A poly approach to ploidy:
Polyploid evolution and taxonomic implications

Marte Holten Jørgensen

CEES

Dissertation presented for the degree of Philosophiae Doctor

Centre for Ecological and Evolutionary Synthesis, Department of Biology

Faculty of Mathematics and Natural Sciences

University of Oslo

2011

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1136

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Kan du'kke bare levere'a?

Lille far, januar 2011.

Contents

Preface

When my project is about to finish, it's time for acknowledgements:

- Annen. ANNEANNEANNE! Fineste sjefstøsen! My main supervisor Anne K. Brysting. Best supervisor in the world, I'm sure. You are my colleague, my friend, my third mum (I have a modern family), and my hero. As well as scientifically strong, you show a care for and a faith in your students that is really remarkable. Your door is always open, even literally, for all kinds of problems, personal as well as scientifical. If I didn't have you, I would have quitted a long time ago. But you were there, and you kept your faith in me when I didn't. Thanks for all your chatter, your support, and your hugs! I wonder if you'll ever be able to say Vålenga kjærke!
- Barbs. My wonderful co-supervisor Barbara Mable at the Glasgow Uni. Your hospitality when welcoming me to Glasgow meant a lot. And your patience with my ignorance on S-alleles and inheritage and crossings. From now on, my focus will be on our joint project!
- Hilde, agent Empetrum, thanks for always being there. Thanks for sharing my tears and my laughter. Thanks to you, Julie, Ragnhild, Flua, Helge, and Einar for quiz nights every Tuesday. And thanks to you, Vir, and Inger for wine colloquia and a nice master period. Snilleskrede: thanks for reading!
- Torbæibi, my wise and nice colleague Tor Carlsen, thanks for all the nice scientific discussions and all the coffea!
- Thanks to Svend-Ole Nielsen (my Danish friend), Elisabeth Hubbuck, and Cecilia Leslie for professional help with a messy brain.
- Thanks to my family, particularly my father for being supportive, taking me on boat trips, trips to our cottage, and for using me as a source of knowledge in molecular

biology. Thanks to my sister for always being there, and letting me use her kids in hug therapy. Thanks to my brother for breaks in Svalbard and particularly for taking me to Paris on a two day's notice. Thanks to my grandpa for financial help throughout my studies. I'm sorry you couldn't be there to see me finish, farfar!

- Thanks to all my co-authors, particularly Doro, Andi (my hero! Thanks for the title!), Tor, and Reidar.
- Thanks to Cecilie for help in the lab and all the chatter and cigarettes.
- Thanks to my supervisors in my master; Christian, Reidar, Brita, and Tove, for making me learn enough to be able to do a PhD (I hope!).
- Thanks to Mika, Ola, Unni, Inger, and the rest of the botanical group at Blindern for all your support and all the discussions. And Ave for chatter in the office.
- Thanks to Karin for giving me an anxiety attack and sweeping up the remains.
- Thanks to the Department of Biology and the CEES for giving me three months extra salary. Thanks to Kristine Bonnevie travelling stipend, S. G. Sønneland Foundation, and Hans Henrik Maschmann's legacy for financial support.
- Thanks to the people at the phytotron for taking care of all my plants.
- Thanks to Marc, Katya, Liza, Barry, Peter, Clair, Ash, Andy, and Katherine at the Glasgow Uni for making my stays in Glasgow full of joy.
- Thanks to Hege, Hilde L., Anette, Martha, Helle, Heidi, Gard, and other dear old friends for their patience with a depressed PhD student.
- Thanks to Herr Trøst for all the hugs throughout this period of trial! And to my father for giving him to me.

Oslo, 080811,

List of papers

I. **Jørgensen MH**, Elven R, Tribsch A, Gabrielsen TM, Stedje B, Brochmann C. 2006. Taxonomy and evolutionary relationships in the *Saxifraga rivularis* complex. *Systematic Botany,* 31: 702-729.

II. Westergaard K, **Jørgensen MH**, Gabrielsen TM, Alsos IG, Brochmann C. 2010. The extreme Beringian/Atlantic disjunction in *Saxifraga rivularis* (Saxifragaceae) has formed at least twice. *Journal of Biogeography,* 37: 1262-1276.

III. **Jørgensen MH**, Carlsen T, Skrede I, Elven R. 2008. Microsatellites resolve the taxonomy of the polyploid *Cardamine digitata* aggregate (Brassicaceae). *Taxon,* 57: 882-892.

IV. Schmickl R, **Jørgensen MH**, Brysting AK, Koch MA. 2010. The evolutionary history of the *Arabidopsis lyrata* complex: A hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. *BMC Evolutionary Biology,* 10: 98.

V. **Jørgensen MH**, Brysting AK, Takebayashi N, Steets JA, Beecher JJ, Wolf DE. Digging deeper into the allopolyploid origin of *Arabidopsis kamchatica* and its effect on molecular diversity. Manuscript.

VI. **Jørgensen MH**, Ehrich D, Schmickl R, Koch MA, Brysting AK. Interspecific and interploidal gene flow in Central European *Arabidopsis* (Brassicaceae). Manuscript submitted to *BMC Evolutionary Biology*.

