

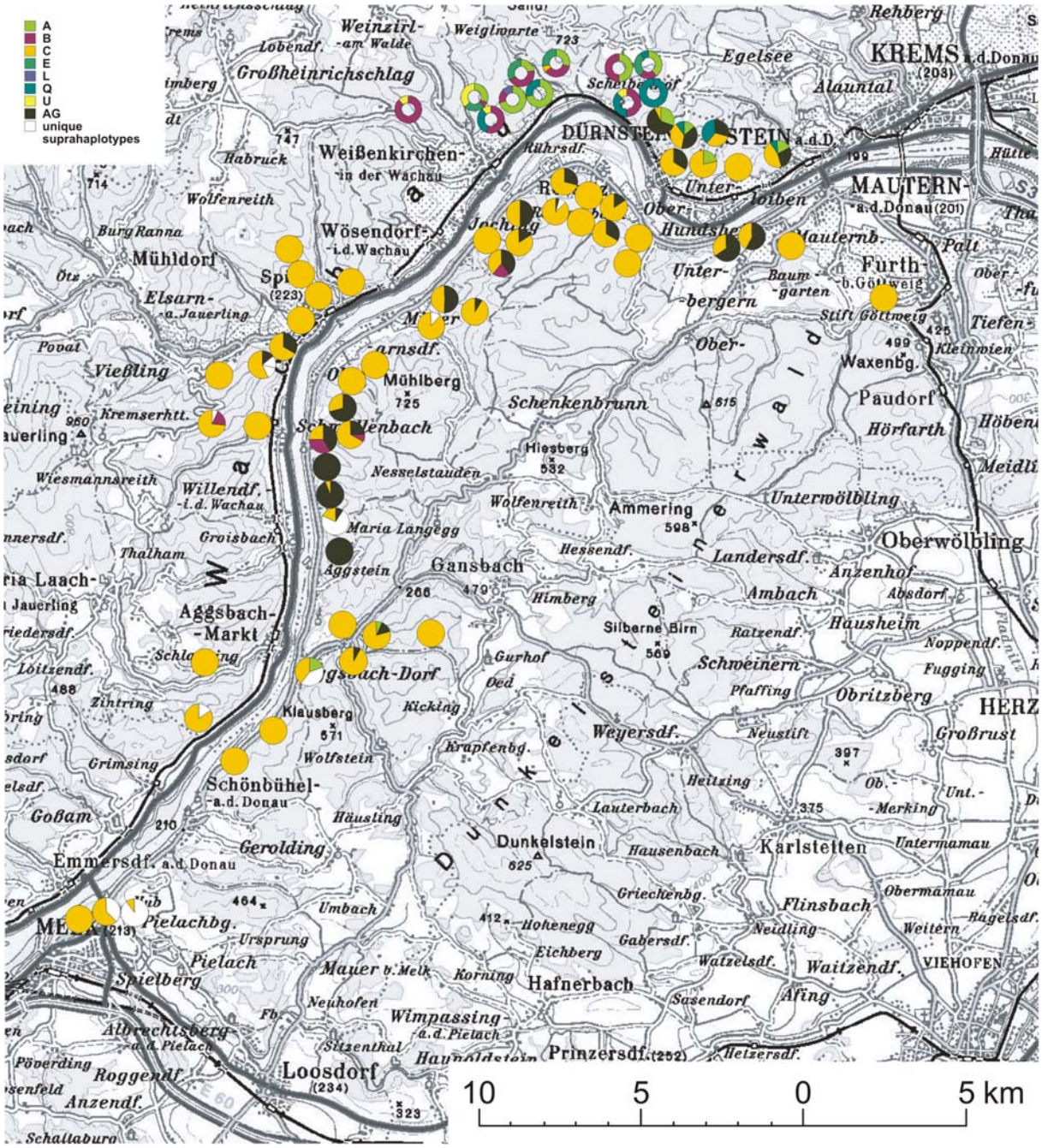


Fig. 10. Distribution of cpDNA *trn*L/F suprahaplotypes of *Arabidopsis arenosa* (rings) and *Arabidopsis lyrata* ssp. *petraea* (circles) in the Wachau. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

Although incomplete lineage sorting was already desribed for the genus *Arabidopsis* (Clauss and Koch, 2006; Koch and Matschinger, 2007), hybridisation in form of chloroplast capture can not be ruled out: Suprahaplotype Q was diagnostic for *Arabidopsis arenosa*, but also found in three *Arabidopsis lyrata* ssp. *petraea* populations in the northern Wachau close to *Arabidopsis arenosa* populations (Fig. 10), which frequently showed this suprahaplotype. Furthermore, suprahaplotype A, also diagnostic for *Arabidopsis arenosa*, was observed in

four of those *Arabidopsis lyrata* ssp. *petraea* populations, indicating unidirectional *Arabidopsis arenosa* seed flow into the distribution range of *Arabidopsis lyrata* ssp. *petraea*. The rare occurence of suprahaplotype C, an *Arabidopsis lyrata* ssp. *petraea* sequence type, in e.g. one *Arabidopsis arenosa* population from the Eastern Alps (Pop402192) could also be explained by chloroplast capture (Fig. 11).

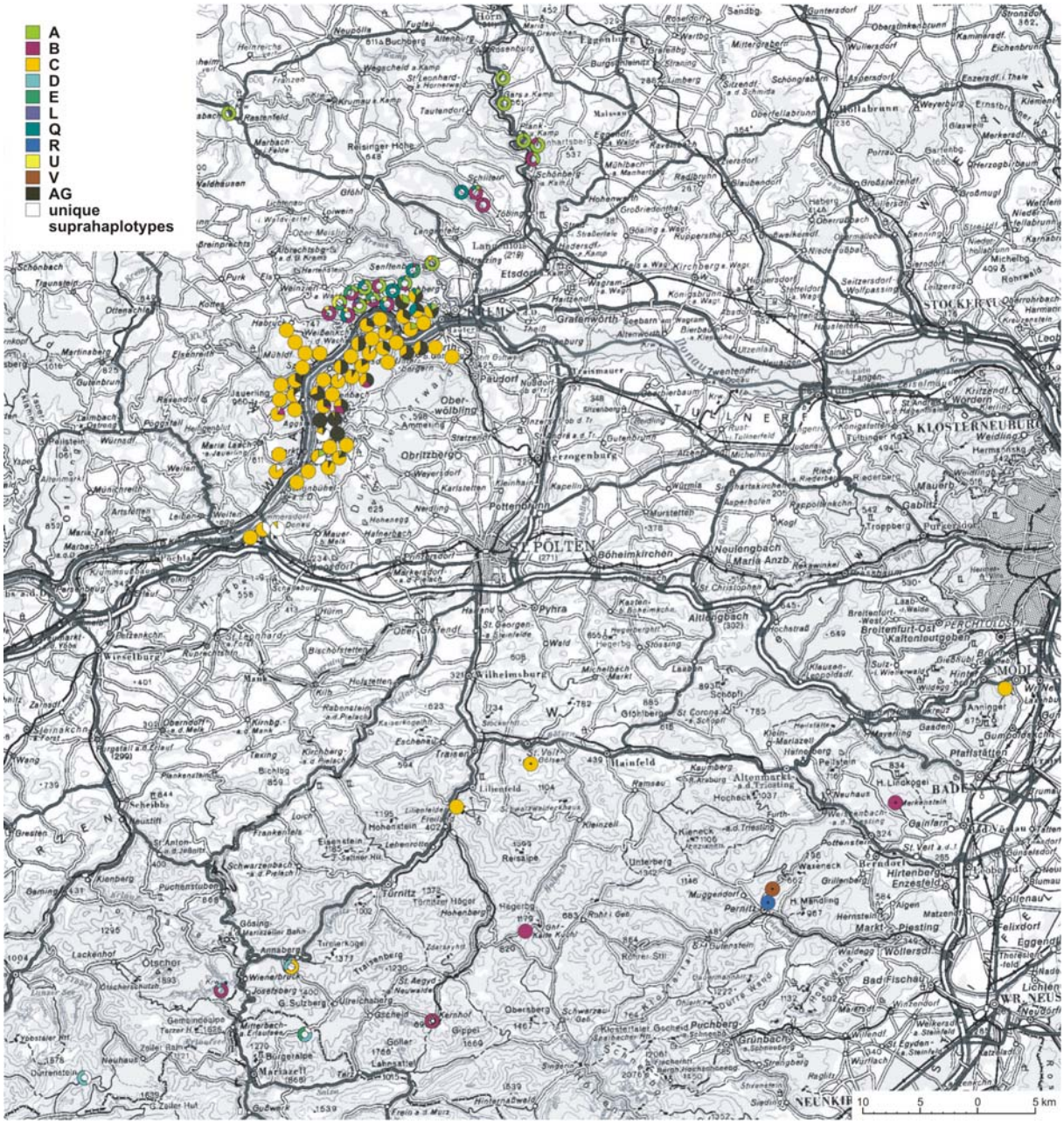


Fig. 11. Distribution of cpDNA *trn*L/F suprahaplotypes of *Arabidopsis arenosa* (rings) and *Arabidopsis lyrata* ssp. *petraea* (circles) in Eastern Austria (Eastern Alps, Wachau, Krems/Kamp Valley). The four diploid *Arabidopsis lyrata* ssp. *petraea* populations are marked with an asterisk. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

3.3.2. High genetic diversity within diploid and tetraploid Arabidopsis arenosa populations, low genetic diversity within diploid and tetraploid Arabidopsis lyrata ssp. petraea populations

Genetic diversity after Nei (1987) (TABLE 4) was approximately fourfold (He = 0.5021) in a combined dataset of diploid and tetraploid *Arabidopsis arenosa* in contrast to *Arabidopsis lyrata* ssp. *petraea* (He = 0.1252), also as a combination of both cytotypes. Interestingly, the diploid *Arabidopsis arenosa* populations from the Carpathians had the highest value of genetic diversity (He = 0.6268). Seven *trn*L/F suprahaplotypes were found throughout all investigated diploid *Arabidopsis arenosa* populations (Fig. 12A), although only A and E with a higher frequency (Fig. 12B). Three suprahaplotypes (BE, P, Y) were only found in Carpathian populations (Fig. 12C). Tetraploids of *Arabidopsis arenosa* showed reduced genetic diversity of He = 0.3773 (TABLE 4). Altogether nine suprahaplotypes were detected in tetraploid *Arabidopsis arenosa* (Fig. 12A), four of them with a higher frequency (A, B, E, Q) (Fig. 12B). Populations with only suprahaplotype A were found in the northernmost part of the Kamp Valley (Pop119, Pop124, Pop125), close to Zwettl (Pop402140), and in Upper Austria (Pop54) (Fig. 11). Furthermore, this suprahaplotype was the most common one in Central Europe and Scandinavia (this PhD thesis, second chapter), suggesting postglacial colonisation of individuals with this sequence type from refuge areas like the Wachau in Eastern Austria. Despite reduced genetic diversity in the Krems/Kamp Valley, most suprahaplotypes were shared between this region and the Wachau (A, B, C, E, Q) (Fig. 12C). Populations from the Eastern Alps were partly distinct, due to the frequent occurrence of suprahaplotypes D and AU, which were unique for this region (Fig. 12C). Genetic diversity in diploid *Arabidopsis lyrata* ssp. *petraea* was He = 0.0000 (TABLE 4). Only one suprahaplotype occurred in each of the four investigated populations (Fig. 11), leading to extremely low genetic diversity. However, these sequence types were all different (B, C, R, V) (Fig. 11, 12D). This strong genetic differentiation is probably due to long-term geographic isolation of these populations, which occupy relict habitats on single, exposed rocks and rocky slopes within an otherwise forested area. Nowadays, a decline of population size of numerous diploid populations can be observed, as only few individuals are found (personal observation), which could contribute to a reduction of geneflow between the remaining populations. Tetraploid *Arabidopsis lyrata* ssp. *petraea* showed increased genetic diversity of He = 0.2504 (TABLE 4), due to several singletons in the Wachau (AH, AI, AJ, AK, AL) and the additional suprahaplotype AC (Fig. 10, 11, 12D). However, genetic diversity was strongly reduced in

comparison to *Arabidopsis arenosa*, as especially suprahaplotype C was extremely frequent. Consequently, from the ten suprahaplotypes found in tetraploid *Arabidopsis lyrata* ssp. *petraea*, only two, C and AC, were mainly found in the populations (Fig. 12B). Diversity statistics according to Nei (1987) was used in this analysis, as *Arabidopsis arenosa* sequences were mostly separated by only one mutational step in contrast to four mutational steps between *Arabidopsis lyrata* ssp. *petraea* suprahaplotypes C and AC. Nucleotide diversity would have been biased towards an increased diversity of *Arabidopsis lyrata* ssp. *petraea* over *Arabidopsis arenosa*, which does not reflect the actual situation.

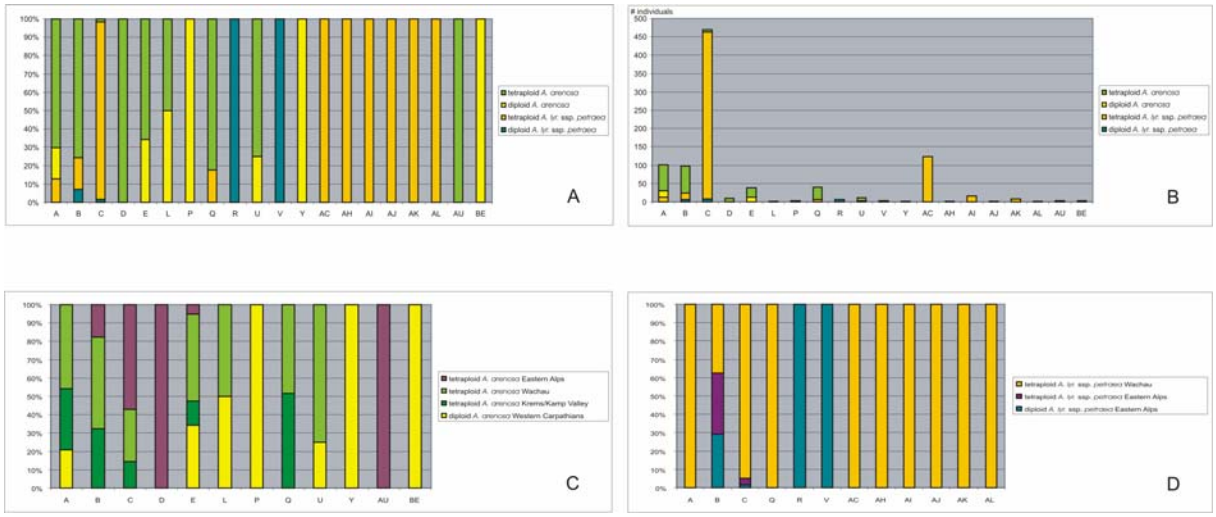


Fig. 12. Frequencies and numbers of cpDNA *trn*L/F suprahaplotypes of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*). A. Suprahaplotype frequencies of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. B. Suprahaplotype numbers of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. C. Suprahaplotype frequencies of *Arabidopsis arenosa* in four geographic regions (Eastern Alps, Wachau, Krems/KampValley, Western Carpathians). D. Suprahaplotype frequencies of *Arabidopsis lyrata* ssp. *petraea* in two geographic regions (Eastern Alps, Wachau).

Species	Geographic region	Population number	He (Nei)
diploid *A. arenosa*	Western Carpathians	126	0.7500 +/- 0.1391
diploid *A. arenosa*	Western Carpathians	127	0.8333 +/- 0.2224
diploid *A. arenosa*	Western Carpathians	128	0.3333 +/- 0.2152
diploid *A. arenosa*	Western Carpathians	131	0.6000 +/- 0.1753
diploid *A. arenosa*	Western Carpathians	915140	0.7333 +/- 0.1005
diploid *A. arenosa*	Western Carpathians	915141	0.5111 +/- 0.1643
tetraploid *A. arenosa*	Krems/Kamp Valley	44	0.2857 +/- 0.1964
tetraploid *A. arenosa*	Krems/Kamp Valley	67	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Krems/Kamp Valley	89	0.5357 +/- 0.1232
tetraploid *A. arenosa*	Krems/Kamp Valley	100	0.4643 +/- 0.2000
tetraploid *A. arenosa*	Krems/Kamp Valley	101	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Krems/Kamp Valley	104	0.6212 +/- 0.0867
tetraploid *A. arenosa*	Krems/Kamp Valley	118	0.4000 +/- 0.2373
tetraploid *A. arenosa*	Krems/Kamp Valley	119	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Krems/Kamp Valley	124	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Krems/Kamp Valley	125	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Wachau	3	0.7417 +/- 0.0569
tetraploid *A. arenosa*	Wachau	4	0.6389 +/- 0.1258
tetraploid *A. arenosa*	Wachau	5	0.5238 +/- 0.2086
tetraploid *A. arenosa*	Wachau	6	0.7222 +/- 0.0967
tetraploid *A. arenosa*	Wachau	7	0.7302 +/- 0.0352
tetraploid *A. arenosa*	Wachau	19	0.0000 +/- 0.0000
tetraploid *A. arenosa*	Wachau	20	0.7778 +/- 0.0907
tetraploid *A. arenosa*	Wachau	21	0.6786 +/- 0.1220
tetraploid *A. arenosa*	Wachau	22	0.6000 +/- 0.1291
tetraploid *A. arenosa*	Wachau	23	0.2500 +/- 0.1802
tetraploid *A. arenosa*	Wachau	106	0.4396 +/- 0.1120
tetraploid *A. arenosa*	Eastern Alps	402140	0.8000 +/- 0.1640
tetraploid *A. arenosa*	Eastern Alps	402159	0.5714 +/- 0.1195
tetraploid *A. arenosa*	Eastern Alps	402168	0.2500 +/- 0.1802
tetraploid *A. arenosa*	Eastern Alps	402192	0.5333 +/- 0.1721
tetraploid *A. arenosa*	Eastern Alps	915142	0.0000 +/- 0.0000
diploid *A. lyr.* ssp. *petraea*	Eastern Alps	115	0.0000 +/- 0.0000
diploid *A. lyr.* ssp. *petraea*	Eastern Alps	120	0.0000 +/- 0.0000
diploid *A. lyr.* ssp. *petraea*	Eastern Alps	915143	0.0000 +/- 0.0000
diploid *A. lyr.* ssp. *petraea*	Eastern Alps	915145	0.0000 +/- 0.0000
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	116	0.0000 +/- 0.0000
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	402196	0.0000 +/- 0.0000
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	915144	0.0000 +/- 0.0000
tetraploid *A. lyr.* ssp. *petraea*	Wachau	1	0.7333 +/- 0.0764