Background

In school I learned about evolution as a gradual change of states. I remember memorising the jingle mutation - adaptation - selection - evolution: random mutations sometimes lead to adaptation, advantageous character changes which increase an organism's chance of survival and reproduction, i.e. fitness. Natural selection will spread the character state in the population until it becomes fixed. Evolution is then the process from mutation to fixation. Speciation, or the origin of species, was illustrated as a process where a population is split into two by a geographical barrier, and the two are gradually changed by adaptation to different environments. When they have accumulated so many differences that they no longer can reproduce if the barrier is removed, then they are considered different species. Horses and donkeys are different species because the mule is sterile. Furthermore the evolutionary relationship among organisms was presented as a tree where the root represents the beginning of life, the nodes are ancestors, and the leaves are extant organisms (Fig. 1). Now, this childhood learning of mine is not erroneous. The speciation description is the book example of allopatric speciation (Fig. 2), which accounts for the majority of speciation events in animals (Coyne and Orr, 2004). Reproductive isolation is the only defining criteria of Mayr's biological species concept (e.g., Mayr, 2000), widely accepted and used

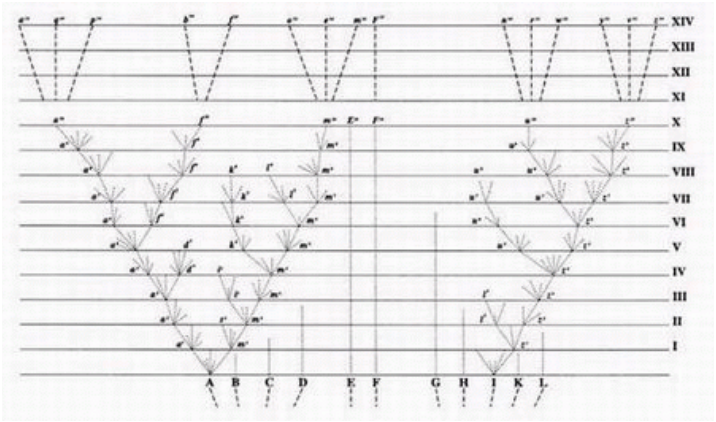


Figure 1. Darwin's Tree of Life. From The Origin of Species (1859).

among zoologists. And finally, trees are unquestionably the most common way of illustrating evolutionary ancestor-descendant or sibling relationships. However, moving from animals to the Plant Kingdom complicates the matter, and the primary reason is the prevalence of polyploidy among plants.

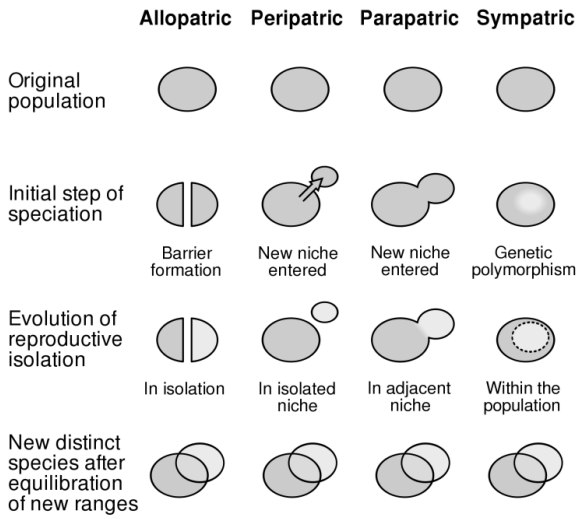


Figure 2. Different modes of speciation. From Wikipedia (http://en.wikipedia.org/wiki/Allopatric_speciation).

The word polyploidy is derived from Greek. The literal meaning is manifold and refers to number of chromosome sets in an organism, i.e. a diploid has two chromosome sets, a triploid has three and a polyploid has many (Fig. 3A). Throughout most of the 20th century it was generally believed that polyploids mostly arose through a somatic increase in chromosome number (e.g. Levin, 2002), i.e. through a chromosome replication without subsequent cell division in the mitotic cycle. Such a mitotic non-reduction is indeed not uncommon and gives for instance rise to the polyploid tissue in the human liver (Biesterfeld et al., 1994). However, the formation of polyploid organisms is mostly through meiotic non-reduction, i.e. through the production of diploid gametes (Harlan and De Wet, 1975). Ramsey