Species	Geographic region	Population number	He (Nei)
tetraploid *A. lyr*. ssp. *petraea*	Wachau	9	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	10	0.5091 +/- 0.1008
tetraploid *A. lyr*. ssp. *petraea*	Wachau	11	0.5714 +/- 0.1195
tetraploid *A. lyr*. ssp. *petraea*	Wachau	12	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	13	0.3889 +/- 0.1644
tetraploid *A. lyr*. ssp. *petraea*	Wachau	14	0.4848 +/- 0.1059
tetraploid *A. lyr*. ssp. *petraea*	Wachau	15	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	16	0.7455 +/- 0.0978
tetraploid *A. lyr*. ssp. *petraea*	Wachau	17	0.7143 +/- 0.0568
tetraploid *A. lyr*. ssp. *petraea*	Wachau	18	0.5641 +/- 0.1117
tetraploid *A. lyr*. ssp. *petraea*	Wachau	24	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	25	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	26	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	27	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	28	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	29	0.5000 +/- 0.1283
tetraploid *A. lyr*. ssp. *petraea*	Wachau	30	0.7143 +/- 0.1809
tetraploid *A. lyr*. ssp. *petraea*	Wachau	31	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	32	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	33	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	34	0.4476 +/- 0.1345
tetraploid *A. lyr*. ssp. *petraea*	Wachau	35	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	36	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	37	0.5217 +/- 0.0301
tetraploid *A. lyr*. ssp. *petraea*	Wachau	39	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	41	0.2451 +/- 0.1126
tetraploid *A. lyr*. ssp. *petraea*	Wachau	42	0.4762 +/- 0.1713
tetraploid *A. lyr*. ssp. *petraea*	Wachau	43	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	45	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	46	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	47	0.3333 +/- 0.2152
tetraploid *A. lyr*. ssp. *petraea*	Wachau	48	0.8000 +/- 0.1640
tetraploid *A. lyr*. ssp. *petraea*	Wachau	49	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	50	0.6000 +/- 0.1291
tetraploid *A. lyr*. ssp. *petraea*	Wachau	51	0.1818 +/- 0.1436
tetraploid *A. lyr*. ssp. *petraea*	Wachau	52	0.2500 +/- 0.1802
tetraploid *A. lyr*. ssp. *petraea*	Wachau	53	0.2000 +/- 0.1541
tetraploid *A. lyr*. ssp. *petraea*	Wachau	55	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	56	0.5357 +/- 0.1232
tetraploid *A. lyr*. ssp. *petraea*	Wachau	57	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	58	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	59	0.8000 +/- 0.1640

Species	Geographic region	Population number	He (Nei)
tetraploid *A. lyr*. ssp. *petraea*	Wachau	60	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	61	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	62	0.7121 +/- 0.0691
tetraploid *A. lyr*. ssp. *petraea*	Wachau	63	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	64	0.5303 +/- 0.1359
tetraploid *A. lyr*. ssp. *petraea*	Wachau	65	0.4667 +/- 0.1318
tetraploid *A. lyr*. ssp. *petraea*	Wachau	68	0.4762 +/- 0.0920
tetraploid *A. lyr*. ssp. *petraea*	Wachau	69	0.1667 +/- 0.1343
tetraploid *A. lyr*. ssp. *petraea*	Wachau	70	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	71	0.3846 +/- 0.1494
tetraploid *A. lyr*. ssp. *petraea*	Wachau	72	0.0000 +/- 0.0000
tetraploid *A. lyr*. ssp. *petraea*	Wachau	73	0.4727 +/- 0.1617
tetraploid *A. lyr*. ssp. *petraea*	Wachau	75	0.3333 +/- 0.2152
tetraploid *A. lyr*. ssp. *petraea*	Wachau	102	0.2941 +/- 0.1193
tetraploid *A. lyr*. ssp. *petraea*	Wachau	103	0.1176 +/- 0.1012

TABLE 4. Genetic diversity after Nei (1987) He +/- standard deviation, based on cpDNA *trn*L/F suprahaplotypes, of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*).

Microsatellites

3.3.3. Summary statistics of the seven microsatellite loci

According to their repeat type, the microsatellite markers could be grouped into two classes (TABLE 5): (1) Perfect repeat type with continuous progression of repeat units (ATTS0392, trinucleotide repeat; ICE13, trinucleotide repeat; NGA162, dinucleotide repeat; ICE14, trinucleotide repeat; AthZFPG, dinucleotide repeat) (2) Imperfect repeat type with discontinuous progression of repeat units (SLL2, dinucleotide repeat; ICE7, trinucleotide repeat). Due to their degree of allelic polymorphism, the seven microsatellite loci could be arranged in three classes (TABLE 5, Fig. 13): (1) High degree of allelic polymorphism (ATTS0392, ICE13, AthZFPG). Throughout the whole dataset of diploids and tetraploids fragments ranged from 126 bp to 198 bp in ATTS0392, 197 bp to 272 bp in ICE13, and 120 bp to 206 bp in AthZFPG. Considering only tetraploids, around 20 alleles were found altogether. (2) Medium degree of allelic polymorphism (ICE7, NGA162, ICE14). Here alleles were detected from 79 bp to 142 bp in ICE7, from 75 bp to 99 bp in NGA162, and from 206 bp to 260 bp in ICE14. In tetraploids approximately 10 alleles were found altogether. (3) Low degree of allelic polymorphism (SLL2). Fragment sizes in SLL2 ranged from 263 bp to 313 bp, and only around 5 alleles were observed in tetraploids altogether. Each tetraploid individual had, at least for one of the highly polymorphic markers, four different alleles, confirming tetraploidy. Chromosome counts in flower buds (Baldauf, 2006) had already indicated tetraploidy for all Wachau populations, but only five individuals per population were measured on average. Summing up all seven SSR markers for each species and cytotype, the following allele numbers could be found: 72 different alleles in 122 accessions of diploid *Arabidopsis arenosa*, 97 different alleles in 431 accessions of tetraploid *Arabidopsis arenosa*, 49 different alleles in 99 accessions of diploid *Arabidopsis lyrata* ssp. *petraea*, and 102 different alleles in 1118 accessions of tetraploid *Arabidopsis lyrata* ssp. *petraea*. The vast majority of alleles in diploids was also found in tetraploids of the same species: all in SLL2, ATTS0392, and ICE14, 89% in NGA162, 86% in ICE13 and AthZFPG, and 80% in ICE7. New alleles in tetraploids mainly had very small or large fragment sizes, e.g. in ICE13 and ATTS0392, indicating increased mutation rates in SSR regions of polyploids. Allele sharing was not only observed between diploids and tetraploids of the same species, but also between diploids of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*: SLL2 – 100% in *Arabidopsis arenosa*, 50% in *Arabidopsis lyrata* ssp. *petraea*, ICE13 – 43% in *Arabidopsis arenosa*, 86% in *Arabidopsis lyrata* ssp. *petraea*; ATTS0392 – 64% in *Arabidopsis arenosa*,

90% in *Arabidopsis lyrata* ssp. *petraea*; AthZFPG – 57% in *Arabidopsis arenosa*, 80% in *Arabidopsis lyrata* ssp. *petraea*; NGA162 – 44% in *Arabidopsis arenosa*, 100% in *Arabidopsis lyrata* ssp. *petraea*; ICE14 – 71% in *Arabidopsis arenosa*, 83% in *Arabidopsis lyrata* ssp. *petraea*; ICE7 – 40% in *Arabidopsis arenosa*, 67% in *Arabidopsis lyrata* ssp. *petraea*. Summarised across all seven microsatellite markers, around 60% alleles were shared in diploid *Arabidopsis arenosa* and approximately 80% in diploid *Arabidopsis lyrata* ssp. *petraea*. The percentage of allele sharing was similar between tetraploids of both species: Altogether around 80% alleles were shared both in tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. Allele sharing impeded the detection of hybrid individuals, as there was only a very low number of species-specific alleles, which could clearly been allocated to one of the parental genomes in a hybrid individual. However, allele frequencies were significantly different between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* (Fig. 14), and, consequently, hybrids could be detected by intermediate allele frequencies. The most significant difference in allele frequencies between the two species was found in the marker SLL2. Here the frequency of allele 313 bp reached approximately 100% in *Arabidopsis arenosa*, of allele 303 bp around 90% in *Arabidopsis lyrata* ssp. *petraea*. Therefore, SLL2 could be used as a species-specific marker. Also in other SSR markers allele frequencies were different between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, e.g. in NGA162 with 20% frequency of allele 79 bp in *Arabidopsis arenosa* and around 80% in *Arabidopsis lyrata* ssp. *petraea*.

SLL2	ICE13	ATTS0392	AthZFPG	NGA162	ICE14	ICE7
263 bp	197 bp	126 bp	120 bp	75 bp	206 bp	79 bp
273 bp	209 bp	130 bp	124 bp	77 bp	209 bp	82 bp
289 bp	212 bp	133 bp	126 bp	79 bp	218 bp	94 bp
295 bp	215 bp	136 bp	128 bp	81 bp	221 bp	97 bp
303 bp	218 bp	139 bp	132 bp	83 bp	224 bp	100 bp
305 bp	221 bp	142 bp	134 bp	85 bp	227 bp	105 bp
307 bp	224 bp	145 bp	136 bp	87 bp	230 bp	109 bp
309 bp	227 bp	148 bp	138 bp	89 bp	233 bp	118 bp
311 bp	230 bp	151 bp	140 bp	91 bp	236 bp	123 bp
313 bp	233 bp	154 bp	142 bp	93 bp	239 bp	126 bp
	236 bp	157 bp	144 bp	95 bp	242 bp	130 bp
	239 bp	160 bp	146 bp	97 bp	245 bp	134 bp
	242 bp	163 bp	148 bp	99 bp	248 bp	139 bp
	245 bp	165 bp	150 bp		260 bp	142 bp
	248 bp	168 bp	152 bp			
	251 bp	171 bp	154 bp			
	254 bp	174 bp	156 bp			
	257 bp	177 bp	158 bp			
	260 bp	180 bp	160 bp			
	263 bp	183 bp	162 bp			
	266 bp	186 bp	164 bp			
	269 bp	188 bp	166 bp			
	272 bp	198 bp	168 bp			
			170 bp			
			172 bp			
			174 bp			
			176 bp			
			178 bp			
			180 bp			
			198 bp			
			200 bp			
			202 bp			
			206 bp			

TABLE 5. List of alleles found at seven microsatellite loci throughout the complete dataset of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*). Perfect repeats are observed in ICE13, ATTS0392, AthZFPG, NGA162, and ICE14, imperfect repeats are predominantly found in SLL2 and ICE 7.

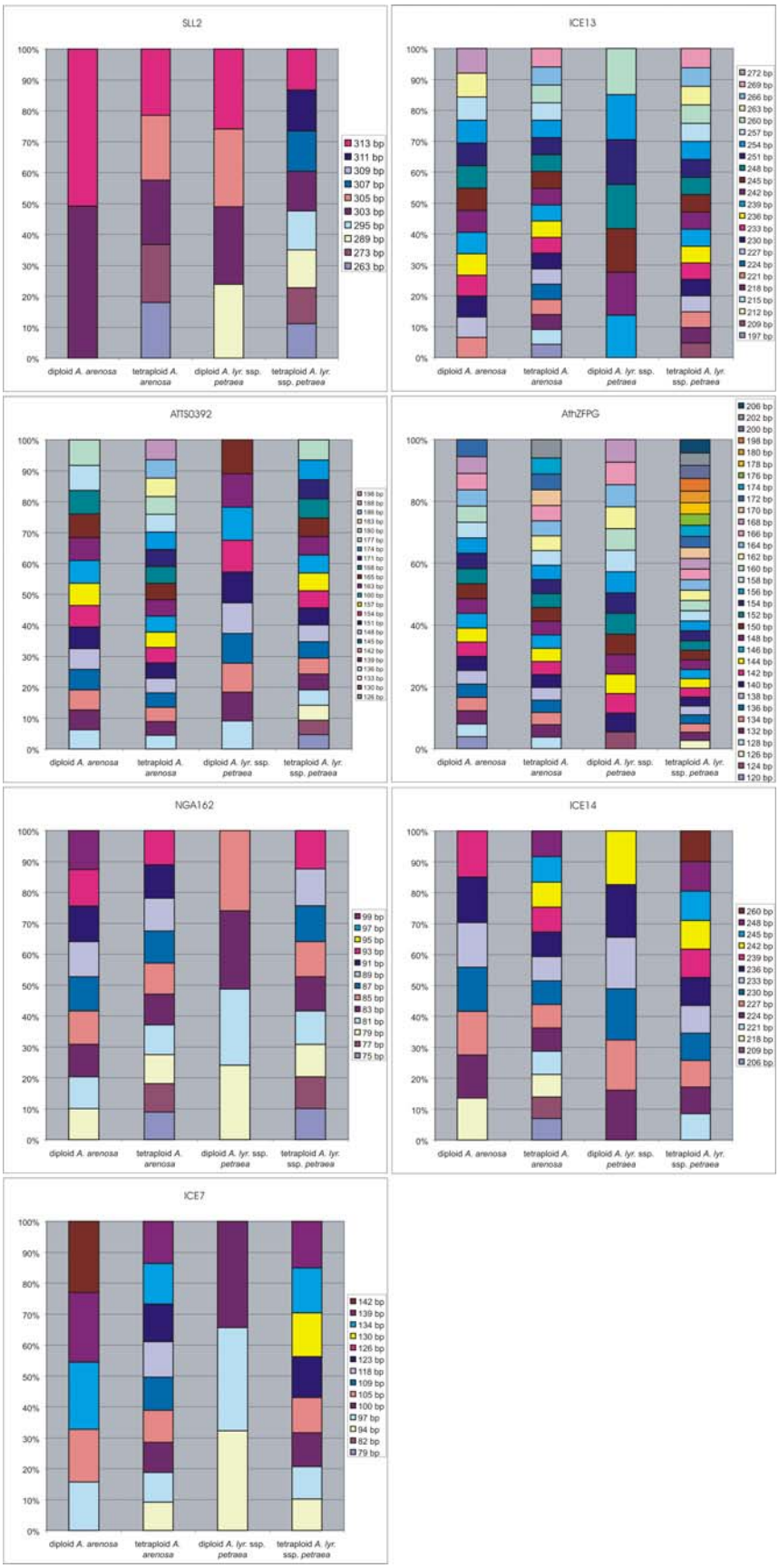


Fig. 13. Allele numbers of seven microsatellite loci of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*).

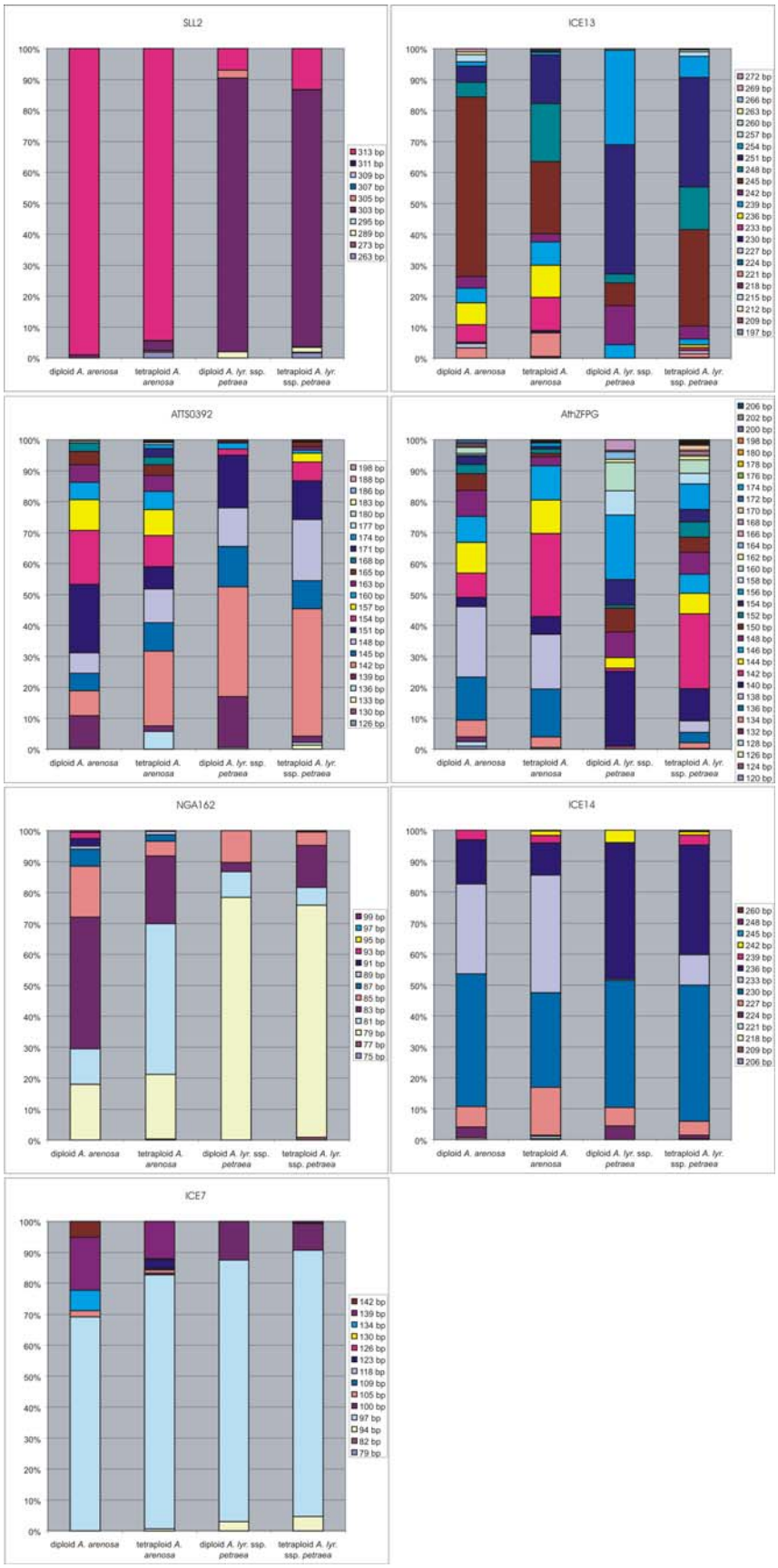


Fig. 14. Allele frequencies of seven microsatellite loci of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*).

3.3.4. Bayesian Analyses reveal four main genetic groups: diploid and tetraploid Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Based on Bayesian Algorithm with various data subsets, a mainly clear differentiation between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* was observed (Fig. 15). Group clustering with K = 2 detected one group of diploid and tetraploid *Arabidopsis arenosa* and a second group of diploid and tetraploid *Arabidopsis lyrata* ssp. *petraea*. Within the *Arabidopsis lyrata* ssp. *petraea* group one outlier was found (Pop402196), indicating an interspecies hybrid population. However, increased substructuring of the dataset was gained from Bayesian Analyses with increasing K values. As the right K value for the mixed dataset of diploids and tetraploids could not be calculated unequivocally, results of K = 3, 4, 5, 9 will be discussed. At K = 9 the log(ml) graph plateaus (data not shown). Group clustering of K = 3 revealed a shared group of diploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, indicating difficulties of the algorithm to properly handle a mixed dataset of diploids and tetraploids. With K = 4 formation of the four main groups could already be observed, comprising diploid *Arabidopsis arenosa* from the Western Carpathians, tetraploid *Arabidopsis arenosa*, and diploid and tetraploid *Arabidopsis lyrata* ssp. *petraea*, all from Eastern Austria. With increasing K values (K = 5, 9) tetraploid *Arabidopsis arenosa* population Pop915142 formed a distinct cluster, probably due to hybridisation with *Arabidopsis lyrata* ssp. *petraea*, as further indicated by additional analyses. Moreover, the group of diploid *Arabidopsis lyrata* ssp. *petraea* showed substructuring: Within the four Eastern Austrian populations three distinct genetic clusters could be differentiated, with the geographically neighbouring populations Pop120 and Pop915143 forming one group.