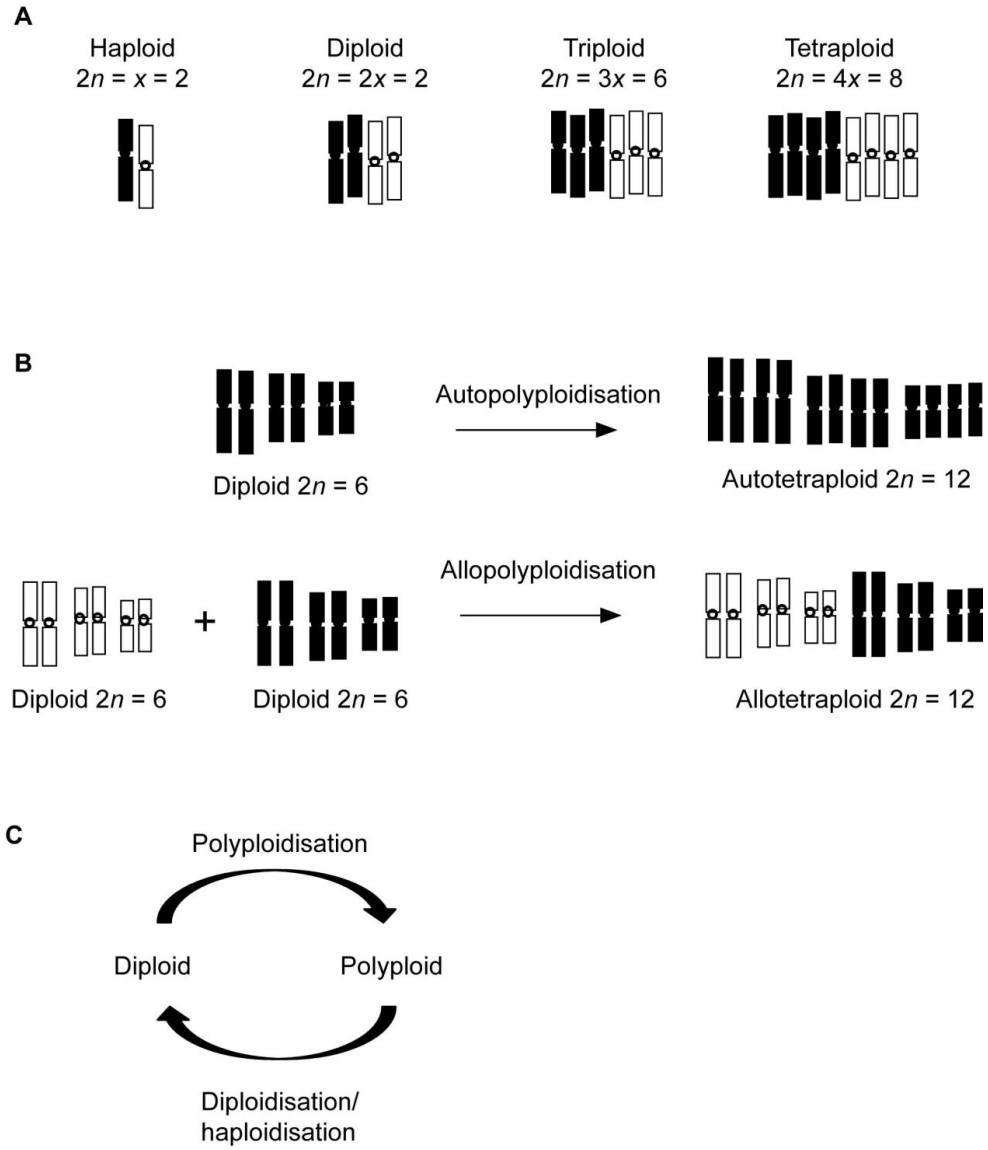


Figure 3. Ploidy definitions. **A:** A haploid organism contains a single set of chromosomes, a diploid contains two, a triploid three, a tetraploid four, etc. **B:** The difference between auto- and allopolyploidy. Autopolyploidy is the duplication of a genome within a single species. Allopolyploidy is the result of hybridisation between species with subsequent duplication. The alternative cytogenetic definitions are that autopolyploids contain duplicated genomes that create multivalents during meiosis, whereas allopolyploids continue to have disomic inheritance with bivalent formation of homologous chromosomes. **C:** Polyploidy as a dynamic process.

and Schemske (1998) estimated that the mean production of diploid gametes in flowering plants is 0.56 %. The number may seem small, but considering that a single grass shoot could produce half a million pollen grains (McKone, 1989), about 250 of these would be diploid. Traditionally, polyploidisation has been considered a rare process, and each polyploid species has therefore typically been thought to be of single origin (Soltis and Soltis, 1993). Contrary to this, Soltis and Soltis (1993, 1999) demonstrate that the majority of polyploid plants is of recurrent origins, and that single origin of polyploids is rare (Soltis et al., 2003). Recurrent polyploid formation might be facilitated by presence of triploids through a so-called triploid bridge: A triploid plant may produce gametes with one ($n = x$), two ($n = 2x$) and three ($n = 3x$) chromosome sets, and simulations suggest that this may suffice to cause recurrent polyploid formation, or gene flow from diploids to polyploids subsequent to the polyploid formation (Husband and Sabara, 2003, Husband, 2004).

We often distinguish between two different kinds of polyploids: autopolyploids arise from a single species, and allopolyploids are the results of hybridisation between two different species (Fig. 3B; Ramsey and Schemske, 1998, Levin, 2002). As these definitions are based on species circumscriptions, they are often referred to as taxonomic (e.g., Stebbins, 1947, Grant, 1971). The original definitions are cytogenetic and based on chromosome pairing behaviour, first published by Kihara and Ono (1926). In both kinds of definitions, however, allo- and autopolyploids are considered extremes of a continuum from the doubling of identical genomes to the doubling of highly differentiated genomes (e.g., Parisod et al., 2010). Darlington (1937) suggested an inverse relationship between the fertility of a diploid hybrid and that of a tetraploid to which it gives rise, i.e., if the diploid hybrid is fully fertile, the tetraploid is not. Three recent studies have supported this, suggesting that parental divergence drives polyploidy (Chapman and Burke, 2007, Buggs et al., 2008, Paun et al., 2009). This might explain why allopolyploids are more common than autopolyploids, but as stressed by

Soltis et al. (2007); the prevalence of autopolyploids is probably underestimated as they are more difficult to detect in natural populations. Furthermore, calculations by Ramsey and Schemske (1998) made them conclude that the rate of autopolyploid formation may often be higher than the rate of allopolyploid formation.

Authors' views on just *how* prevailing polyploidy is among plants have changed through time; the newer the paper, the higher the number. The good old guys thought that about half (or less) of the angiosperms are polyploid (Müntzing, 1936, Darlington, 1937, Stebbins, 1950, Grant, 1963). In the 80's and 90's, several authors calculated that about 70-80 % are polyploid based on available chromosome numbers (Goldblatt, 1980, Lewis, 1980) and stomata size (Masterson, 1994). Whereas today, authors seem to doubt the existence of any true diploid angiosperms (e.g., Soltis et al., 2003). Similarly, in the year 2000, Otto and Whitton estimated that 2-4 % of speciations in angiosperms, and 7 % in ferns, involve polyploidisation (Otto and Whitton, 2000). In 2009 Wood et al. adjusted these numbers to 15 and 31 %, respectively (Wood et al., 2009). Regardless of where we end up; polyploidy is the single most important mode of sympatric speciation in plants (Soltis et al., 2003).