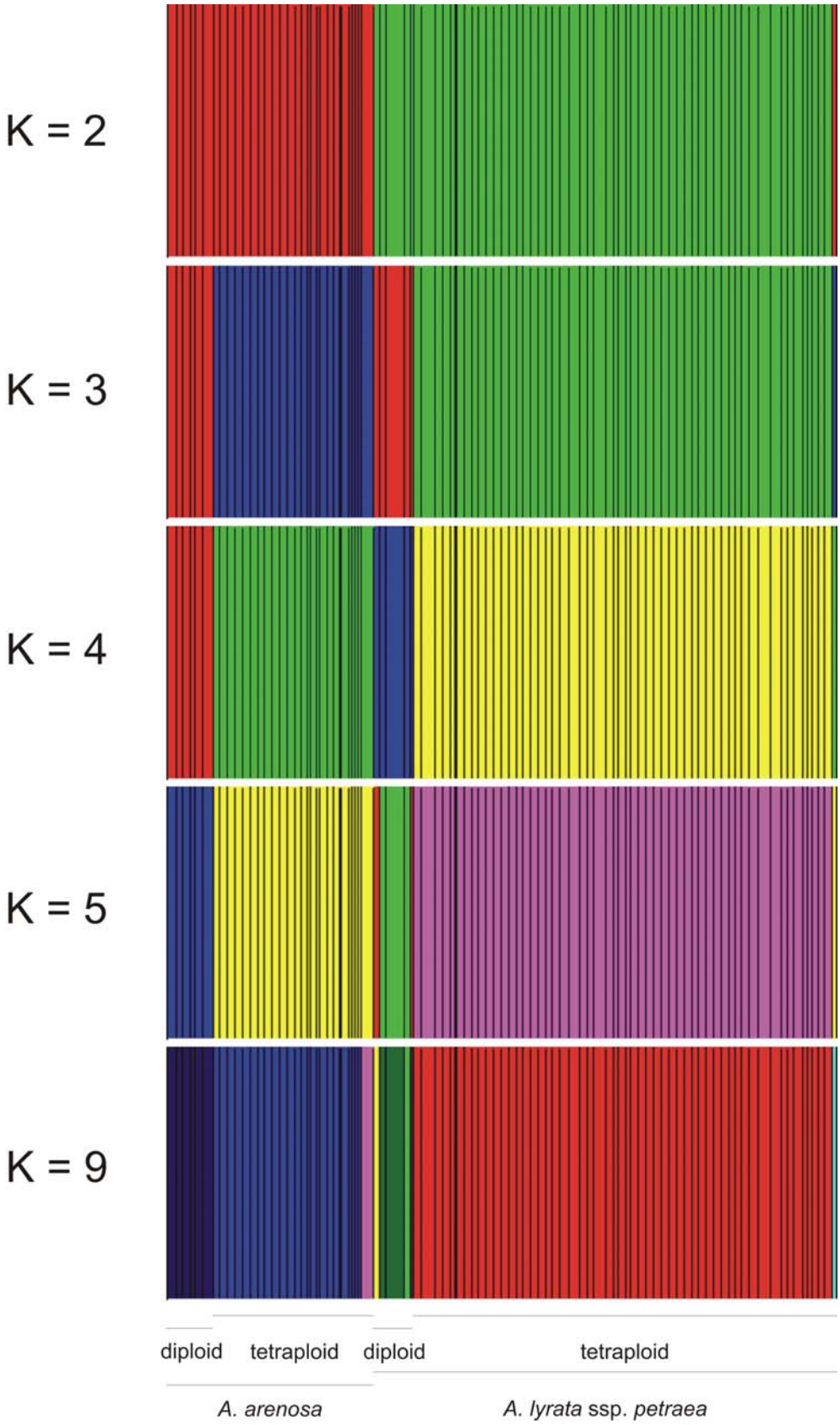


Fig. 15. Population structure estimated by BAPS analyses of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*), based on seven microsatellite loci. Calculations were performed with different K values (K = 2, 3, 4, 5, 9), which correspond to the numbers of ancestral populations. Each population is represented by a vertical bar.

In a separate analysis of exclusively diploid populations of various taxonomic ranking with K = 14 (Fig. 16) each of the four diploid *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria resembled a distinct genetic cluster (Pop115, Pop120, Pop915143, Pop915145). Thus, diploid *Arabidopsis lyrata* ssp. *petraea* is characterised by strong genetic differentiation between the populations, probably due to long-term geographic isolation during Holocene warming. The populations were restricted to single, exposed rocks and rocky slopes within a mainly forested landscape. Long-term geographic isolation might also be the explanation for the separate clustering of diploids from Czech Republic (Pop96) and Franconian Switzerland/Germany (Pop112, Pop133). In contrast, diploid *Arabidopsis arenosa* from the Carpathians was a well-defined group. The six investigated populations formed only one cluster in an analysis of exclusively diploids. Additional diploid *Arabidopsis arenosa* taxa from the Carpathians, *Arabidopsis neglecta* and *Arabidopsis petrogena* nom. prov., formed distinct genetic groups.

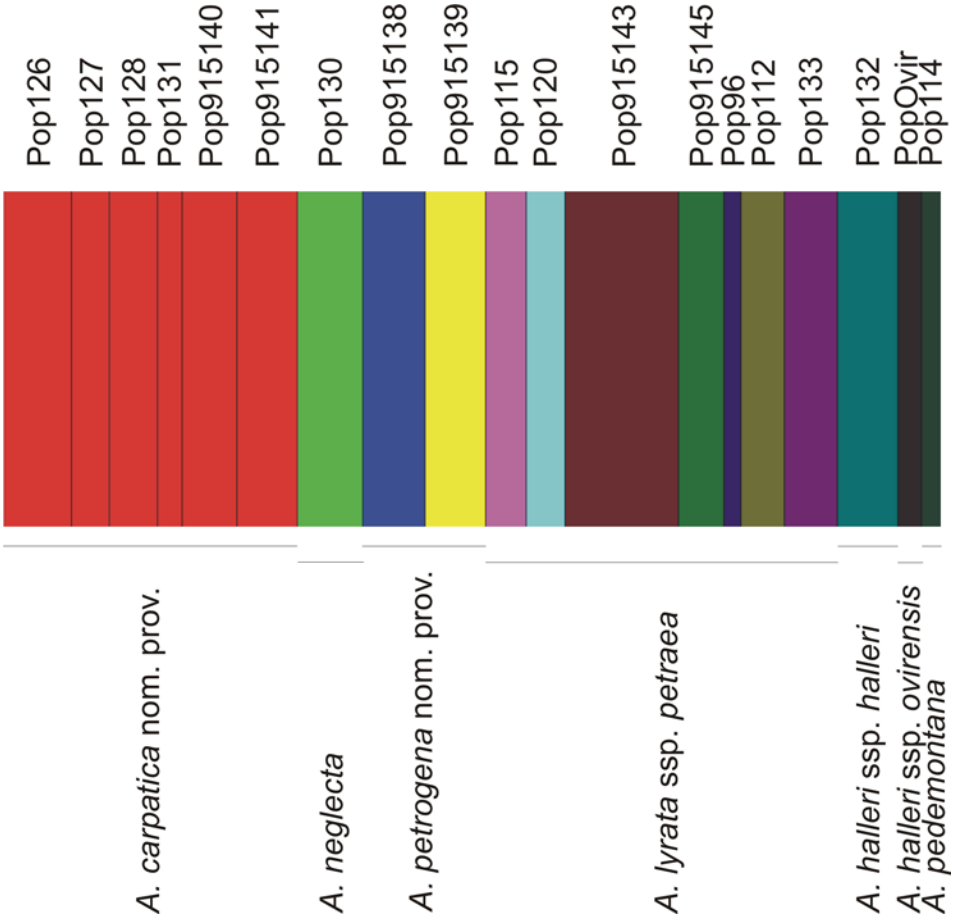


Fig. 16. Population structure estimated by BAPS analysis of exclusively diploid *Arabidopsis* populations of various taxonomic ranking (*Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis petrogena* nom. prov., *Arabidopsis lyrata* ssp. *petraea*, *Arabidopsis halleri* ssp. *halleri*, *Arabidopsis halleri* ssp. *ovirensis*, *Arabidopsis pedemontana*), based on seven microsatellite loci. Calculations were performed with K = 14.

3.3.5. Population genetic statistics of diploids across the genus Arabidopsis

In total 19 diploid populations were analysed (TABLE 6): 9 populations from the *Arabidopsis arenosa* species complex, all from the Carpathians (6x *Arabidopsis carpatica* nom. prov., 1x *Arabidopsis neglecta*, 2x *Arabidopsis petrogena* nom. prov.); 7 *Arabidopsis lyrata* ssp. *petraea* populations (4x from Eastern Austria, 1x from Czech Republic, 2x from Franconian Switzerland/Germany); 2 populations from the *Arabidopsis halleri* species complex (ssp. *halleri* from Sauerland/Germany, ssp. *ovirensis* from Carinthia/Austria); 1 *Arabidopsis pedemontana* population. Mean gene diversity He over all seven microsatellite loci ranged in members of the *Arabidopsis arenosa* complex from He = 0.540 to 0.640 (TABLE 6), with an average of He = 0.591 (TABLE 7), in *Arabidopsis lyrata* ssp. *petraea* from He = 0.308 to 0.538 (TABLE 6), with an average of He = 0.442 (TABLE 7), and in *Arabidopsis halleri*/*Arabidopsis pedemontana* from He = 0.214 to 0.305 (TABLE 6), with an average of He = 0.264 (TABLE 7). Thus, gene diversity was highest in *Arabidopsis arenosa*, modest in *Arabidopsis lyrata*, and lowest in *Arabidopsis halleri*/*Arabidopsis pedemontana*. As *Arabidopsis halleri* and *Arabidopsis pedemontana* both propagate clonally to a large extent, this could, at least partly, explain the low gene diversity in those two species. Genetic bottlenecks, resulting from high endemism of *Arabidopsis halleri* ssp. *ovirensis* and *Arabidopsis pedemontana*, could additionally have contributed to low gene diversity. A similar tendency was observed for allelic richness. Mean allelic richness (TABLE 7) in *Arabidopsis arenosa* was A = 3.780, in *Arabidopsis lyrata* A = 2.719, and in *Arabidopsis halleri*/*Arabidopsis pedemontana* A = 1.680. Inbreeding coefficient for the three species groups was close to zero (TABLE 7), indicating mainly outcrossing: F_{IS} = 0.076 in *Arabidopsis arenosa*, F_{IS} = 0.012 in *Arabidopsis lyrata* ssp. *petraea*, and F_{IS} = -0.332 in *Arabidopsis halleri*/*Arabidopsis pedemontana*. Observed heterozygosity (TABLE 7) was highest in *Arabidopsis arenosa* (Ho = 0.542), followed by *Arabidopsis lyrata* ssp. *petraea* (Ho = 0.443), and *Arabidopsis halleri*/*Arabidopsis pedemontana* (Ho = 0.360). In *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* values of observed and expected heterozygosity were nearly the same (TABLE 7): Ho = 0.542/He = 0.586 in *Arabidopsis arenosa*, Ho = 0.443/He = 0.449 in *Arabidopsis lyrata* ssp. *petraea*. Populations of those two species were in Hardy-Weinberg Equilibrium. Probably due to clonal propagation, observed versus expected heterozygosity values significantly diverged in *Arabidopsis halleri*/*Arabidopsis pedemontana* (Ho = 0.360/He = 0.270).

Species	Population number	He (Nei)	A	Total allele number	F_{IS}
A. arenosa [*A. carpatica* nom. prov.]	126	0.605	4.180	55	0.157
A. arenosa [*A. carpatica* nom. prov.]	127	0.629	4.118	42	0.109
A. arenosa [*A. carpatica* nom. prov.]	128	0.604	3.923	42	0.050
A. arenosa [*A. carpatica* nom. prov.]	131	0.573	3.768	33	0.182
A. arenosa [*A. carpatica* nom. prov.]	915140	0.640	4.069	42	0.235
A. arenosa [*A. carpatica* nom. prov.]	915141	0.576	3.788	45	0.002
A. neglecta	130	0.575	3.619	42	0.009
A. petrogena nom. prov.	915138	0.540	3.061	29	-0.012
A. petrogena nom. prov.	915139	0.575	3.493	34	0.046
A. lyr. ssp. *petraea*	115	0.308	2.087	17	-0.141
A. lyr. ssp. *petraea*	120	0.538	2.872	25	0.007
A. lyr. ssp. *petraea*	915143	0.447	2.790	31	-0.033
A. lyr. ssp. *petraea*	915145	0.515	2.872	26	0.111
A. lyr. ssp. *petraea*	96	0.450	2.611	20	0.057
A. lyr. ssp. *petraea*	112	0.526	3.217	29	0.048
A. lyr. ssp. *petraea*	133	0.312	2.583	23	0.152
A. halleri ssp. *halleri*	132	0.272	1.697	13	-0.288
A. halleri ssp. *ovirensis*	Ovir	0.305	1.915	14	-0.042
A. pedemontana	114	0.214	1.429	10	-1.000

TABLE 6. Multilocus statistics of seven microsatellite loci of exclusively diploid *Arabidopsis* populations of various taxonomi**c** ranking (*Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, *Arabidopsis petrogena* nom. prov., *Arabidopsis lyrata* ssp. *petraea*, *Arabidopsis halleri* ssp. *halleri*, *Arabidopsis halleri* ssp. *ovirensis*, *Arabidopsis pedemontana*). Genetic diversity after Nei (1987) He, allelic richness A, total allele number, and inbreeding coefficient F_{IS}.

Taxonomic group	Number of populations	He (Nei)	A	F_{IS}	Ho	He
A. arenosa (*A. carpatica* nom. prov., *A. neglecta*, *A. petrogena* nom. prov.)	9	0.591	3.780	0.076	0.542	0.586
A. lyr. ssp. *petraea*	7	0.442	2.719	0.012	0.443	0.449
A. halleri / *A. pedemontana*	3	0.264	1.680	-0.332	0.360	0.270

TABLE 7. Multilocus statistics of seven microsatellite loci of exclusively diploid *Arabidopsis* populations of various taxonomic ranking, assembled in three taxonomic groups: *Arabidopsis arenosa* group including *Arabidopsis carpatica* nom. prov., *Arabidopsis neglecta*, and *Arabidopsis petrogena* nom. prov., *Arabidopsis lyrata* group comprising only *Arabidopsis lyrata* ssp. *petraea*, and *Arabidopsis halleri* / *Arabidopsis pedemontana* group with *Arabidopsis halleri* ssp. *halleri*, *Arabidopsis halleri* ssp. *ovirensis*, and *Arabidopsis pedemontana*. Genetic diversity after Nei (1987) He, allelic richness A, inbreeding coefficient F_{IS}, observed heterozygosity Ho, and expected heterozygosity He.

3.3.6. Bayesian Test rejects both pure disomic and tetrasomic inheritance in tetraploid Arabidopsis lyrata ssp. petraea

Population genetic statistics in polyploids is dependent on the segregation pattern of alleles. Autotetraploids are normally characterised by tetrasomic inheritance, allopolyploids by disomic inheritance. To test these two modes of inheritance is, therefore, the prerequisite for further statistical analyses of e.g. gene diversity, allelic richness, and heterozygosity. For the majority of investigated populations, disomic inheritance was slightly favored over tetrasomic inheritance, supported by positive Bayes factors (TABLE 8). Marginal likelihood values for disomic inheritance were insignificantly higher than for tetrasomic inheritance, e.g. -0.58 versus -2.35 for Pop1. Across the 37 analysed populations favoring disomic inheritance the total Bayes factor was 5.05, which was very low compared to a total Bayes factor of 65.09 in *Bordera chouardii* and 82.90 in *Bordera pyrenaica* (Catalán et al., 2006). Either the Bayesian Test was not significant, due to marker choice, or tetraploid *Arabidopsis lyrata* ssp. *petraea* showed a mixed di- and tetrasomic inheritance pattern. Such a mixed pattern could resemble an autopolyploid, which underwent genetic diploidisation.

Species	Geographic region	Population number	D $\text{Log}_e\ p(y \mid H_x)$	T $\text{Log}_e\ p(y \mid H_x)$	2 Log_e $(A_{allo,auto})$
tetraploid *A. lyr.* ssp. *petraea*	Wachau	1	-2.35	-0.58	3.54
tetraploid *A. lyr.* ssp. *petraea*	Wachau	9	-7.57	-3.07	9.00
tetraploid *A. lyr.* ssp. *petraea*	Wachau	10	-3.13	-0.79	4.68
tetraploid *A. lyr.* ssp. *petraea*	Wachau	11	-3.97	-0.61	6.72
tetraploid *A. lyr.* ssp. *petraea*	Wachau	12	-3.72	-1.14	5.17
tetraploid *A. lyr.* ssp. *petraea*	Wachau	13	-1.32	-4.11	-
tetraploid *A. lyr.* ssp. *petraea*	Wachau	14	-3.59	-2.27	2.64
tetraploid *A. lyr.* ssp. *petraea*	Wachau	15	-3.37	-2.24	2.28
tetraploid *A. lyr.* ssp. *petraea*	Wachau	17	-4.00	-1.30	5.40
tetraploid *A. lyr.* ssp. *petraea*	Wachau	18	-2.93	-0.23	5.41
tetraploid *A. lyr.* ssp. *petraea*	Wachau	24	-3.19	-4.41	-
tetraploid *A. lyr.* ssp. *petraea*	Wachau	25	-7.99	-10.49	-
tetraploid *A. lyr.* ssp. *petraea*	Wachau	26	-0.90	1.65	5.10
tetraploid *A. lyr.* ssp. *petraea*	Wachau	27	-2.19	0.19	4.75
tetraploid *A. lyr.* ssp. *petraea*	Wachau	28	-1.76	1.01	5.55
tetraploid *A. lyr.* ssp. *petraea*	Wachau	29	-3.80	-1.91	3.78
tetraploid *A. lyr.* ssp. *petraea*	Wachau	30	-3.49	-0.16	6.67
tetraploid *A. lyr.* ssp. *petraea*	Wachau	31	-0.34	2.20	5.08
tetraploid *A. lyr.* ssp. *petraea*	Wachau	34	-3.56	-4.82	-
tetraploid *A. lyr.* ssp. *petraea*	Wachau	35	-2.50	0.07	5.16
tetraploid *A. lyr.* ssp. *petraea*	Wachau	37	-2.23	-0.33	3.79
tetraploid *A. lyr.* ssp. *petraea*	Wachau	39	-2.76	-0.49	4.55
tetraploid *A. lyr.* ssp. *petraea*	Wachau	41	-2.56	-4.03	-

Species	Geographic region	Population number	D $\text{Log}_e\ p(y \mid H_x)$	T $\text{Log}_e\ p(y \mid H_x)$	2 Log_e ($A_{allo,auto}$)
tetraploid *A. lyr.* ssp. *petraea*	Wachau	42	-1.72	1.37	6.18
tetraploid *A. lyr.* ssp. *petraea*	Wachau	46	-0.45	2.13	5.18
tetraploid *A. lyr.* ssp. *petraea*	Wachau	48	-2.65	0.30	5.91
tetraploid *A. lyr.* ssp. *petraea*	Wachau	53	-10.18	-9.46	1.46
tetraploid *A. lyr.* ssp. *petraea*	Wachau	57	-1.47	1.41	5.76
tetraploid *A. lyr.* ssp. *petraea*	Wachau	59	-7.60	-3.79	7.61
tetraploid *A. lyr.* ssp. *petraea*	Wachau	60	-7.24	-4.67	5.14
tetraploid *A. lyr.* ssp. *petraea*	Wachau	62	-5.57	-3.35	4.44
tetraploid *A. lyr.* ssp. *petraea*	Wachau	64	-1.61	1.78	6.77
tetraploid *A. lyr.* ssp. *petraea*	Wachau	65	-4.88	-4.18	1.40
tetraploid *A. lyr.* ssp. *petraea*	Wachau	68	-4.92	-2.24	5.36
tetraploid *A. lyr.* ssp. *petraea*	Wachau	102	-1.23	1.55	5.56
tetraploid *A. lyr.* ssp. *petraea*	Wachau	103	-2.83	-0.64	4.38
tetraploid *A. lyr.* ssp. *petraea*	Wachau	107	-1.93	0.02	3.89
tetraploid *A. lyr.* ssp. *petraea*	Wachau	108	-1.95	0.07	4.03
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	85	-1.55	-3.06	-

TABLE 8. Bayesian Test after Catalán et al. (2006) of the SLL2 locus. Marginal probability of data {$\text{Log}_e\ p(y \mid H_x)$} obtained for the alternative hypotheses of allotetraploidy (D, disomic inheritance) and autotetraploidy (T, tetrasomic inheritance). Bayes factors {$2\ \text{Log}_e\ (A_{allo,auto}) = 2 \times (\text{Log}_e\ p(y \mid H_{allo}) - \text{Log}_e\ p(y \mid H_{auto}))$} are given for each tetraploid *Arabidopsis lyrata* ssp. *petraea* population. Investigated populations are mainly from the Wachau.

3.3.7. Calculation of heterozygosity based on tetrasomic inheritance supports tetrasomic inheritance for both Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Assumption of either disomic or tetrasomic inheritance accounts for different calculations of allele- and genotype frequencies. The Bayesian Test after Catalán et al. (2006) failed to detect the mode of inheritance, but summary statistics of the seven microsatellite loci favored autopolyploidy. With the following approach it was tested, if populations were in Hardy-Weinberg Equilibrium under the assumption of tetrasomic inheritance. Therefore, observed versus expected heterozygosity was calculated (TABLE 9). Tetraploid populations of both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* showed similar values for observed and expected heterozygosity, indicating Hardy-Weinberg Equilibrium. As selfing is mostly inhibited both within *Arabidopsis arenosa* (personal observation) and *Arabidopsis lyrata* ssp. *petraea*, and population sizes were relatively large, Hardy-Weinberg Equilibrium was expected. Hence, the reverse conclusion could be drawn, that tetrasomic inheritance is the correct mode of inheritance, suggesting autopolyploid origin of both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. Observed heterozygosity values in the herein investigated tetraploid *Arabidopsis arenosa* populations reached from Ho = 0.350 to 0.750, with an average of Ho = 0.575. In tetraploid *Arabidopsis lyrata* ssp. *petraea* populations values ranged from Ho = 0.328 to 0.825, with an average of Ho = 0.598. Consequently, heterozygosity in tetraploids was twice as much in contrast to diploids: According to our analyses, observed heterozygosity was Ho = 0.542 in diploid *Arabidopsis arenosa* from the Western Carpathians and Ho = 0.443 in diploid *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria, Czech Republic and Germany (TABLE 7). In diploid *Arabidopsis lyrata* ssp. *petraea* populations analysed by Ansell (2004) from Franconian Switzerland/Germany, Eastern Austria, and Scotland observed heterozygosity values ranged from Ho = 0.101 to 0.399, with an average Ho = 0.250.

Species	Geographic region	Population number	Ho	He
tetraploid *A. arenosa*	Wachau	3	0.536	0.520
tetraploid *A. arenosa*	Wachau	4	0.500	0.484
tetraploid *A. arenosa*	Wachau	6	0.550	0.543
tetraploid *A. arenosa*	Wachau	7	0.556	0.565
tetraploid *A. arenosa*	Wachau	21	0.500	0.629
tetraploid *A. arenosa*	Wachau	22	0.708	0.649
tetraploid *A. arenosa*	Wachau	23	0.607	0.632
tetraploid *A. arenosa*	Wachau	106	0.604	0.637
tetraploid *A. arenosa*	Krems/Kamp Valley	44	0.500	0.482
tetraploid *A. arenosa*	Krems/Kamp Valley	89	0.625	0.643
tetraploid *A. arenosa*	Krems/Kamp Valley	118	0.350	0.450
tetraploid *A. arenosa*	Krems/Kamp Valley	125	0.417	0.465
tetraploid *A. arenosa*	Eastern Alps	402140	0.750	0.717
tetraploid *A. arenosa*	Eastern Alps	402168	0.708	0.786
tetraploid *A. arenosa*	Eastern Alps	915142	0.713	0.796
Mean			0.579	0.600
tetraploid *A. lyr.* ssp. *petraea*	Wachau	1	0.733	0.726
tetraploid *A. lyr.* ssp. *petraea*	Wachau	9	0.524	0.565
tetraploid *A. lyr.* ssp. *petraea*	Wachau	10	0.500	0.476
tetraploid *A. lyr.* ssp. *petraea*	Wachau	11	0.583	0.585
tetraploid *A. lyr.* ssp. *petraea*	Wachau	12	0.643	0.650
tetraploid *A. lyr.* ssp. *petraea*	Wachau	13	0.583	0.677
tetraploid *A. lyr.* ssp. *petraea*	Wachau	14	0.719	0.713
tetraploid *A. lyr.* ssp. *petraea*	Wachau	15	0.750	0.752
tetraploid *A. lyr.* ssp. *petraea*	Wachau	16	0.825	0.824
tetraploid *A. lyr.* ssp. *petraea*	Wachau	17	0.700	0.663
tetraploid *A. lyr.* ssp. *petraea*	Wachau	18	0.591	0.625
tetraploid *A. lyr.* ssp. *petraea*	Wachau	24	0.596	0.669
tetraploid *A. lyr.* ssp. *petraea*	Wachau	25	0.554	0.681
tetraploid *A. lyr.* ssp. *petraea*	Wachau	26	0.682	0.641
tetraploid *A. lyr.* ssp. *petraea*	Wachau	27	0.625	0.642
tetraploid *A. lyr.* ssp. *petraea*	Wachau	28	0.604	0.547
tetraploid *A. lyr.* ssp. *petraea*	Wachau	29	0.518	0.565
tetraploid *A. lyr.* ssp. *petraea*	Wachau	30	0.671	0.666
tetraploid *A. lyr.* ssp. *petraea*	Wachau	31	0.500	0.486
tetraploid *A. lyr.* ssp. *petraea*	Wachau	32	0.609	0.670
tetraploid *A. lyr.* ssp. *petraea*	Wachau	34	0.656	0.639
tetraploid *A. lyr.* ssp. *petraea*	Wachau	35	0.607	0.630
tetraploid *A. lyr.* ssp. *petraea*	Wachau	36	0.472	0.447
tetraploid *A. lyr.* ssp. *petraea*	Wachau	37	0.617	0.569
tetraploid *A. lyr.* ssp. *petraea*	Wachau	39	0.547	0.609

Species	Geographic region	Population number	Ho	He
tetraploid *A. lyr.* ssp. *petraea*	Wachau	41	0.477	0.494
tetraploid *A. lyr.* ssp. *petraea*	Wachau	42	0.571	0.581
tetraploid *A. lyr.* ssp. *petraea*	Wachau	45	0.375	0.368
tetraploid *A. lyr.* ssp. *petraea*	Wachau	46	0.559	0.581
tetraploid *A. lyr.* ssp. *petraea*	Wachau	50	0.500	0.538
tetraploid *A. lyr.* ssp. *petraea*	Wachau	51	0.600	0.510
tetraploid *A. lyr.* ssp. *petraea*	Wachau	53	0.667	0.695
tetraploid *A. lyr.* ssp. *petraea*	Wachau	55	0.694	0.658
tetraploid *A. lyr.* ssp. *petraea*	Wachau	56	0.718	0.709
tetraploid *A. lyr.* ssp. *petraea*	Wachau	57	0.643	0.621
tetraploid *A. lyr.* ssp. *petraea*	Wachau	58	0.525	0.556
tetraploid *A. lyr.* ssp. *petraea*	Wachau	59	0.479	0.535
tetraploid *A. lyr.* ssp. *petraea*	Wachau	60	0.600	0.664
tetraploid *A. lyr.* ssp. *petraea*	Wachau	61	0.667	0.631
tetraploid *A. lyr.* ssp. *petraea*	Wachau	62	0.604	0.690
tetraploid *A. lyr.* ssp. *petraea*	Wachau	64	0.579	0.568
tetraploid *A. lyr.* ssp. *petraea*	Wachau	65	0.534	0.571
tetraploid *A. lyr.* ssp. *petraea*	Wachau	68	0.538	0.575
tetraploid *A. lyr.* ssp. *petraea*	Wachau	70	0.571	0.541
tetraploid *A. lyr.* ssp. *petraea*	Wachau	71	0.500	0.540
tetraploid *A. lyr.* ssp. *petraea*	Wachau	102	0.500	0.509
tetraploid *A. lyr.* ssp. *petraea*	Wachau	103	0.773	0.732
tetraploid *A. lyr.* ssp. *petraea*	Wachau	107	0.400	0.383
tetraploid *A. lyr.* ssp. *petraea*	Wachau	108	0.800	0.773
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	85	0.682	0.744
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	116	0.654	0.657
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	402196	0.750	0.693
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	915144	0.328	0.362
Mean			0.598	0.607

TABLE 9. Multilocus statistics of four microsatellite loci (SLL2, NGA162, ICE14, ICE7) of exclusively tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria. Calculations are based on tetrasomic inheritance. Observed heterozygosity Ho and expected heterozygosity He.

3.3.8. Bayesian Analyses and population genetic statistics detect rare geneflow between Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Hybrid individuals were found in several populations of mainly *Arabidopsis lyrata* ssp. *petraea*, based on Structure analysis of the tetraploid dataset with K = 4 (Fig. 17, 18). This K value was strongly supported by both statistics (Fig. 6) and biological interpretation. Proportions of each genetic cluster of individuals from the same population were summed up to the total proportions of all genetic clusters within this population, and resulting pie charts were plotted on a map. The hybrid index of each tetraploid *Arabidopsis lyrata* ssp. *petraea* population corresponds to the fraction (in percentage) of the genetic cluster characteristic for tetraploid *Arabidopsis arenosa* in this tetraploid *Arabidopsis lyrata* ssp. *petraea* population. A strong hybrid index of HI = 0.496 was only reported from one population, Pop402196 (*Arabidopsis lyrata* ssp. *petraea*), from the Eastern Alps (TABLE 10, Fig. 18). Weaker hybridisation was observed in *Arabidopsis lyrata* ssp. *petraea* populations from the Eastern Alps (Pop116) and the Wachau, especially in populations northeast of the Danube River (Pop1, Pop13, Pop14, Pop15, Pop16, Pop17), with a hybrid index of about HI = 0.052-0.118. The southwestern populations Pop35 and Pop75 showed similar levels of introgressive hybridisation. All other Wachau populations had a very low hybrid index of HI = 0.006-0.048. It became obvious, that hybrid populations of *Arabidopsis lyrata* ssp. *petraea* were predominantly located in the geographic contact zones of both species (Fig. 17, 18). In total two contact zones were localised, one in the northern Wachau and another in the Eastern Alps. In both regions tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations were only one kilometre apart from each other. Therefore, these findings could be interpreted as recurrent geneflow between the two species, mainly unidirectionally from *Arabidopsis arenosa* into *Arabidopsis lyrata* ssp. *petraea*.

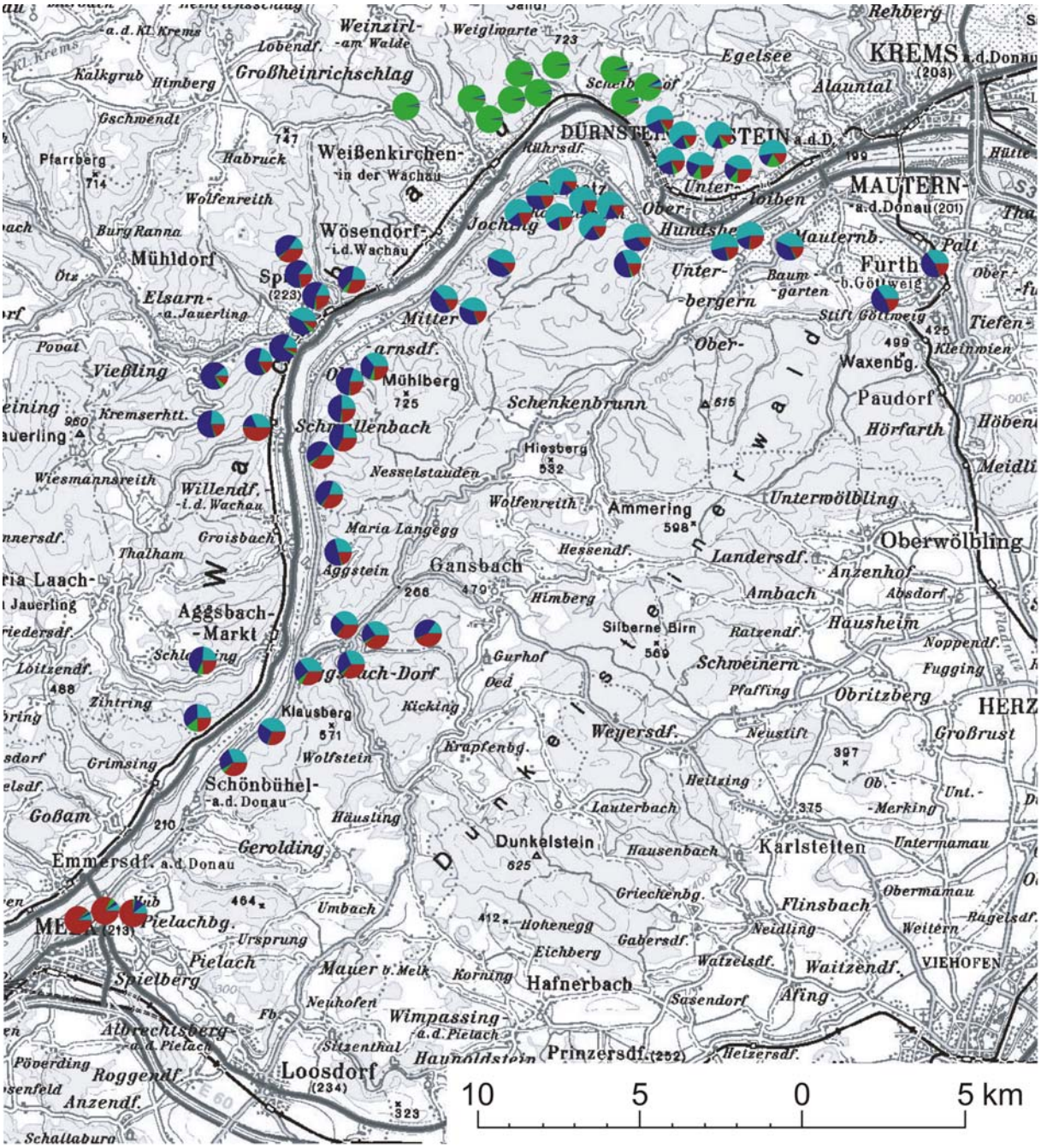


Fig. 17. Population structure estimated by Structure analysis of tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations (20 individuals/population) in the Wachau, based on seven microsatellite loci. Calculations were performed with K = 4, which corresponds to the numbers of ancestral populations. *Arabidopsis arenosa* populations are characterised by the predominantly green genetic cluster, *Arabidopsis lyrata* ssp. *petraea* populations mainly comprise three genetic clusters (red, light and dark blue), but also fractions of the green cluster. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

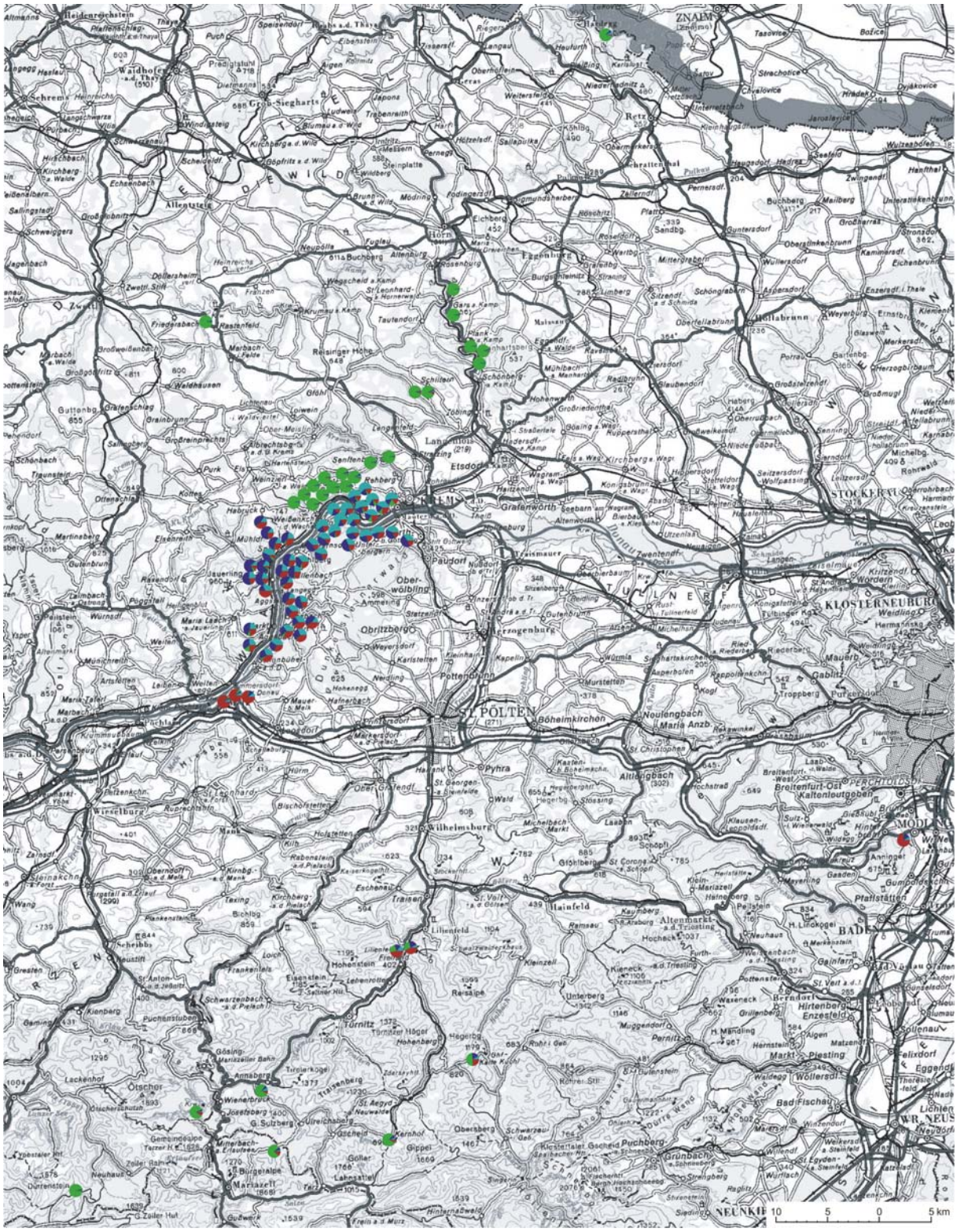


Fig. 18. Population structure estimated by Structure analysis of tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations (20 individuals/population) in Eastern Austria (Eastern Alps, Wachau, Krems/Kamp Valley), based on seven microsatellite loci. Calculations were performed with K = 4, which corresponds to the numbers of ancestral populations. *Arabidopsis arenosa* populations are characterised by the predominantly green genetic cluster, *Arabidopsis lyrata* ssp. *petraea* populations mainly comprise three genetic clusters (red, light and dark blue), but also fractions of the green cluster. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