If all plants are polyploid and polyploidisations occur frequently, one would expect that mean genome size increase through time. However, even though *Arabidopsis thaliana* has three, possibly four, whole genome duplications in its evolutionary history, its genome is really small ($2n = 10$, 1C = 0.215 pg, The Arabidopsis Genome Initiative, 2000, Vision et al., 2000, Simillion et al., 2002, Bowers et al., 2003, Schmuts et al., 2004, Maere et al., 2005, Fawcett et al., 2009). The reason for this is that polyploidisation often is followed by massive chromosomal rearrangements and loss (e.g., Wolfe, 2001, Tate et al., 2009, Mandáková et al., 2010). This process, called diploidisation, will through time make the polyploid into a functional diploid (Fig. 3C). As DNA replication costs energy, removing unneeded genetic material will be an advantage for the polyploid. Another way of becoming diploid is through

haploidisation; the reverse of polyploidisation (e.g., Turcotte and Feaster, 1963, Raven and Thompson, 1964). This has been experimentally demonstrated in the crop plant *Sorghum* (Duara and Stebbins, 1952) and has also been suggested for other plant groups (Turcotte and Feaster, 1963, De Wet, 1965, 1968, Anderson, 1972). Raven and Thompson (1964) suggest that there is no theoretical reason why such polyhaploids haven't played a part in the evolution of some taxa, and De Wet (1965, 1968, 1971) presents what he believes are valid data to support this phenomenon. Stebbins (1970, 1971) argues that polyhaploidy will add nothing to the diploid gene pool, and that polyploids are incapable of producing viable polyhaploids because of diploidisation, or because the polyhaploids are weak or sterile or both. Jackson (1976) argues that until experimental evidence shows competitive superiority of polyhaploids, Stebbin's argument should be accepted as valid. Later authors have ignored the subject completely (e.g., Soltis and Soltis, 1999, Soltis et al., 2003, Soltis et al., 2010). Haploidisation may, however, play a role in making gene flow from polyploids to their diploid progenitors possible (Kloda et al., 2008, paper 6).

Polyploidisation may lead to new ecological preferences through new gene combinations and doses, and may therefore be advantageous (Ramsey, 2011). Polyploids usually spread from the periphery of the diploid's niche space, but the ecological contrasts may vary in character (Levin, 2002). For example in *Chamerion angustifolium*, the diploids occupy the coldest climate, the hexaploids the warmest, and the tetraploids are in between (Mosquin, 1967). In *Empetrum nigrum* it is the other way around; the tetraploids are more cold tolerant than the diploids (Hagerup, 1927, Suda, 2002). The frequency of polyploids increases with latitude in the Northern hemisphere, and Hagerup (1932) suggested that polyploids are better adapted than diploids to extreme climates. Stebbins (1984, 1985) suggested that there is a correlation between polyploid frequencies and degree of glaciation, rather than with latitude *per se*. Brochmann et al. (2004) found that the frequency of

polyploids in the Arctic is higher in previously glaciated than non-glaciated areas, but they also found a correlation between latitude and ploidy level. They conclude that the evolutionary success of polyploids in the Arctic may be based on their fixed-heterozygous genomes, which buffer against inbreeding and genetic drift through periods of dramatic climate change (Brochmann et al., 2004). This is in agreement with studies dating whole genome duplication events, suggesting that plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event (Fawcett et al., 2009, Soltis and Burleigh, 2009).

Polyploidy has also been associated with a change in mating system; the allotetraploid *Arabidopsis kamchatica* is self-compatible while both its diploid parents are self-incompatible (Mable et al., 2004, Shimizu and Purugganan, 2005). In *Empetrum* the tetraploid has hermaphroditic flowers, while the diploid has uni-sexual flowers (Hagerup, 1927). As selfing may play an important role in polyploid establishment, Stebbins (1950) suggested that polyploid plants have higher rates of self-fertilization than their diploid progenitors. This is supported by Barringer (2007) who compared data for 235 species of flowering plants with known ploidy levels. However, Mable (2004) found no strong association between self-compatibility and ploidy.

In this study I discuss polyploid origins in four different plant groups. I discuss which molecular methods that are suitable to use when studying the origins of polyploids, if recurrent origins really is the rule for polyploids, and the taxonomic implications of polyploidy. Even though developments in molecular biology and technology have dramatically increased our knowledge on polyploid evolution the last couple of decades, there is still a lot that we don't know or where we don't have a general agreement (Soltis et al., 2010). The majority of polyploids seem to be of recurrent origins (Soltis and Soltis, 1993, Soltis and Soltis, 1999), but introgression from the diploid parents may give the same patterns

as recurrence (Soltis et al., 2010). There is no general agreement on the taxonomic treatment of polyploids (e.g., Soltis et al., 2007); shall we classify polyploids as species or as conspecific with the progenitors? I thereby hope to elucidate some of the still controversial aspects of polyploid evolution.