Species	Geographic region	Population number	Hybrid index HI
tetraploid *A. lyr.* ssp. *petraea*	Wachau	1	0.087
tetraploid *A. lyr.* ssp. *petraea*	Wachau	9	0.017
tetraploid *A. lyr.* ssp. *petraea*	Wachau	10	0.019
tetraploid *A. lyr.* ssp. *petraea*	Wachau	11	0.013
tetraploid *A. lyr.* ssp. *petraea*	Wachau	12	0.028
tetraploid *A. lyr.* ssp. *petraea*	Wachau	13	0.118
tetraploid *A. lyr.* ssp. *petraea*	Wachau	14	0.062
tetraploid *A. lyr.* ssp. *petraea*	Wachau	15	0.100
tetraploid *A. lyr.* ssp. *petraea*	Wachau	16	0.173
tetraploid *A. lyr.* ssp. *petraea*	Wachau	17	0.052
tetraploid *A. lyr.* ssp. *petraea*	Wachau	18	0.048
tetraploid *A. lyr.* ssp. *petraea*	Wachau	24	0.058
tetraploid *A. lyr.* ssp. *petraea*	Wachau	25	0.062
tetraploid *A. lyr.* ssp. *petraea*	Wachau	26	0.032
tetraploid *A. lyr.* ssp. *petraea*	Wachau	27	0.023
tetraploid *A. lyr.* ssp. *petraea*	Wachau	28	0.015
tetraploid *A. lyr.* ssp. *petraea*	Wachau	29	0.039
tetraploid *A. lyr.* ssp. *petraea*	Wachau	30	0.027
tetraploid *A. lyr.* ssp. *petraea*	Wachau	31	0.016
tetraploid *A. lyr.* ssp. *petraea*	Wachau	32	0.058
tetraploid *A. lyr.* ssp. *petraea*	Wachau	34	0.019
tetraploid *A. lyr.* ssp. *petraea*	Wachau	35	0.076
tetraploid *A. lyr.* ssp. *petraea*	Wachau	36	0.007
tetraploid *A. lyr.* ssp. *petraea*	Wachau	37	0.011
tetraploid *A. lyr.* ssp. *petraea*	Wachau	39	0.011
tetraploid *A. lyr.* ssp. *petraea*	Wachau	41	0.034
tetraploid *A. lyr.* ssp. *petraea*	Wachau	42	0.010
tetraploid *A. lyr.* ssp. *petraea*	Wachau	45	0.008
tetraploid *A. lyr.* ssp. *petraea*	Wachau	46	0.010
tetraploid *A. lyr.* ssp. *petraea*	Wachau	48	0.011
tetraploid *A. lyr.* ssp. *petraea*	Wachau	50	0.013
tetraploid *A. lyr.* ssp. *petraea*	Wachau	51	0.016
tetraploid *A. lyr.* ssp. *petraea*	Wachau	53	0.030
tetraploid *A. lyr.* ssp. *petraea*	Wachau	55	0.024
tetraploid *A. lyr.* ssp. *petraea*	Wachau	56	0.061
tetraploid *A. lyr.* ssp. *petraea*	Wachau	57	0.047
tetraploid *A. lyr.* ssp. *petraea*	Wachau	58	0.014
tetraploid *A. lyr.* ssp. *petraea*	Wachau	59	0.057
tetraploid *A. lyr.* ssp. *petraea*	Wachau	60	0.053
tetraploid *A. lyr.* ssp. *petraea*	Wachau	61	0.018
tetraploid *A. lyr.* ssp. *petraea*	Wachau	62	0.036

Species	Geographic region	Population number	Hybrid index HI
tetraploid *A. lyr.* ssp. *petraea*	Wachau	64	0.008
tetraploid *A. lyr.* ssp. *petraea*	Wachau	65	0.026
tetraploid *A. lyr.* ssp. *petraea*	Wachau	68	0.014
tetraploid *A. lyr.* ssp. *petraea*	Wachau	69	0.025
tetraploid *A. lyr.* ssp. *petraea*	Wachau	70	0.012
tetraploid *A. lyr.* ssp. *petraea*	Wachau	71	0.027
tetraploid *A. lyr.* ssp. *petraea*	Wachau	72	0.015
tetraploid *A. lyr.* ssp. *petraea*	Wachau	75	0.148
tetraploid *A. lyr.* ssp. *petraea*	Wachau	102	0.014
tetraploid *A. lyr.* ssp. *petraea*	Wachau	103	0.019
tetraploid *A. lyr.* ssp. *petraea*	Wachau	107	0.006
tetraploid *A. lyr.* ssp. *petraea*	Wachau	108	0.010
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alp	85	0.231
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	116	0.290
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	402196	0.496
tetraploid *A. lyr.* ssp. *petraea*	Eastern Alps	915144	0.016

TABLE 10. Hybrid index HI of tetraploid *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria, based on seven microsatellite loci. HI corresponds to the fraction (in percentage) of the green genetic cluster characteristic for *Arabidopsis arenosa* in *Arabidopsis lyrata* ssp. *petraea*, according to a Structure analysis with K = 4. Underlined HI values (HI ≥ 0.05) mark populations with significant introgression from *Arabidopsis arenosa*.

Calculation of pairwise d_0, based on tetrasomic inheritance with the software Tetraploide, plotted against geographic distance along a linear transect (Fig. 20), supported geneflow between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* (Fig. 19), however, only for few *Arabidopsis lyrata* ssp. *petraea* populations from the Eastern Alps (Pop85, Pop116, Pop402196). These populations showed low genetic distance of an average $d_0 = 0.1$ to *Arabidopsis arenosa* populations from that region (Pop402140, Pop402168, Pop915142). In contrast, *Arabidopsis lyrata* ssp. *petraea* populations from the northern Wachau displayed strong genetic differentiation of an average $d_0 = 0.4$ to *Arabidopsis arenosa* populations from the same area (Fig. 19), which contradicts the genetic introgression observed with Bayesian Analyses. However, this is probably due to the reduced dataset, excluding individuals with missing data for d_0 calculation. So far only recent geneflow between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* was discussed. In this study there was little evidence for past geneflow. The very low hybrid index of HI = 0.006-0.048 in many *Arabidopsis lyrata* ssp. *petraea* populations from the Wachau could possibly be interpreted as past geneflow between the two species.

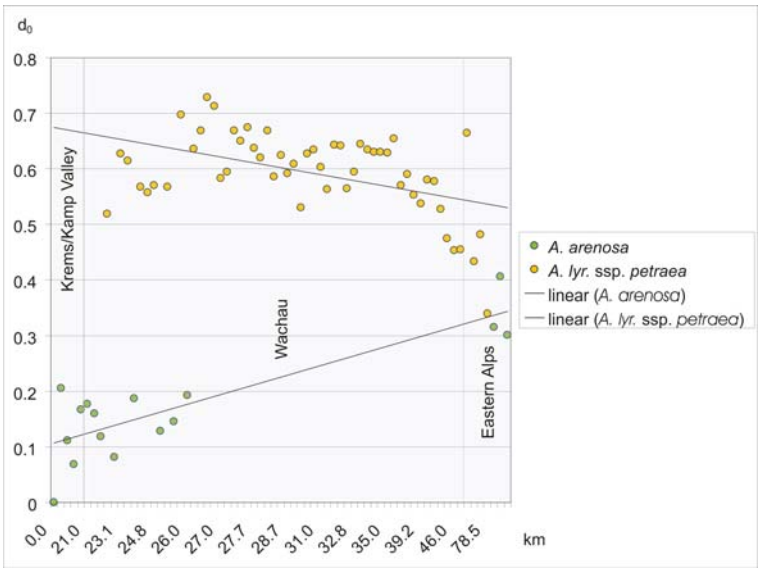


Fig. 19. Genetic distance d_0 according to Gregorius (1974) and Prevosti et al. (1975) of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria (Eastern Alps, Wachau, Krems/Kamp Valley), plotted against geographic distance. Calculation of d_0 is based on four microsatellite loci (SLL2, NGA162, ICE14, ICE7) and tetrasomic inheritance.

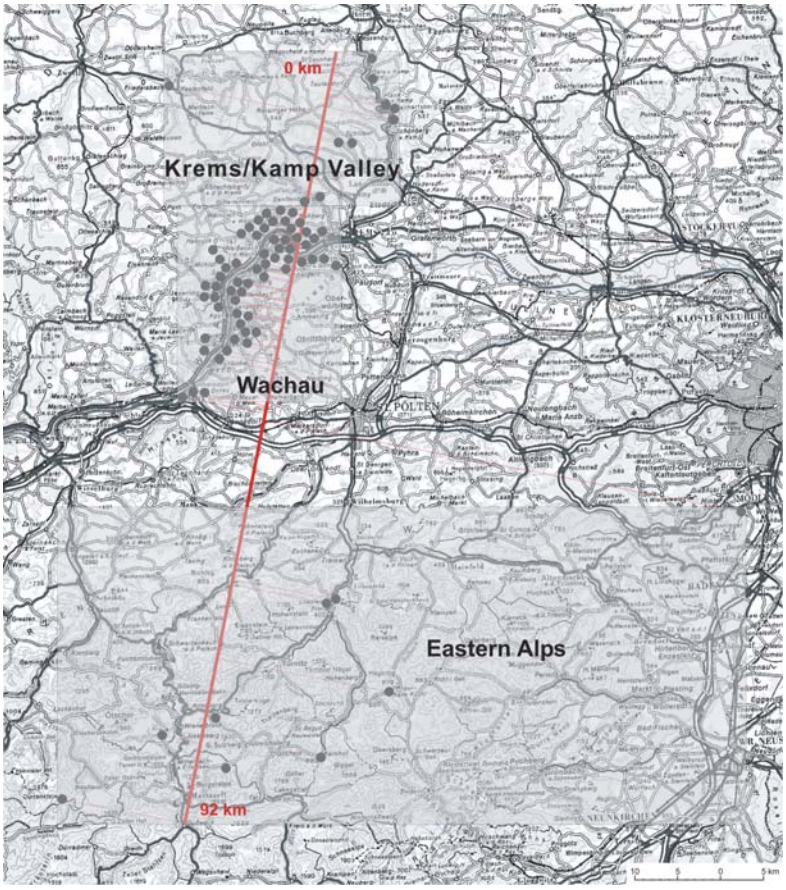


Fig. 20. Geographic distance of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria (Eastern Alps, Wachau, Krems/Kamp Valley). Populations are plotted on a linear transect. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

3.3.9. Bayesian Analysis visualises migrational movements of Arabidopsis lyrata ssp. petraea

Structure analysis of exclusively tetraploids with K = 4 revealed several genetic clusters (Fig. 17, 18). According to the geographic distribution of these clusters, a migration model of populations could be developed: As there was a high percentage of the red cluster both in the Eastern Alps and the southern Wachau, colonisation of the Wachau from the Eastern Alps can be assumed. The red cluster gradually decreased towards the northern Wachau, being replaced by the light and dark blue clusters. The dark blue cluster was mainly present in the south- and northwestern part of the Wachau, and the light blue cluster in the northeastern part on both riversides. In contrast, *Arabidopsis arenosa* was characterised by a strongly homogenuous genepool over a vast geographic range, including the Eastern Alps, the Wachau, Krems and Kamp Valley, and even the northern Waldviertel. In both types of Bayesian Analysis, Structure (Fig. 17, 18) and BAPS (Fig. 15), tetraploid *Arabidopsis arenosa* populations from all those regions formed one group. However, in an analysis of genetic distance d_0 assigned with geographic distance (Fig. 19) *Arabidopsis arenosa* populations from the Eastern Alps were genetically slightly differentiated with an average $d_0 = 0.2$, and weak geographic isolation could be assumed.

Morphology

3.3.10. Principal Component Analysis of the complete dataset reveals morphological intermediates in tetraploid Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Data were analysed with PCA and displayed in a two-dimensional graph (Fig. 21). Eigenvalue of axis one was 0.25, of axis two 0.11, comprising only 36% total variance of the data. As the following axes showed even lower eigenvalues (λ_3, λ_4 = 0.07; λ_5 = 0.06; λ_6, λ_7 = 0.05), they are not further discussed. Ordination revealed two groups, *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, but they were not strongly separated: Tetraploids of both species overlapped, in contrast to the diploids, which remained separate. Tetraploids, in particular *Arabidopsis lyrata* ssp. *petraea*, showed increased variance over their diploid ancestors. Group formation was only partly correlated with the variables, probably as a result of the high overall variance within tetraploids (Fig. 22). The group of tetraploid *Arabidopsis arenosa* was formed by the following variables, which highly correlated with axis one: length of main stem/length of leafy part of main stem; hairs on leafless part of main stem; hairs on flower buds simple, bifurcate, trifurcate; hairs on second stem leaf bifurcate, trifurcate; petal colour dark pink. Diploid *Arabidopsis arenosa* comprised the following variables: number of flowers on main stem; length of main stem from ground rosette to smallest stem leaf/length of main stem from ground rosette to biggest stem leaf; hairs on leafy part of main stem; hairs on main stem from ground rosette to first stem leaf simple, bifurcate, trifurcate; hairs on second stem leaf simple; petal colour light pink. However, also several tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* individuals were characterised by those variables. The group consisting of diploid and tetraploid *Arabidopsis lyrata* ssp. *petraea* was formed based on the following variables: number of leaves on upper part of main stem; number of leaves on main stem; length of second stem leaf/width of second stem leaf; length of uppermost stem leaf/width of uppermost stem leaf; length of biggest rosette leaf/length of biggest rosette leaf from first leaf tooth; basic number of leaf teeth of longest rosette leaf; hairs on main stem to first stem leaf; petal colour white. Due to PCA, morphological characters could roughly be grouped into hair and leaf characters discriminating between *Arabidopsis arenosa*, characterised by mainly hair characters, and *Arabidopsis lyrata* ssp. *petraea*, characterised by mainly leaf characters. Petal colour distinguished between *Arabidopsis lyrata* ssp. *petraea* (white), diploid *Arabidopsis arenosa* (light pink), and tetraploid *Arabidopsis arenosa* (dark pink).

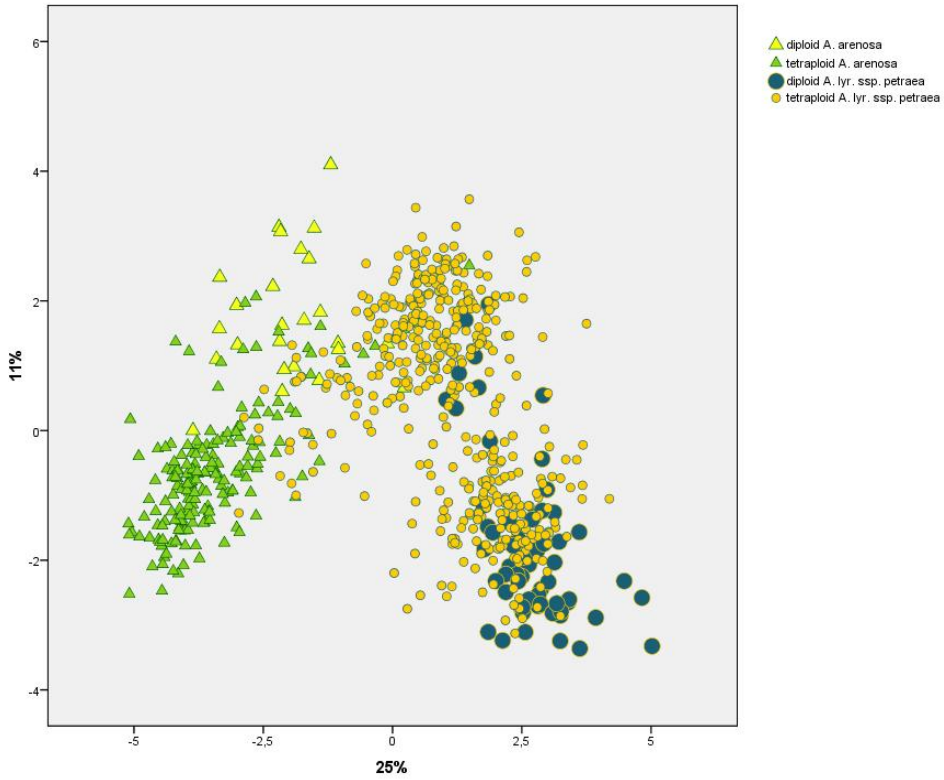


Fig. 21. Principal Component Analysis based on 29 morphological characters. Each symbol (circle, triangle) represents a population of either diploid or tetraploid *Arabidopsis arenosa* or *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*).

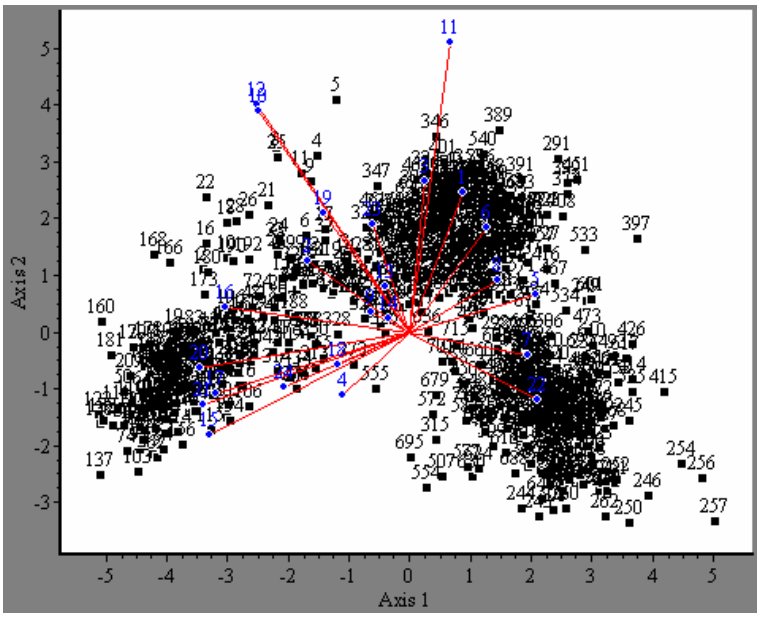


Fig. 22. Biplot of objects and variables of a Principal Component Analysis based on 29 morphological characters. Objects (black squares) are the individuals of both diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria and the Western Carpathians (diploid *Arabidopsis arenosa*). Variables (blue circles, red lines) are the morphological characters (see TABLE 3): 1 – PLHTP, 2 – PK, 3 – PL, 4 – VR/DBPS, 5 – D2BL/S2BL, 6 – DHL/SHL, 7 – MDLR/DC, 8 – MDLR/PZ, 9 – VNKL/VNDL, 10 – TRB, 11 – TRB, 12 – TRB, 13 – TRB, 14 – TRB, 15 – TRS, 16 – OLJT, 17 – OLVT, 18 – OLTT, 19 – 2BLJT, 20 – 2BLVT, 21 – 2BLTT, 22 – FLB, 23 – FKR, 24 – FKRT.

3.3.11. Principal Component Analysis of tetraploid Arabidopsis lyrata ssp. petraea detects a gradient of morphological hybrids

PCA was used to calculate the hybrid degree of tetraploid *Arabidopsis lyrata* ssp. *petraea* populations. If we want to test, if tetraploid *Arabidopsis lyrata* ssp. *petraea* is an allopolyploid between diploid *Arabidopsis lyrata* ssp. *petraea* and diploid *Arabidopsis arenosa*, the diploid taxa have to be used as reference groups. Diploid *Arabidopsis lyrata* ssp. *petraea* populations from Eastern Austria were taken as first "pure" ancestral group. Diploid *Arabidopsis arenosa* was not found in Eastern Austria. The geographically closest diploid *Arabidopsis arenosa* populations were located in the Carpathians. However, these populations were, at least partly, genetically distinct from tetraploid populations from the Alps, according to chloroplast sequence data. Therefore, diploid *Arabidopsis arenosa* from the Carpathians could not be included as reference group. Instead, tetraploid *Arabidopsis arenosa* from Eastern Austria, which was assumed to have originated via autopolyploidisation, was taken as second ancestral group. The hybrid index of each population was calculated from the two-dimensional PCA bar plot (Fig. 23) as deviation of the mean value of the investigated tetraploid *Arabidopsis lyrata* ssp. *petraea* population from the mean value of diploid *Arabidopsis lyrata* ssp. *petraea*. In the pie charts plotted on the map morphological proportion of diploid *Arabidopsis lyrata* ssp. *petraea* was coloured in yellow, of tetraploid *Arabidopsis arenosa* in green (Fig. 24). All tetraploid *Arabidopsis lyrata* ssp. *lyrata* populations showed morphological proportions of *Arabidopsis arenosa*. Especially the populations from the northern Wachau, north of the Danube River, which were in close geographic contact to *Arabidopsis arenosa*, had a high hybrid index of HI = 0.58 (TABLE 11). The average hybrid index of all other Wachau populations was moderate (HI = 0.26). The hybrid index of the Mödling population (Pop915144) from the Eastern Alps was very low (HI = 0.05), indicating limited to no geneflow with *Arabidopsis arenosa*.

Tetraploid *Arabidopsis arenosa* could not be analysed in that way, as "pure" ancestral diploid *Arabidopsis arenosa* was missing in Eastern Austria, a limitation, which was already discussed regarding the analysis of tetraploid *Arabidopsis lyrata* ssp. *petraea*.

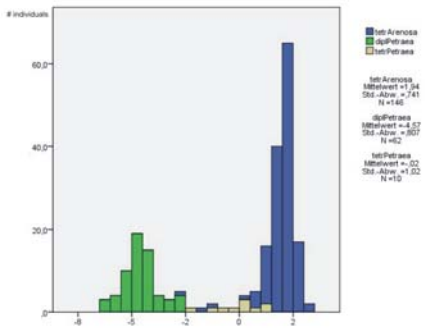


Fig. 23. Two-dimensional PCA bar plot of a tetraploid *Arabidopsis lyrata* ssp. *petraea* population (beige). Diploid *Arabidopsis lyrata* ssp. *petraea* (green) and tetraploid *Arabidopsis arenosa* (blue) populations from Eastern Austria are used as "pure" ancestral populations. The hybrid index HI is calculated as deviation of the mean value ("Mittelwert") of the investigated tetraploid *Arabidopsis lyrata* ssp. *petraea* population from the mean value ("Mittelwert") of diploid *Arabidopsis lyrata* ssp. *petraea*.

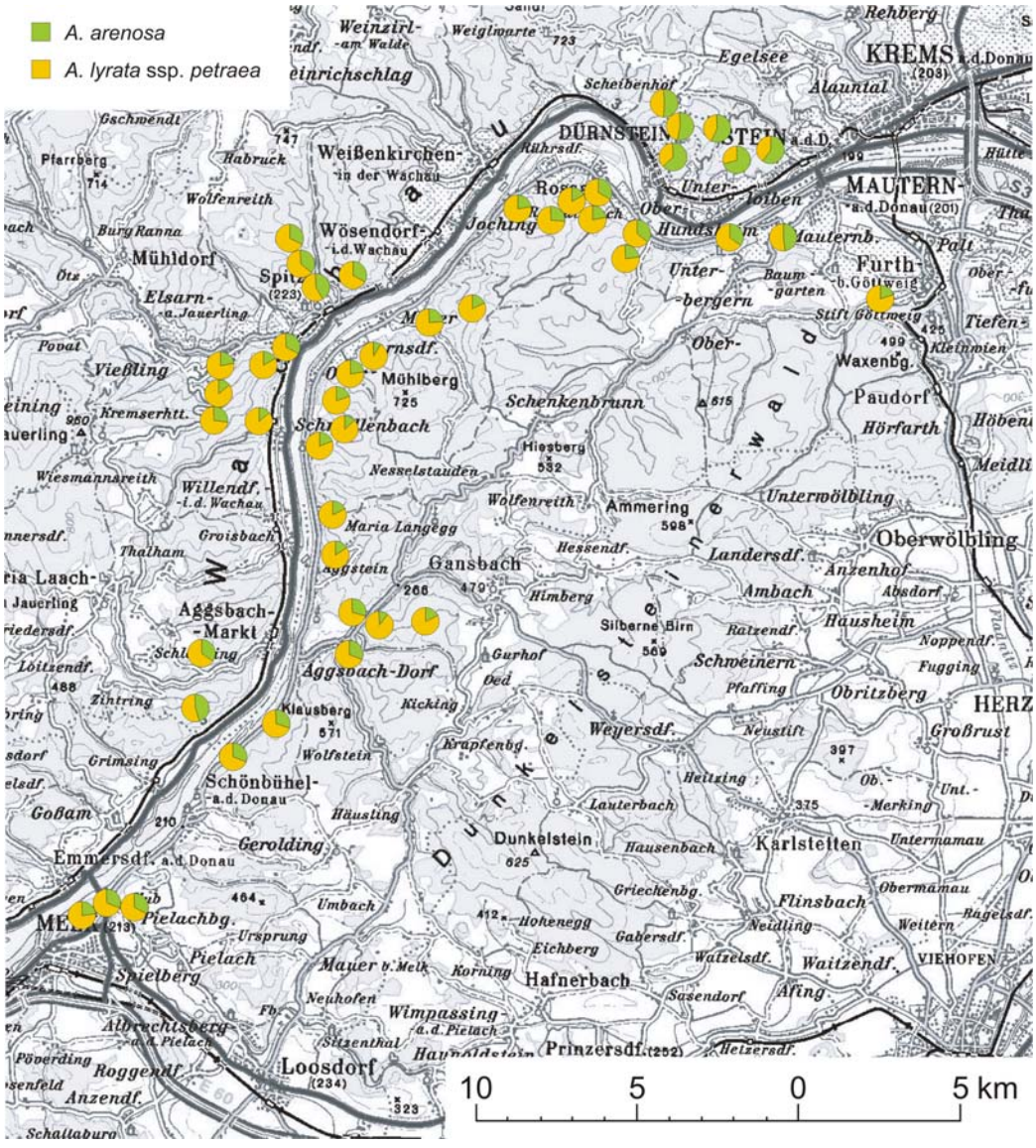


Fig. 24. Morphological proportion of *Arabidopsis arenosa* in tetraploid *Arabidopsis lyrata* ssp. *petraea* populations in the Wachau. Calculation see Fig. 23. Map taken from "Austrian Map Ost", BEV (Bundesamt für Eich- und Vermessungswesen).

Species	Geographic region	Population number	Hybrid index HI
tetraploid *A. lyr.* ssp. *petraea*	Wachau	1	0.56
tetraploid *A. lyr.* ssp. *petraea*	Wachau	9	0.48
tetraploid *A. lyr.* ssp. *petraea*	Wachau	10	0.35
tetraploid *A. lyr.* ssp. *petraea*	Wachau	12	0.24
tetraploid *A. lyr.* ssp. *petraea*	Wachau	14	0.62
tetraploid *A. lyr.* ssp. *petraea*	Wachau	15	0.70
tetraploid *A. lyr.* ssp. *petraea*	Wachau	16	0.60
tetraploid *A. lyr.* ssp. *petraea*	Wachau	17	0.52
tetraploid *A. lyr.* ssp. *petraea*	Wachau	18	0.50
tetraploid *A. lyr.* ssp. *petraea*	Wachau	24	0.33
tetraploid *A. lyr.* ssp. *petraea*	Wachau	26	0.41
tetraploid *A. lyr.* ssp. *petraea*	Wachau	27	0.33
tetraploid *A. lyr.* ssp. *petraea*	Wachau	28	0.36
tetraploid *A. lyr.* ssp. *petraea*	Wachau	29	0.33
tetraploid *A. lyr.* ssp. *petraea*	Wachau	30	0.17
tetraploid *A. lyr.* ssp. *petraea*	Wachau	31	0.14
tetraploid *A. lyr.* ssp. *petraea*	Wachau	32	0.23
tetraploid *A. lyr.* ssp. *petraea*	Wachau	33	0.15
tetraploid *A. lyr.* ssp. *petraea*	Wachau	34	0.28
tetraploid *A. lyr.* ssp. *petraea*	Wachau	35	0.35
tetraploid *A. lyr.* ssp. *petraea*	Wachau	36	0.36
tetraploid *A. lyr.* ssp. *petraea*	Wachau	37	0.23
tetraploid *A. lyr.* ssp. *petraea*	Wachau	39	0.33
tetraploid *A. lyr.* ssp. *petraea*	Wachau	41	0.27
tetraploid *A. lyr.* ssp. *petraea*	Wachau	43	0.17
tetraploid *A. lyr.* ssp. *petraea*	Wachau	45	0.20
tetraploid *A. lyr.* ssp. *petraea*	Wachau	46	0.18
tetraploid *A. lyr.* ssp. *petraea*	Wachau	51	0.19
tetraploid *A. lyr.* ssp. *petraea*	Wachau	52	0.25
tetraploid *A. lyr.* ssp. *petraea*	Wachau	53	0.34
tetraploid *A. lyr.* ssp. *petraea*	Wachau	55	0.24
tetraploid *A. lyr.* ssp. *petraea*	Wachau	56	0.32
tetraploid *A. lyr.* ssp. *petraea*	Wachau	57	0.31
tetraploid *A. lyr.* ssp. *petraea*	Wachau	58	0.31
tetraploid *A. lyr.* ssp. *petraea*	Wachau	60	0.09
tetraploid *A. lyr.* ssp. *petraea*	Wachau	61	0.23
tetraploid *A. lyr.* ssp. *petraea*	Wachau	62	0.20
tetraploid *A. lyr.* ssp. *petraea*	Wachau	64	0.14
tetraploid *A. lyr.* ssp. *petraea*	Wachau	65	0.20
tetraploid *A. lyr.* ssp. *petraea*	Wachau	68	0.23
tetraploid *A. lyr.* ssp. *petraea*	Wachau	69	0.31

Species	Geographic region	Population number	Hybrid index HI
tetraploid *A. lyr.* ssp. *petraea*	Wachau	70	0.30
tetraploid *A. lyr.* ssp. *petraea*	Wachau	71	0.11
tetraploid *A. lyr.* ssp. *petraea*	Wachau	72	0.16
tetraploid *A. lyr.* ssp. *petraea*	Wachau	73	0.18
tetraploid *A. lyr.* ssp. *petraea*	Wachau	75	0.45

TABLE 11. Hybrid index HI of tetraploid *Arabidopsis lyrata* ssp. *petraea* populations from the Wachau, based on 29 morphological characters. HI corresponds to the deviation of the mean value of the investigated tetraploid *Arabidopsis lyrata* ssp. *petraea* population from the mean value of diploid *Arabidopsis lyrata* ssp. *petraea* in a two-dimensional PCA bar plot (in percentage). Calculation see Fig. 23, graphical display see Fig. 24. Underlined HI values (HI ≥ 0.05) mark populations with significant introgression from *Arabidopsis arenosa*.

3.4. Discussion

CpDNA

3.4.1. Chloroplast sequence data indicate rare chloroplast capture between Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Chloroplast capture is frequently involved in intraspecies and interspecies hybridisation, and has been reported in e.g. *Helianthus*, *Gossypium*, and *Quercus* (Rieseberg and Soltis, 1991). If chloroplast sequences, diagnostic for another species, are found within the analysed species, interspecific introgression can be assumed. It can additionally be supported by both other independently evolving characters, e.g. nuclear microsatellite markers and morphology, and the geographic distribution of the taxa (Wendel and Albert, 1992; Rieseberg, 1995; Rieseberg et al., 1996). Arising incompatibilities between nuclear and cytoplasmic genes of the two hybridising species are often weakened by partial or complete cytoplasmic male sterility (Lewis, 1941; Frank, 1989). F1 hybrids with cytoplasmic male sterility could, in contrary, show increased female fitness, leading to increased seed production (Lewis, 1941). In this way, hybrids with chloroplast capture might successfully be dispersed within a population. However, interspecific sharing of chloroplast sequence types can also be the result of ancestral polymorphism and indicate incomplete lineage sorting.

In our study suprahaplotypes A and Q, diagnostic for *Arabidopsis arenosa*, were found in tetraploid *Arabidopsis lyrata* ssp. *petraea* populations in the northern Wachau, indicating hybrid origin in form of chloroplast capture. The populations were restricted to an area of close geographic contact between populations of both species, obviously resulting in rare, unidirectional *Arabidopsis arenosa* seed flow into the distribution range of *Arabidopsis lyrata* ssp. *petraea*. Repeated backcrossing with *Arabidopsis lyrata* ssp. *petraea* resulted in individuals with the nuclear genome of *Arabidopsis lyrata* ssp. *petraea* and the chloroplast genome of *Arabidopsis arenosa*. In contrast, suprahaplotype B, diagnostic for *Arabidopsis arenosa*, was detected both in diploid and tetraploid populations of *Arabidopsis lyrata* ssp. *petraea*, and ancestral polymorphism can be assumed, predated to the main radiation of the genus *Arabidopsis* approximately two million years ago (Koch and Matschinger, 2007).

Microsatellites

3.4.2. Autopolyploid origin of Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

An increasing number of studies implies, that autopolyploidisation, the self-doubling of a genome, is more frequent than previously assumed (Soltis et al., 2004, 2007). Autopolyploidy is normally indicated by tetrasomic inheritance (Soltis and Soltis, 1993), characterised by equal frequencies of all possible allelic combinations (Muller, 1914). Meanwhile, there is increasing evidence, that tetrasomic inheritance actually occurres in natural populations of autopolyploids, e.g. *Heuchera micrantha* (Soltis and Soltis, 1989), *Dactylis glomerata* (Tomekpe and Lumaret, 1991), and *Tolmiea menziesii* (Soltis and Rieseberg, 1986). For natural populations of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* from Eastern Austria we also suggest autopolyploidisation and tetrasomic inheritance, although the Bayesian Test failed, as our microsatellite marker was not variable enough. But calculation with other SSRs was not possible in *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. No species-specific alleles were found, which hindered the correct assignment of alleles according to the segregation patterns of either tetra- or disomic inheritance. However, as allele frequencies between diploids and tetraploids of the same species were extremely similar, we assume autopolyploid formation in both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. New alleles in tetraploids occurred at very low frequencies and were mostly not shared between the two species.

3.4.3. Past and recent geneflow from Arabidopsis arenosa into Arabidopsis lyrata ssp. petraea and the formation of two hybrid zones

Genetic introgression takes place, if a hybrid individual backcrosses with an individual of one of the parental populations over generations. Consequently, the backcross has part of the genetic material of the introgressed taxon integrated in its genome. In the case of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*, introgressive hybridisation is strongly indicated by our study with seven nuclear microsatellite markers. In Bayesian Analyses hybrid genotypes were rarely intermediate between the parental genotypes. In the majority of hybrid individuals only part of the genome could be assigned to the other species. Therefore, recent allopolyploidisation can be ruled out. For recent allopolyploids equal proportions of both parental genotypes is characteristic, as e.g. in the hybrid fern *Osmunda x intermedia*, based on analyses of ten nuclear gene loci (Yatabe, 2009). Geneflow between the two *Arabidopsis*

species mostly occurred unidirectionally from *Arabidopsis arenosa* into *Arabidopsis lyrata* ssp. *petraea*. We assume, that the hybrid individual between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* repeatedly backcrossed with individuals of parental *Arabidopsis lyrata* ssp. *petraea*, leading to an *Arabidopsis lyrata* ssp. *petraea* genome with fractions of the *Arabidopsis arenosa* genome. Interestingly, unidirectional geneflow was also responsible for the formation of the natural hybrid *Arabidopsis suecica*, which is of allopolyploid origin, with the maternal parent *Arabidopsis thaliana* and the paternal parent *Arabidopsis arenosa*, confirmed by artificial crosses (Comai et al., 2000). In both cases of hybridisation *Arabidopsis arenosa* was the pollen donor, which could be explained by strong cytoplasmatic and nuclear incompatibility reactions in case of *Arabidopsis arenosa* as seed donor. The fraction of the *Arabidopsis arenosa* genome was often rather small in introgressed *Arabidopsis lyrata* ssp. *petraea*. This could either be the result of backcrossing of the hybrid with *Arabidopsis lyrata* ssp. *petraea* over many generations or limited permeability of introgressive *Arabidopsis arenosa* alleles in *Arabidopsis lyrata* ssp. *petraea*. It was recently found out that the genome is not uniformly porous for introgressive alleles (Baack and Rieseberg, 2007). Parental alleles seem to vary in their fitness within different genetic backgrounds. The different introgression patterns can both be explained by selection and recombination. If neutral or advantageous alleles are linked to sites that contribute to reproductive isolation, less introgression is expected, as selection acts against it. Less introgression is also true for genomic regions with low levels of recombination, e.g. close to chromosomal breakpoints, as seen in *Helianthus* (Yatabe et al., 2007) and *Drosophila* (Machado et al., 2007). However, as our SSR markers were neither linked to sites accounting for reproductive isolation nor located in genomic regions with low levels of recombination, uniform permeability of foreign alleles can be expected. Thus, individuals with a low degree of introgression have probably backcrossed over many generations, and individuals with a high degree of introgression can be assumed to be of recent hybrid origin, lacking repeated backcrossing with *Arabidopsis lyrata* ssp. *petraea*.