Case studies

Papers 1 and 2: The *Saxifraga rivularis* complex

Saxifraga rivularis L. is part of a small arctic-alpine species complex in *Saxifraga* section *Mesogyne*. The section has a complex evolutionary history of polyploidisation and reticulation: The local endemics *S. svalbardensis* and *S. oppdalensis* are both allopolyploids with the same parents (Brochmann et al., 1998). More than one ploidy level are known in both *S. cernua* and *S. radiata* (Elven et al., 2010). And *S. hyperborea*, traditionally considered diploid, is palaeotetraploid (Brochmann et al., 1998). Earlier studies have suggested that *S. rivularis* is an allopolyploid with *S. hyperborea* as the maternal parent (Brochmann et al.,

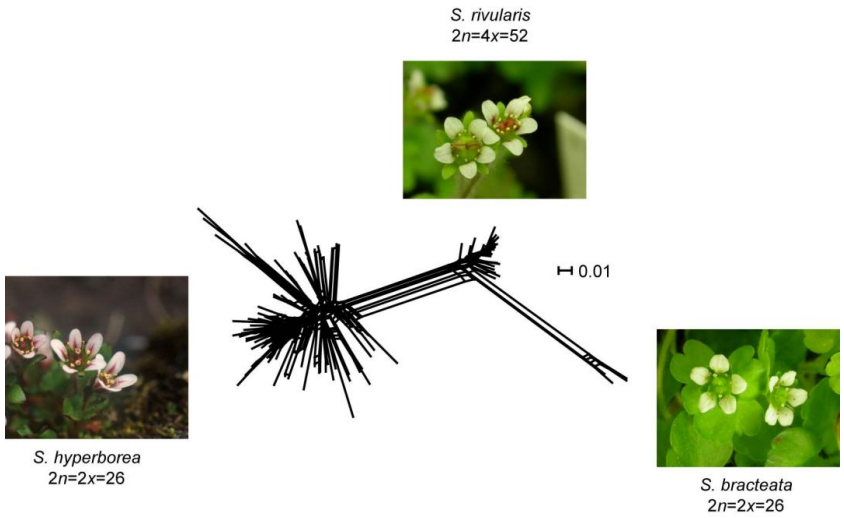


Figure 4. The allopolyploid origin of *Saxifraga rivularis*. The neighbour network is based on Dice similarity and 250 AFLP markers. Photo: HH Grundt and MH Jørgensen.

1998, Guldahl et al., 2005). Using AFLPs (amplified fragment length polymorphisms) we show in paper 1 that the second parent is *S. bracteata*, an amphi-Pacific diploid (Fig. 4). In paper 2 *S. rivularis* is studied in depth using AFLPs and five plastid sequences. The lack of diversity found in both studies combined with *S. rivularis* being almost strictly selfing (Brochmann and Steen, 1999), suggests that a single polyploidisation event sufficed for the establishment of this species.

Paper 3: The *Cardamine digitata* aggregate

Cardamine L. is a large, nearly global genus in Brassicaceae. It contains a few extant diploids, but most species are polyploid with up to 32 genomes ($2n = 32x = 256$; Kučera et al., 2005), and are the result of many independent polyploidisation events (Carlsen et al., 2009). The small *C. digitata* aggregate (Fig. 5) in the amphi-Beringian area is well separated from its close relatives. All plants studied so far are tetra- to dodecaploid. In paper 3 we use six microsatellite loci to identify evolutionary lineages, and find that the aggregate contains four genetically equidistant units. These correspond in morphology with the described species *C. blaisdelli*, *C. digitata*, *C. microphylla*, and *C. purpurea*. As each of them contains at least two ploidy levels, we suggest that recurrent autopolyplodisation events are important within this aggregate.

Papers 4 and 5: *Arabidopsis kamchatica*

The small genus *Arabidopsis* (Brassicaceae) is an excellent model system for studying the evolutionary effects of genome duplications as it includes both allo- and autoployploids, and powerful molecular tools are available. The allotetraploid *A. suecica* is already studied extensively, for instance, through artificial neo-polyploidisations (e.g., Wang et al., 2004), made possible by the identification of its origin: a single hybridisation and polyploidisation between *A. thaliana* and *A. arenosa* (O'Kane et al., 1996, Lind-Halldén et al., 2002, Jakobsson et al., 2006). In paper 4 and 5 we address the parental origin of another allopolyploid: *A.*

Figure 5. The *Cardamine digitata* aggregate: *C. blaisdellii* (top), *C. digitata* (above middle), *C. microphylla* (below middle), and *C. purpurea* (bottom).

kamchatica (Fig. 6). The combination of nuclear ITS and plastid *trn*L-F sequences (paper 4) and six low-copy nuclear markers (paper 5) tells a convincing story of hybridisations between

Figure 6. The allotetraploid *Arabidopsis kamchatica*. Photo: J. Beecher (left) and N. Takebayashi (right).

A. lyrata ssp. *umbrosa* and *A. halleri* ssp. *gemmifera* in NE Asia leading to the establishment of the allopolyploid. Both studies also suggest that this has happened several times, or alternatively that there has been subsequent gene flow from the parents to the descendant. A comparison of *A. suecica* and *A. kamchatica* may give insights as to what causes or facilitates recurrent polyploidisation events.

Paper 6: *Arabidopsis arenosa* and *A. lyrata*

Effects of polyploidisation on gene flow between natural populations are yet little known. In paper 6 we study diploid and auotetraploid populations of the two sister species *A. arenosa* and *A. lyrata* (Fig. 7). We use a combination of plastid and low-copy nuclear sequences, and show that there's interspecific gene flow between the tetraploids even though the diploids are reproductively isolated. This suggests that polyploidy may buffer against negative effects of interspecific

Figure 7. *Arabidopsis arenosa* (left) and *A. lyrata* (right). Or: What to do when you brain's messed up by science. Photo: MH Jørgensen and J Bråte.

hybridisation, as has been previously suggested (De Wet and Harlan, 1970, Harlan and De Wet, 1975). Within each species, we demonstrate bidirectional interploidal gene flow, particularly in *A. lyrata*, suggesting that ploidy changes not necessarily lead to almost total reproductive isolation, as traditionally believed.