Those individuals with presumably recent geneflow from *Arabidopsis arenosa* were found in so called hybrid zones, areas of sympatry of closely related taxa, which are incompletely reproductively isolated from each other and still can cross-fertilise, resulting in hybrid formation (Barton and Hewitt, 1985, 1989; Harrison, 1993). During the last decades increasing interest is paid to studies of hybrid zones. Such zones have been found in a vast range of both plant and animal groups, e.g. orchids (Aagaard et al., 2005), sunflowers (Rieseberg et al., 1999), mussels (Comesaña et al, 1999; Toro et al., 2004), honeybees

(Beekman et al., 2007), birds (Lockley, 1992; Haas and Brodin, 2005; Ruegg, 2008), sea water fish (Van Herwerden and Doherty, 2006), and freshwater fish (Costedoat et al., 2007). As Barton and Hewitt stated (1985), "The widespread occurrence of natural hybrid zones, the taxonomic questions they pose, and the opportunities they present for investigating processes of species isolation make them a research field in their own right". The characterisation of the dynamics of a hybrid zone in space and time can radically change evolutionary scenarios (Buggs, 2007), thereby also providing important knowledge of the dynamics of hybrid species´ invasion for natural conservation (Whitham et al., 1991; Ellstrand, 1992; Rhymer and Simberloff, 1996; Whitham et al., 1999). The structure of a hybrid zone varies largely, according to the degree of genetic and ecological differentiation between the two meeting taxa, their rates of dispersal, and the fitness of their hybrid offspring (Harrison, 1993). If the fitness of the hybrid is increased over its parents´ fitness, there is either the possibility of a stable hybrid zone or hybrid zone movement. In case of a stable hybrid zone, the hybrid´s fitness is restricted to a hybrid zone habitat, which is ecologically intermediate to the habitats occupied by its parental taxa, an "ecotonal" zone. If not thus dependent on an "ecotonal" zone, hybrid zone movement into the parental ranges might occur, thereby gradually eliminating parental populations. If only one parental taxa shows lower fitness than the hybrid, hybrid zone expansion takes place unidirectionally. Hybrid zone movement in favour of one parental taxon is likely to occur in cases of introgressive hybridisation. Repeated backcrossing with one of the parental taxa could result in increased fitness of the hybrid over its parent, and parental populations could finally be completely replaced, even driven to extinction. In contrast, if the fitness of the parents is increased over the hybrid´s fitness, parental range expansion into the hybrid zone occurs, leading to narrow "tension zones" (Barton, 1979; Barton and Hewitt, 1985). In our study two hybrid zones of approximately 100 km^2 each were detected in areas of close geographic neighbourhood of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. One hybrid zone was located in the northern Wachau, the second in the foothills of the Eastern Alps. Hybrid zone movement in favour of *Arabidopsis lyrata* ssp. *petraea* was observed, probably due to increased fitness of the hybrid over its parent.

So far we only discussed recent geneflow between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* in hybrid zones. However, Bayesian Analyses also revealed genetic clusters of *Arabidopsis arenosa* in *Arabidopsis lyrata* ssp. *petraea* populations, which were not in close geographic contact to *Arabidopsis arenosa* populations, e.g. throughout the southern part of the Wachau. Recent introgression of *Arabidopsis arenosa* can be ruled out, as long-distance insect pollination via small bees and hoverflies (Fig. 1E, F), the main

pollinators of the two species, is extremely improbable. This finding could be interpreted as past geneflow between the two species. Similar to the two recent hybrid zones, *Arabidopsis arenosa x Arabidopsis lyrata* ssp. *petraea* hybrids may have formed, either on the diploid or tetraploid level, which backcrossed with parental *Arabidopsis lyrata* ssp. *petraea* individuals over many generations. Hence, there is a weak possibility of hybrid instead of autopolyploid origin of Eastern Austrian tetraploid *Arabidopsis lyrata* ssp. *petraea*. Past geneflow could be assumed to have occurred during the climate oscillations of the Pleistocene. Both species survived the glacials in this unglaciated part of Eastern Austria, according to their high overall genetic diversity in that region. But do we have more evidence for past geneflow from *Arabidopsis arenosa* into *Arabidopsis lyrata* ssp. *petraea* except for weak admixture in Bayesian Analyses? One possible consequence of introgressive hybridisation can be the transfer of adaptations from one taxon to the other. In introgressed *Arabidopsis lyrata* ssp. *petraea* the transfer of edaphic adaptations from *Arabidopsis arenosa* could be assumed. Diploid *Arabidopsis lyrata* ssp. *petraea* is restricted to calcarous bedrock in Central Europe. However, tetraploid *Arabidopsis lyrata* ssp. *petraea* mainly colonised siliceous bedrock of the Bohemian Massif in Eastern Austria. As *Arabidopsis arenosa* is also found on siliceous bedrock, this new adaptation of tetraploid *Arabidopsis lyrata* ssp. *petraea* could be due to geneflow from *Arabidopsis arenosa*. Introgressed plants and animals frequently show adaptations transferred from their hybrid partner (Arnold, 2004). Choler et al. (2002) had evidence for novel adaptations in introgressed populations of the high alpine *Carex curvula*, which enabled them to occupy ecologically marginal habitats. An introgressed taxa of the silversword alliance in Hawai´i, *Dubautia ciliolata*, was able to colonise new lava flows formerly occupied by *Dubautia scabra*, which introgressed into *Dubautia ciliolata* in areas of sympatry (Caraway et al., 2001). A traditional study object for introgressive hybridisation in plant taxa is the *Iris* species complex with *Iris fulva*, *Iris hexagona* and *Iris brevicaulis* (Riley, 1938; Cruzan and Arnold, 1993, 1994). Introgressed *Iris hexagona* individuals showed higher shade tolerance than non-introgressed individuals, indicating that introgressing, shade tolerant *Iris fulva* had transferred this particular adaptation to *Iris hexagona* (Bennett and Grace, 1990). The *Helianthus* species complex is another classical example for studies on introgressive hybridisation. Adaptive introgression can be assumed from *Helianthus debilis* var. *cucumerifolius* into *Helianthus annuus* (Heiser, 1951). However, the hybrid index of most of the tetraploid *Arabidopsis lyrata* ssp. *petraea* populations beyond the two hybrid zones is extremely weak, and it remains doubtful, if past geneflow between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* can still be measured after multitudinous backcrossing.

3.4.4. Evolutionary history of Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea in Eastern Austria

We assume autopolyploid origin of both *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* and recent introgressive hybridisation between them. Recently introgressed individuals are concentrated in hybrid zones. But are there any indications how the Wachau was colonised by the tetraploids of both species? Regarding *Arabidopsis lyrata* ssp. *petraea*, autopolyploidisation in the foothills of the Eastern Alps can be assumed, as both diploid and tetraploid populations were found in that region. The diploid populations were restricted to the eastern part of that area and showed strong genetic differentiation between each other. According to Bayesian Analyses the four investigated populations formed four distinct genetic groups, indicating long-term isolation. This is a similar pattern observed in populations of *Cochlearia macrorrhiza* from that area, where also strong genetic differentiation between populations was found, based on AFLP fingerprinting (Koch et al., 2003). Both species grow in relict habitats of exposed rocks or rocky slopes (*Arabidopsis lyrata* ssp. *petraea*) or along cold headwaters (*Cochlearia macrorrhiza*). In the case of *Arabidopsis lyrata* ssp. *petraea* the establishment of large, forested areas resulted in its restriction to these warm-stage refugia, where competition with other plants is extremely low. Due to Bayesian Analysis, tetraploid *Arabidopsis lyrata* ssp. *petraea* spread northwestwards from the Eastern Alps along several mountain rivers and finally migrated into the southernmost part of the Wachau, probably along the Pielach River. Colonisation of the Wachau continued from south to north. Due to the perennial life cycle of *Arabidopsis lyrata* ssp. *petraea*, migration was probably slow. During the LGM large population sizes can be assumed, as the unglaciated Eastern Alps and the Wachau climatically corresponded to arctic tundra. This vegetation type is nowadays found e.g. on Iceland, where large *Arabidopsis lyrata* ssp. *petraea* populations are established. During Holocene warming *Arabidopsis lyrata* ssp. *petraea* was restricted to exposed rocks, rocky slopes, and gravel. Forestation and anthropogenic landuse, in the Wachau especially viniculture, even reduced the number of suitable habitats and contributed to a population mosaic. Consequently, geneflow between the populations was limited.

For tetraploid *Arabidopsis arenosa* migrational movements were not that obvious. According to Bayesian Analysis, tetraploids from the Eastern Alps, the Wachau, and areas north of it were extremely homogenuous and displayed no genetic gradient. This can probably be explained by rapid colonisation of those regions from a common area of origin. Rapid

migration was certainly facilitated by the mostly biennial life cycle of *Arabidopsis arenosa*. Subsequent range fragmentation between the Eastern Alps and the Wachau led to a disjunct distribution of *Arabidopsis arenosa* in Eastern Austria, but the relatively large population sizes of approximately 5000 individuals per population and landscape dynamics resulted in ongoing geneflow within and between the populations in each of the two regions. Consequently, rare alleles were preserved, and random genetic drift played a minor role. Tetraploid *Arabidopsis arenosa* is, in contrast to *Arabidopsis lyrata* ssp. *petraea*, a highly colonising species, as seen from rapid postglacial colonisation of northwestern and northern Europe (Meusel et al., 1965). Habitat preference of tetraploid *Arabidopsis arenosa* is broader than of *Arabidopsis lyrata* ssp. *petraea*: Next to relict habitats of exposed rocks and rocky slopes also gravel along streets and railroads can easily be colonised. Especially in the Eastern Alps and north of the Wachau, in the Krems and Kamp Valley, tetraploid *Arabidopsis arenosa* populations could even spread in recent times, fascilitating geneflow between the populations. In contrast, diploid *Arabidopsis arenosa* from the Western Carpathians is much more restricted to natural habitats of single rocks and gravel in the herbal layer of forests of class *Querco-Fagetea*, order *Fagetalia*, association *Carpinion betuli*, with the four different association types *Dentario-enneaphylli-Fagetum*, *Cephalanthero-Fagetum*, *Carici albae (Abieti)-Fagetum*, and *Seslerio-Fagetum* (Ujházyová and Ujházy, 2004). This cytotype is obviously not that drought tolerant as tetraploid *Arabidopsis arenosa*. Nevertheless, population sizes are large, as this forest type is extremely common in the Western Carpathians, contributing to geneflow between the populations. According to Bayesian Analyses, genetic homogenisation was observed between all investigated diploid *Arabidopsis arenosa* populations except for *Arabidopsis neglecta* and *Arabidopsis petrogena* nom. prov.. *Arabidopsis neglecta* grows in higher alpine habitats, and, therefore, geneflow with other taxa of the *Arabidopsis arenosa* species complex is prevented. Although *Arabidopsis petrogena* nom. prov. occupies similar habitats than diploid *Arabidopsis arenosa*, genetic isolation was even observed between the two investigated populations. Probably the annual life cycle of this species (personal observation) contributes to that finding.

Morphology

3.4.5. Indication of morphological hybrids between Arabidopsis arenosa and Arabidopsis lyrata ssp. petraea

Morphological intermediacy can be used in the early stages of genetic introgression to indicate the hybrid degree, but in general its application is limited. Rieseberg and Ellstrand (1993) showed in artificial hybrids, that 10% of the morphological characters measured in the F1 hybrids and 30% in the later generation hybrids were novel or extreme relative to the parental species. Therefore, a precise classification of hybrid backcrosses and parental species is usually not possible, based on exclusively morphometrics. In our study of *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* detection of hybrids, based on multivariate statistics of morphological characters, was even more critical, as we had to use tetraploid *Arabidopsis arenosa* as "pure" ancestral outgroup because of the lack of geographically close diploid populations. Morphological plasticity was significantly increased in the polyploids of both species, probably resulting from increased genotypic plasticity as a result of genome doubling. In the PCA of the complete dataset the group of tetraploid *Arabidopsis arenosa* partly overlapped with the group of tetraploid *Arabidopsis lyrata* ssp. *petraea*. Thus, we have to be aware, that part of *Arabidopsis arenosa´s* morphological proportion, contributing to *Arabidopsis lyrata* ssp. *petraea´s* morphological hybrid degree, is due to the large intraspecies´ variance of the tetraploids. Consequently, morphological measurements were mainly used to discriminate between diploid and tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. The hybrid index was significantly higher in populations from the northeastern Wachau north of the Danube River. Both chloroplast sequence data and nuclear SSR markers indicated a hybrid zone in this area, which is supported by morphological data.

3.5. Conclusions

What can we learn from this study? One major finding is ongoing geneflow between members of the genus *Arabidopsis*. Geneflow not only between diploids, which resulted in the origin of the allotetraploids *Arabidopsis suecica* and *Arabidopsis kamchatica*, but also between tetraploids, as in the two hybrid zones in Eastern Austria. In contrast to the allopolyploids, tetraploid *Arabidopsis arenosa x Arabidopsis lyrata* ssp. *petraea* so far failed to form a new species, probably due to the absence of factors, which normally contribute to homoploid hybrid speciation: On the one hand, reproductive isolation of the hybrids from parental individuals was not observed. Instead, hybrid individuals backcrossed with parental *Arabidopsis lyrata* ssp. *petraea*. On the other hand, ecological and/or spatial separation of hybrids from their parents, which can rapidly prevent backcrossing with parental individuals, was not observed. Instead, hybrids successfully grow on the same bedrock as their parents, which is Gföhler Gneis and amphibolite of the Bohemian Massif in the northern Wachau and limestone in the Eastern Alps. Regarding drought tolerance, hybrid individuals were found in more exposed habitats characteristic for *Arabidopsis lyrata* ssp. *petraea*, which additionally enhances the frequency of backcrossing with this subspecies. Hence, we can conclude, that geneflow between tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* results in introgressive hybridisation. For the establishment of a new hybrid species factors contributing to reproductive and ecological isolation would have to change drastically.

The other major finding concerns the origin of tetraploid *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea*. Based on our results (this PhD thesis, chapters 2 and 3), several independent autopolyploidisation events are very likely to have occurred in *Arabidopsis arenosa*. Also for tetraploid *Arabidopsis lyrata* ssp. *petraea* our data tend towards the interpretation of autopolyploid origin of this cytotype. However, we can not completely rule out, that hybridisation between *Arabidopsis arenosa* and *Arabidopsis lyrata* ssp. *petraea* and subsequent backcrossing with *Arabidopsis lyrata* ssp. *petraea* was the initial mode, although with our combined approach of cpDNA sequence data, ntDNA microsatellites, and morphology we were only able to detect more recent geneflow between these two species.

3.6. LITERATURE CITED

AAGAARD, S. M., S. M. SÅSTAD, J. GREILHUBER, AND A. MOEN. 2005. A secondary hybrid zone between diploid *Dactylorhiza incarnata* ssp. *cruenta* and allotetraploid *D. lapponica* (Orchidaceae). *Heredity* 94: 488-496.

ABBOTT, R. J., M. G. RITCHIE, AND P. M. HOLLINGSWORTH. 2008. Introduction. Speciation in plants and animals: pattern and process. *Philos Trans R Soc Lond B Biol Sci* 363: 2965-2969.

ADAMS, K. L., AND J. F. WENDEL. 2005. Novel patterns of gene expression in polyploid plants. *Trends Genet* 21: 539-543.

ADAMS, K. L., AND J. F. WENDEL. 2005. Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* 8: 135-141.

ANDERSON, E. 1948. Hybridization of the habitat. *Evolution* 2: 1-9.

ANSELL, S. W. 2004. Historical biogeography and population genetics of the plant *Arabidopsis petraea*. Wolfson College Cambridge.

ARNOLD, M. L. 2004. Transfer and origin of adaptations through natural hybridization: were Anderson and Stebbins right? *Plant Cell* 16: 562-570.

BAACK, E. J., AND L. H. RIESEBERG. 2007. A genomic view of introgression and hybrid speciation. *Curr Opin Genet Dev* 17: 513-518.

BARTON, N. H. 1979. The dynamics of hybrid zones. *Heredity* 43: 341-359.

BARTON, N. H., AND G. M. HEWITT. 1985. Analysis of hybrid zones. *Ann Rev Ecol Syst* 16: 113-148.

BARTON, N. H., AND G. M. HEWITT. 1989. Adaptation, speciation and hybrid zones. *Nature* 341: 497-503.

BEEKMAN, M., M. H. ALLSOPP, T. C. WOSSLER, AND B. P. OLDROYD. 2008. Factors affecting the dynamics of the honeybee (*Apis mellifera*) hybrid zone of South Africa. *Heredity* 100: 13-18.

BENNETT, B. D., AND J. B. GRACE. 1990. Shade tolerance and its effect on the segregation of two species of Louisiana *Iris* and their hybrids. *Am J Bot* 77: 100-107.

BRETAGNOLLE, F., AND J. D. THOMPSON. 1995. Gametes with the stomatic chromosome number: Mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytol* 129: 1-22.

BUGGS, R. J. 2007. Empirical study of hybrid zone movement. *Heredity* 99: 301-312.

CARAWAY, V., G. D. CARR, AND C. W. MORDEN. 2001. Assessment of hybridization and introgression in lava-colonizing Hawaiian *Dubautia* (Asteraceae: Madiinae) using RAPD markers. *Am J Bot* 88: 1688-1694.

CATALÁN, P., J. G. SEGARRA-MORAGUES, M. PALOP-ESTEBAN, C. MORENO, AND F. GONZÁLEZ-CANDELAS. 2006. A Bayesian approach for discriminating among alternative inheritance hypotheses in plant polyploids: the allotetraploid origin of genus *Borderea* (Dioscoreaceae). *Genetics* 172: 1939-1953.

CHOLER, P., B. ERSCHBAMER, A. TRIBSCH, L. GIELLY, AND P. TABERLET. 2004. Genetic introgression as a potential to widen a species' niche: insights from alpine *Carex curvula*. *Proc Natl Acad Sci U S A* 101: 171-176.

CLAUSS, M. J., H. COBBAN, AND T. MITCHELL-OLDS. 2002. Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaeae). *Mol Ecol* 11: 591-601.

CLAUSS, M. J., AND M. A. KOCH. 2006. Poorly known relatives of *Arabidopsis thaliana*. *Trends Plant Sci* 11: 449-459.