Discussion

Pros and cons of markers - is there a perfect one?

During the last decade the most commonly used molecular markers for plant phylogeny and phylogeography have been sequencing of a single or several plastid sequences, often in combination with the nuclear ribosomal DNA region ITS (internal transcribed spacer), cloning and sequencing of low-copy nuclear DNA regions, and DNA fingerprinting techniques such as AFLPs and microsatellites. In this thesis I use them all (Fig. 8): plastid DNA in papers 2 and 4, ITS in paper 4 and 6, low-copy nuclear regions in paper 5 and 6, AFLPs in paper 1 and 2, and microsatellites in paper 3. The choice of markers may depend on the taxonomic level that you study (i.e., if high you need conservative markers), practical

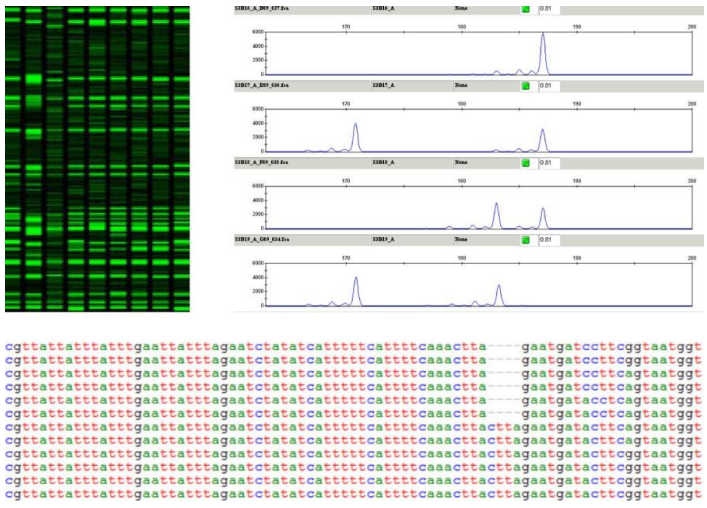


Figure 8. Molecular markers used in this thesis: Amplified fragment length polymorphisms (AFLPs, top left), microsatellites (top right), and DNA sequence data (bottom).

considerations such as time and money consumption, and of course the markers available for the organisms of interest.

Plastids (e.g., chloroplasts) are usually inherited from one parent only, thus most plants contain a single copy of plastid DNA (Birky Jr., 1995). Therefore, plastid DNA regions can be sequenced directly by e.g., Sanger sequencing (Sanger et al., 1977), and the method is low-cost and efficient. Furthermore, there are plenty of general primers available (e.g., Taberlet et al., 1991). Plastid DNA in plants is rather conserved (e.g., Small et al., 2004). This is good when working on higher taxonomic levels (e.g., family or above), but a challenge when working on lower taxonomic levels (e.g., paper 2 where no intraspecific variation was found in five different regions, Fig. 9). Also, the uniparental inheritance of plastids is a drawback when the origin of a polyploid is studied, as the sequences tell only one side of the

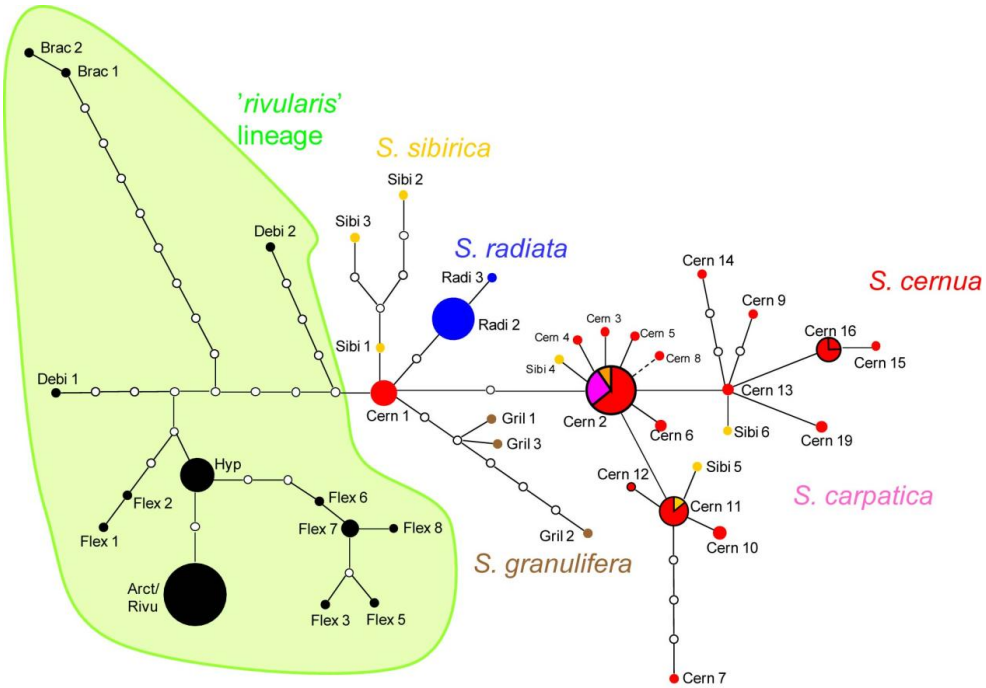


Figure 9. Phylogenetic relationships in *Saxifraga* section *Mesogyne*. The haplotype network is based on the five plastid DNA regions published for *S. rivularis* in paper 2 (for the remaining taxa: T. M. Gabrielsen, unpublished). Taxon names are abbreviated. For the *S. rivularis* lineage: arct/rivu - *S. rivularis*, brac - *S. bracteata*, debi - *S. debilis*, and flex/hyp - *S. hyperborea*.

story. The plastid of the allotetraploid *S. rivularis* is, thus, inherited from *S. hyperborea*, and no signs of the other ancestor *S. bracteata* are visible in the dataset (paper 2; Fig. 9).

Sequencing of ITS is one of the most common markers for phylogeny reconstructions. It is non-coding, i.e. not expressed, and therefore rather fast evolving (Kay et al., 2006). However, the flanking regions on each side are strongly conserved, and universal primers have been available for a long time for most organismal groups (White et al., 1990, Gardes and Bruns, 1993, Ji et al., 2003, Blackwood et al., 2005). As part of the ribosomal DNA, ITS is highly repeated in the genome (e.g., Matyášek et al., 2007). This has the positive consequence that it is easy to PCR amplify. You can mostly sequence the region directly, and the method is therefore rather work efficient and cheap. The repeatedness is, however, also the major drawback of this marker for the study of polyploid evolution, as the region as a result undergoes what we call concerted evolution (Liao, 1999, De Sousa Queiroz et al., 2011). This means that a hybrid with originally two different parental sets of ITS sequences will lose one of them in just a few generations; they all become alike. Thus, an allopolyploid of some age will probably retain the ITS from one parental species only, while the other is lost. Thus, sequencing of ITS alone will not necessarily identify both parents, and might lead to an erroneous conclusion of autopolyploidy. If you're lucky, a combination of ITS and one or several plastid DNA regions will give both sides of the story, and in paper 4, that is the case: Japanese allotetraploid *A. kamchatica* contains *A. halleri*-like plastids and *A. lyrata*-like ITS, whereas Chinese *A. kamchatica* has the other way around.

Low- or single-copy nuclear regions are bi-parentally inherited as they are nuclear, and as they are low-copy, they are not influenced by concerted evolution (e.g., Sang, 2002, Small et al., 2004). In these aspects they are superior to both ITS and cpDNA (hurrah!). Are we closing in on the perfect marker? Not really. Fewer general primers are available, and you mostly need to clone prior to sequencing (or develop paralog-specific primers, e.g.,

Marcussen et al., 2011), making the method time-consuming and expensive. In the resulting dataset, mistakes made by the DNA polymerase are more prevailing than in datasets from direct sequencing; both single nucleotide errors and recombination will occur (e.g., Eckert and Kunkel, 1991). And distinguishing between real and polymerase-made recombination becomes a matter of educated guesswork. Furthermore, the cloning process is often biased as to which alleles are inserted into the vector, so you may have to produce loads of sequences for each specimen in order to make sure you have them all. We initially wanted to include PgiC as well as CHS and *sc*ADH in paper 6, but it became too complicated, we had to make far too many sequences per specimen. Too much work and too expensive to be worth the effort. Using next generation sequencing instead of cloning may resolve the last problem because the high number of reads will increase the chance of catching all copies, but the polymerase error problems remain (MH Jørgensen, unpublished). As to resolve the question of polyploid origin, low-copy regions have worked well both in paper 5 and paper 6. In paper 5 close to every individual of the allotetraploid *A. kamchatica* retained both *A. lyrata*-like and *A. halleri*-like alleles. In paper 6, tetraploid *A. arenosa* contained (with a few exceptions) the same alleles as diploid *A. arenosa*, whereas tetraploid *A.lyrata* contained the same alleles as diploid *A. lyrata*. And as we sequenced so many clones that we are quite certain we got all the alleles, a conclusion of autopolyploidy can be safely drawn. In paper 4, we combine ITS and cpDNA with genotyping of the low-copy nuclear PgiC region. We designed primers that amplified the *A. halleri* version of PgiC1 only, and could thereby distinguish between *A. lyrata* and *A. kamchatica*: both would contain *A. lyrata*-like alleles of ITS and cpDNA, but only *A. kamchatica* would get positive PCRs from our PgiC1 (*A. halleri*) primers.

AFLP is a fingerprinting method based on restriction cutting (Vos et al., 1995). It is therefore a general method that can be used for any organism without prior knowledge about its genome. Thus little methodological development or optimisation is necessary. Depending

on what enzymes and which primers you choose, AFLPs are highly variable. Although we found no variation in *S. rivularis* in five plastid regions, we found enough AFLP variation to be able to see some geographical patterns and to distinguish between subspecies in papers 1 and 2. A major drawback is that AFLPs are dominant markers, i.e. cannot distinguish between homozygous and heterozygous specimens. Available statistical tools are therefore scarce (Bonin et al., 2007). In paper 2, F_{ST} statistics would have been a natural choice as we deal with population genetics, but impossible to include as we use AFLPs. Also the use of parsimony phylogeny reconstruction (paper 1) is controversial with AFLPs (Koopman, 2005). As bands are scored as either present or absent only, you cannot be absolutely certain that present bands are homologous. Also as the method is universal, fragments of endophytes may amplify along with the plant host. However, Koopman (2005) compared an AFLP dataset with an ITS dataset for the same specimens, and found the same pattern in both, suggesting that the phylogenetic signal is higher than the noise of homoplasy or contamination. In paper 1, our AFLP analyses agree with our morphological analyses, suggesting that the AFLPs make sense. Another major drawback is the necessity of high quality DNA for the results to be reproducible (Bonin et al., 2004). In paper 1 the combination of limited amount of leaf tissue and the use of a column DNA extraction protocol gave low concentrations of DNA. The first time I analysed the dataset, the resulting AFLP noise overshadowed any taxonomic or geographical signal. In a PCO plot the specimens simply grouped according to PCR plate. Using replicates and other precautions (e.g. removing specimens with significantly higher or lower number of bands than the average for the given ploidy level) reduced the dataset to approximately 1/3 of the initial number of bands, but removed the noise and brought about the phylogenetic signal. Another effect of the need of high quality DNA is a limitation of available material. The tissue must be quickly dried using silica gel, otherwise the DNA becomes too degraded, and you cannot be certain that the absence of a band isn't simply due