CLAUSS, M. J., AND T. MITCHELL-OLDS. 2003. Population genetics of tandem trypsin inhibitor genes in *Arabidopsis* species with contrasting ecology and life history. *Mol Ecol* 12: 1287-1299.

CLAUSS, M. J., AND T. MITCHELL-OLDS. 2006. Population genetic structure of *Arabidopsis lyrata* in Europe. *Mol Ecol* 15: 2753-2766.

CLEMENT, M., D. POSADA, AND K. A. CRANDALL. 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1659.

COMAI, L., A. P. TYAGI, K. WINTER, R. HOLMES-DAVIS, S. H. REYNOLDS, Y. STEVENS, AND B. BYERS. 2000. Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids. *Plant Cell* 12: 1551-1568.

COMESAÑA, A. S., J. E. TORO, D. J. INNES, AND R. J. THOMPSON. 1999. A molecular approach to the ecology of a mussel (*Mytilus edulis-M. trossulus*) hybrid zone on the east coast of Newfoundland, Canada. *Mar Biol* 133: 213-221.

CORANDER, J., P. MARTTINEN, J. SIRÉN, AND J. TANG. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9: 539.

CORANDER, J., P. WALDMANN, P. MARTTINEN, AND M. J. SILLANPÄÄ. 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20: 2363-2369.

CORANDER, J., P. WALDMANN, AND M. J. SILLANPÄÄ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367-374.

COSTEDOAT, C., N. PECH, R. CHAPPAZ, AND A. GILLES. 2007. Novelties in hybrid zones: crossroads between population genomic and ecological approaches. *PLoS One* 2: e357.

CRONN, R., AND J. F. WENDEL. 2004. Cryptic trysts, genomic mergers, and plant speciation. *New Phytol* 161: 133-142.

CRUZAN, M. B., AND M. L. ARNOLD. 1993. Ecological and genetic associations in an *Iris* hybrid zone. *Evolution* 47: 1432-1445.

CRUZAN, M. B., AND M. L. ARNOLD. 1994. Assortative mating and natural selection in an *Iris* hybrid zone. *Evolution* 48: 1946-1958.

DALE, A., AND T. T. ELKINGTON. 1974. Variation within *Cardamine pratensis* L. in England. *Watsonia* 10: 1-17.

DECARLI, N., AND L. LEINEMANN. 2003. Tetraploide 1.0 – Software zur Analyse genetischer Strukturen autotetraploider Arten.

DOBEŠ, C., C. KIEFER, M. KIEFER, AND M. A. KOCH. 2007. Plastidic *trn*FUUC pseudogenes in North American genus *Boechera* (Brassicaceae): mechanistic aspects of evolution. *Plant Biol (Stuttg)* 9: 502-515.

DOBEŠ, C. H., T. MITCHELL-OLDS, AND M. A. KOCH. 2004. Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). *Mol Ecol* 13: 349-370.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

DOYLE, J. J., L. E. FLAGEL, A. H. PATERSON, R. A. RAPP, D. E. SOLTIS, P. S. SOLTIS, AND J. F. WENDEL. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annu Rev Genet* 42: 443-461.

EHRICH, D., M. GAUDEUL, A. ASSEFA, M. A. KOCH, K. MUMMENHOFF, S. NEMOMISSA, AND C. BROCHMANN. 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Mol Ecol* 16: 2542-2559.

ELLSTRAND, N. C. 1992. Gene flow by pollen: Implications for plant conservation genetics. *Oikos* 63: 77-86.

EXCOFFIER, L., G. LAVAL, AND S. SCHNEIDER. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.

FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.

FLAGEL, L. E., AND J. F. WENDEL. 2009. Gene duplication and evolutionary novelty in plants. *New Phytol* 183: 557-564.

FRANK, S. A. 1989. The evolutionary dynamics of cytoplasmic male sterility. *Am Nat* 133: 345-376.

GOUDET, J. 2000. FSTAT, a program to estimate and test gene diversities and fixation indices.

GREGORIUS, H. R. 1974. Genetischer Abstand zwischen Populationen – I. Zur Konzeption der genetischen Abstandsmessung. *Silvae Genetica* 23: 1-3.

HAAS, F., AND A. BRODIN. 2005. The crow *Corvus corone* hybrid zone in southern Denmark and northern Germany. *Ibis* 147: 649-656.

HARRISON, R. G. 1993. Hybrid zones and the evolutionary process. Oxford University Press, Oxford.

HEGARTY, M. J., AND S. J. HISCOCK. 2005. Hybrid speciation in plants: new insights from molecular studies. *New Phytol* 165: 411-423.

HEISER, C. B., JR. 1951. Hybridization in the annual sunflowers: *Helianthus annuus x H. debilis* var. *cucumerifolius*. *Evolution* 5: 42-51.

HOVAV, R., J. A. UDALL, B. CHAUDHARY, R. RAPP, L. FLAGEL, AND J. F. WENDEL. 2008. Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. *Proc Natl Acad Sci U S A* 105: 6191-6195.

JONSELL, B., K. KUSTÅS, AND I. NORDAL. 1995. Genetic variation in *Arabis petraea*, a disjunct species in northern Europe. *Ecography* 18: 321-332.

KLEIN, J. 2008. Phylogeographie des *Arabidopsis arenosa* (L.) Lawalree - Artkomplexes. University of Heidelberg.

KOCH, M. A., C. DOBEŠ, K. G. BERNHARDT, AND J. KOCHJAROVÁ. 2003. *Cochlearia macrorrhiza* (Brassicaceae): A bridging species between *Cochlearia* taxa from the Eastern Alps and the Carpathians? *Plant Syst Evol* 242: 137-147.

KOCH, M. A., C. DOBEŠ, C. KIEFER, R. SCHMICKL, L. KLIMEŠ, AND M. A. LYSAK. 2007. Supernetwork identifies multiple events of plastid *trn*F(GAA) pseudogene evolution in the Brassicaceae. *Mol Biol Evol* 24: 63-73.

KOCH, M. A., C. DOBEŠ, M. MATSCHINGER, W. BLEEKER, J. VOGEL, M. KIEFER, AND T. MITCHELL-OLDS. 2005. Evolution of the *trn*F(GAA) gene in *Arabidopsis* relatives and the Brassicaceae family: monophyletic origin and subsequent diversification of a plastidic pseudogene. *Mol Biol Evol* 22: 1032-1043.

KOCH, M. A., AND M. KIEFER. 2005. Genome evolution among cruciferous plants: A lecture from the comparison of the genetic maps of three diploid species – *Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. *American Journal of Botany* 92: 761-767.

KOCH, M. A., AND M. MATSCHINGER. 2007. Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 104: 6272-6277.

KOCH, M. A., M. WERNISCH, AND R. SCHMICKL. 2008. *Arabidopsis thaliana´s* wild relatives: an updated overview on systematics, taxonomy and evolution. *Taxon* 57: 933-943.

KUITTINEN, H., A. A. DE HAAN, C. VOGL, S. OIKARINEN, J. LEPPÄLÄ, M. KOCH, T. MITCHELL-OLDS, C. H. LANGLEY, AND O. SAVOLAINEN. 2004. Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. *Genetics* 168: 1575-1584.

LEWIS, D. 1941. Male sterility in natural populations of hermaphrodite plants. *New Phytol* 40: 56-63.

LOCKLEY, A. K. 1992. The position of the hybrid zone between the house sparrow *Passer domesticus domesticus* and the Italian sparrow *P. d. italiae* in the Alpes Martimes. *J Ornith* 133: 77-82.

MACHADO, C. A., T. S. HASELKORN, AND M. A. NOOR. 2007. Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 175: 1289-1306.

MALL, C. 2005. Hybridartbildung innerhalb der Gattung *Cardaminopsis* in Ostösterreich – Untersuchungen zur morphologischen Variabilität. University of Heidelberg.

MALLET, J. 2007. Hybrid speciation. *Nature* 446: 279-283.

MEUSEL, H., E. JÄGER, AND E. WEINERT. 1965. Vergleichende Chorologie der zentraleuropäischen Flora. Gustav Fischer Verlag, Jena.

MULLER, H. J. 1914. A new mode of segregation in Gregory´s tetraploid primulas. *Am Nat* 48: 508-512.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

OZKAN, H., M. TUNA, AND D. W. GALBRAITH. 2006. No DNA loss in autotetraploids of *Arabidopsis thaliana*. *Plant Breeding* 125: 288-291.

PODANI, J. 2001. SYN-TAX 2000: Computer program for data analysis in ecology and systematics. Scientia Publishing, P.O. Box 658, H-1365 Budapest, Hungary.

PREVOSTI, A., J. OCAÑA, AND G. ALONSO. 1975. Distances between populations of *Drosophila subobscura*, based on chromosome arrangement frequencies. *Theor Appl Genet* 45: 231-241.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

RAMOS-ONSINS, S. E., B. E. STRANGER, T. MITCHELL-OLDS, AND M. AGUADÉ. 2004. Multilocus analysis of variation and speciation in the closely related species *Arabidopsis halleri* and *A. lyrata*. *Genetics* 166: 373-388.

RAMSEY, J., AND D. W. SCHEMSKE. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Ann Rev Ecol Syst* 29: 467-501.

RAMSEY, J., AND D. W. SCHEMSKE. 2002. Neopolyploidy in flowering plants. *Ann Rev Ecol Syst* 33: 589-639.

RAPP, R. A., J. A. UDALL, AND J. F. WENDEL. 2009. Genomic expression dominance in allopolyploids. *BMC Biol* 7: 18.

RHYMER, J. M., AND D. SIMBERLOFF. 1996. Extinction by hybridization and introgression. *Ann Rev Ecol Syst* 27: 83-109.

RIESEBERG, L. H. 1995. The role of hybridization in evolution: Old wine in new skins. *Am J Bot* 82: 944-953.

RIESEBERG, L. H. 1997. Hybrid origins of plant species. *Ann Rev Ecol Syst* 28: 359-389.

RIESEBERG, L. H., AND N. C. ELLSTRAND. 1993. What can molecular and morphological markers tell us about plant hybridization? *Critical Reviews in Plant Sciences* 12: 213-241.

RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol Trends Pl* 5: 65-84.

RIESEBERG, L. H., J. WHITTON, AND K. GARDNER. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 152: 713-727.

RIESEBERG, L. H., J. WHITTON, AND C. R. LINDER. 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Bot Neerl* 45: 243-262.

RIESEBERG, L. H., A. WIDMER, A. M. ARNTZ, AND J. M. BURKE. 2003. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos Trans R Soc Lond B Biol Sci* 358: 1141-1147.

RIESEBERG, L. H., AND J. H. WILLIS. 2007. Plant speciation. *Science* 317: 910-914.

RILEY, H. P. 1938. A character analysis of colonies of *Iris fulva*, *Iris hexagona* var. *giganticaerulea* and natural hybrids. *Am J Bot* 25: 727-738.

RUEGG, K. 2008. Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution* 62: 452-466.

SAVOLAINEN, O., C. H. LANGLEY, B. P. LAZZARO, AND H. FRÉVILLE. 2000. Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. *Mol Biol Evol* 17: 645-655.

SCHMICKL, R., M. H. JØRGENSEN, A. K. BRYSTING, AND M. A. KOCH. 2008a. Phylogeographic implications for the North American boreal-arctic *Arabidopsis lyrata* complex. *Plant Ecol Div* 1: 245-254.

SCHMICKL, R., C. KIEFER, C. DOBEŠ, AND M. A. KOCH. 2008b. Evolution of *trn*F(GAA) pseudogenes in cruciferous plants. *Plant Syst Evol*.

SOLTIS, D. E., AND L. H. RIESEBERG. 1986. Autopolyploidy in *Tolmiea menziesii* (Saxifragaceae): Genetic insights from enzyme electrophoresis. *Am J Bot* 73: 310-318.

SOLTIS, D. E., AND P. S. SOLTIS. 1989. Tetrasomic inheritance in *Heuchera micrantha* (Saxifragaceae). *J Hered* 80: 123-126.

SOLTIS, D. E., P. S. SOLTIS, D. W. SCHEMSKE, J. F. HANCOCK, J. N. THOMPSON, B. C. HUSBAND, AND W. S. JUDD. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56: 13-30.

SOLTIS, D. E., P. S. SOLTIS, AND J. A. TATE. 2004. Advances in the study of polyploidy since *Plant speciation*. *New Phytol* 161: 173-191.

SOLTIS, P. S., AND D. E. SOLTIS. 2009. The role of hybridization in plant speciation. *Annu Rev Plant Biol* 60: 561-588.

STEBBINS, G. L. 1947. Types of polyploids: their classification and significance. *Advances in Genetics* 1: 403-429.

SZOMBATHY, J. 1884. Eine paläolithische Fundstätte im Löss bei Willendorf in Niederösterreich. *Mitteilungen der Anthropologischen Gesellschaft in Wien* XIV. Band: 35-36.

SZOMBATHY, J. 1910. Die diluvialen Kulturschichten von Willendorf. *Mitteilungen der Anthropologischen Gesellschaft in Wien* XXXX. Band: 4-9.

TOMEKPE, K., AND R. LUMARET. 1991. Association between quantitative traits and allozyme heterozygosity in a tetrasomic species: *Dactylis glomerata*. *Evolution* 45: 359-370.

TORO, J., D. J. INNES, AND R. J. THOMPSON. 2004. Genetic variation among life-history stages of mussels in a *Mytilus edulis-M. trossulus* hybrid zone. *Mar Biol* 145: 713-725.

UJHÁZYOVÁ, M., AND K. UJHÁZY. 2004. Prehl´ad rastlinných spoločenstiev bukových lesov na vápencoch bradlového pásma. *Bull Slov Bot Spolocn* 11: 1-7.

VAN HERWERDEN, L., AND P. J. DOHERTY. 2006. Contrasting genetic structures across two hybrid zones of a tropical reef fish, *Acanthochromis polyacanthus* (Bleeker 1855). *J Evol Biol* 19: 239-252.

VAN TREUREN, R., H. KUITTINEN, K. KÄRKKÄINEN, E. BAENA-GONZALEZ, AND O. SAVOLAINEN. 1997. Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*. *Mol Biol Evol* 14: 220-229.

WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

WENDEL, J. F., AND V. A. ALBERT. 1992. Phylogenetics of the cotton genus (*Gossypium*): Character-state weighted parsimony analysis of chloroplast-DNA restriction site data and its systematic and biogeographic implications. *Syst Bot* 17: 115-143.

WERNISCH, M. 2007. Molecular systematics, phylogeography and evolution of the genus *Cardaminopsis* HAYEK (Brassicaceae), the closest relatives of the model plant *Arabidopsis thaliana* (L.) HEYNH. Universität für Bodenkultur, Wien.

WHITHAM, T. G., G. D. MARTINSEN, K. D. FLOATE, H. S. DUNGEY, B. M. POTTS, AND P. KEIM. 1999. Plant hybrid zones affect biodiversity: Tools for a genetic-based understanding of community structure. *Ecology* 80: 416-428.

WHITHAM, T. G., P. A. MORROW, AND B. M. POTTS. 1991. Conservation of hybrid plants. *Science* 254: 779-780.

WOLFE, K. H. 2001. Yesterday's polyploids and the mystery of diploidization. *Nat Rev Genet* 2: 333-341.

WRIGHT, S. I., B. LAUGA, AND D. CHARLESWORTH. 2003. Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. *Mol Ecol* 12: 1247-1263.

YATABE, Y., N. C. KANE, C. SCOTTI-SAINTAGNE, AND L. H. RIESEBERG. 2007. Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* 175: 1883-1893.

YATABE, Y., C. TSUTSUMI, Y. HIRAYAMA, K. MORI, N. MURAKAMI, AND M. KATO. 2009. Genetic population structure of *Osmunda japonica*, rheophilous *Osmunda lancea* and their hybrids. *J Plant Res*.

Acknowledgements

This research project was based on many people´s help and advice, and I express my deep gratitude to the following:

Prof. Marcus Koch, my supervisor, for freedom in research, scientific guidance, confidence, and encouragement. It was a great pleasure to work on the "*Arabidopsis*" project.

Prof. Hans-Peter Comes for the agreement of being my second supervisor and for inspiration in textbook and article reading.

Michi (Dr. Michaela Wernisch) for introducing me both to the former "*Cardaminopsis*" and its field localities, for unforgettable times in the Wachau and in Vienna, and for hosting me so many times (many thanks also to Klemens).

My parents, Dr. Willi and Elisabeth Schmickl, for their constant support and encouragement. My father for practical help in mathematical problems, for always figuring out the most appropriate Excel equation, for field trip assistance in Franconian Switzerland, and for taking very good care of all the *Arabidopsis* and numerous other nonstandard plants in my parents´ garden.

Pat (Patrick Weiss) for his constant support and encouragement, for practical help with the hard- and software of my notebook, for construction and mass production of "plant condoms" to protect *Arabidopsis* against evil pollinators, and for field trip assistance in Austria, Franconian Switzerland, and the Black Forest.

Dr. Adolf Polatschek for sharing "*Cardaminopsis*" chromosome counts and for very helpful discussions.

Dr. Jürgen Ehlers and Prof. Phil Gibbard for kindly and rapidly providing me various shapefiles of the maximal extent of the LGM.

Many thanks to all who exchanged plant material with me: Marte Holten Jørgensen, Dr. Barbara Mable, and Dr. David Remington. Thanks to the Natural History Museum and Herbarium Vienna (Dr. Ernst Vitek) and the Natural History Museum and Herbarium London for letting me do destructive sampling on numerous herbarium specimen.

Christina Mall (Staatsexamensarbeit, morphometrics), Holger Baldauf (Staatsexamensarbeit, chromosome counts), Stefan Wötzel (Bachelor thesis, evolutionary history of the *Arabidopsis lyrata* complex), and Johannes Klein (Diploma thesis, evolutionary history of the *Arabidopsis arenosa* complex) for working on various aspects of this project.

Christiane (Dr. Christiane Kiefer) for fantastic "Duo Infernale" times.

All members of the AG Koch and AG Erbar for a great working atmosphere, especially Fabian, Ingrid, Juraj, Lotti, Markus, Myriam, Paola, Peter, Robert, and Susanne.

Markus (Dr. Markus Kiefer) for programming of two software packages and for help in several critical situations with my computer.

Susanne (Susanne Ball) for help with DNA extraction of many, many samples and for the PCRs of the *Pgi*C part.

The team of the Botanical Garden Heidelberg, especially Torsten (Torsten Jakob), for taking very good care of the *Arabidopsis* plants.

DFG for funding this project, the Graduate Academy Excellence Initiative for a scholarship throughout the last months, and SYNTHESYS for funding my stay at the Natural History Museum Vienna.

„Die Bekanntschaft mit den Naturwissenschaften, mit der wissenschaftlichen Methode hat mich immer wachsam bleiben lassen, und ich habe mich bemüht, mein Schreiben dort, wo es möglich war, mit den wissenschaftlichen Gegebenheiten in Einklang zu bringen, wo dies hingegen unmöglich war, zog ich es vor, gar nicht zu schreiben…"

Anton Čechov