to amplification failure. In paper 1 this significantly reduced our AFLP dataset. We had herbarium material of plenty of specimens that we included in the morphological analysis, but couldn't be included in the AFLP analyses. We had to exclude all samples of the potential ancestor *S. debilis* and all but three Beringian samples of the allopolyploid *S. rivularis*. So even though the intermediate position, based on AFLPs, of *S. rivularis* between *S. bracteata* and *S. hyperborea* (Fig. 4) suggests that the latter two are the ancestors of the allopolyploid, *S. debilis* cannot be excluded as an ancestor by the AFLP dataset. Furthermore, as Beringia is the area of origin and therefore the center of diversity for *S. rivularis*, the single polyploidisation event we conclude with in paper 1 could be due to lack of sampling. However, in paper 2, the Beringian sampling is substantially better, and the conclusion stands.

Microsatellites are, as AFLPs, hyper-variable markers, but opposed to AFLPs they are co-dominant (e.g., Li et al., 2002). Thus a higher number of analytical tools are available for microsatellites (e.g., Selkoe and Toonen, 2006). They are cheap and efficient to use once they have been developed, but as they're not general, development may be both expensive and time-consuming. Low resolution was found in both ITS and several cpDNA regions for *Cardamine* by Carlsen et al. (2009), so to study the evolution within the *C. digitata* aggregate (paper 3), we needed markers with better resolution. The study relied on extensive use of herbarium specimens, so AFLPs weren't an alternative (see above). As Skrede was planning to develop microsatellites for her studies in *Draba* (also Brassicaceae, Skrede et al., 2008, Skrede et al., 2009a, Skrede et al., 2009b), we joined forces to produce a microsatellite dataset for the *C. digitata* aggregate (Skrede et al., 2009b£ paper 3). In paper 3 six microsatellites loci give enough resolution to distinguish between the closely related polyploid species within this aggregate, and thereby enough to suggest that the hexaploids are autopolyploid descendants from the tetraploids or extinct (or not sampled) diploids in the same lineage. If a microsatellite marker is single-copied, the maximum number of bands for a diploid specimen is two, for a

tetraploid it's four, etc. Thus microsatellites can also give an indication of ploidy level when flow cytometry analyses or chromosome counts are impossible because of sampling. In paper 3 we find up to nine bands in *C. microphylla* although chromosome counts have identified only tetra-, hexa-, and octoploids. We therefore suggest that higher ploidy levels exist in this lineage. The high number of bands for polyploids is, however, also a drawback with this method. For low-ploid (di- and tetraploid) specimens scoring microsatellites is fairly easy. In diploids, homozygotes have a single allele, heterozygotes have two. In tetraploids, dosage can be used to distinguish between different partial heterozygous individuals (i.e., AAAB, AABB, and ABBB, Schmickl, 2009). But to score dosage differences unambiguously when you have nine different bands is impossible. Thus, in paper 3 we treat the markers as phenotypes rather than genotypes, and thereby lose a lot of potential information and available analytical tools become as limited as for AFLP data.

The different marker systems have different pros and cons, but in combinations, they can all give satisfactory results. The saying 'Many roads lead to Rome' is as true in evolution as in daily affairs (Mayr, 1963). The entry of genomics, i.e. whole genome sequencing, has been welcomed by many as a solution to all problems (e.g., Kahvejian et al., 2008, Rokas and Abbot, 2009). And as the technology improves and becomes cheaper, genomics becomes more available. Different markers tell different stories, so sequencing the whole genome must give us the whole truth and cover up the incongruence often found when comparing single gene trees. Or? Slot and Rokas (2011) demonstrated horisontal transfer of a large gene cluster between fungi, tangling the tree of life. Three studies (Dunn et al., 2008, Philippe et al., 2009, Schierwater et al., 2009) dealing with the early diversification of animals produced highly incongruent findings despite the use of considerable sequence data (49-128 genes). And Phillippe et al. (2011) demonstrate how differences in alignment, taxon sampling, and choice of evolutionary models can influence the same results to incongruence. Thus, phylogenies

depend on much more than simply the number of markers, and even though genomics may bring us closer to the truth, it doesn't necessarily take us all the way.

Recurrence of polyploidy - is there a rule?

I don't know how many times I've read that recurrent formation of polyploids is the rule rather than the exception, originally formulated by Soltis and Soltis (1993, 1999), and cited many times, also by me (papers 3 and 5). In their 1993 paper, they list 46 polyploid species studied with molecular methods, 39 of which have multiple origins, the last seven are listed as "maybe". In their 1999 paper, they expand the list with 15 new examples of polyploid taxa with recurrent origin, whereas in their 2003 review, Soltis et al. (2003) list only four examples of single origin polyploids (Fig. 10). Seven years later they treat the recurrence rule as a widely accepted fact (Soltis et al., 2010). Recurrence obviously has a positive impact on polyploids, as it enables the incorporation of

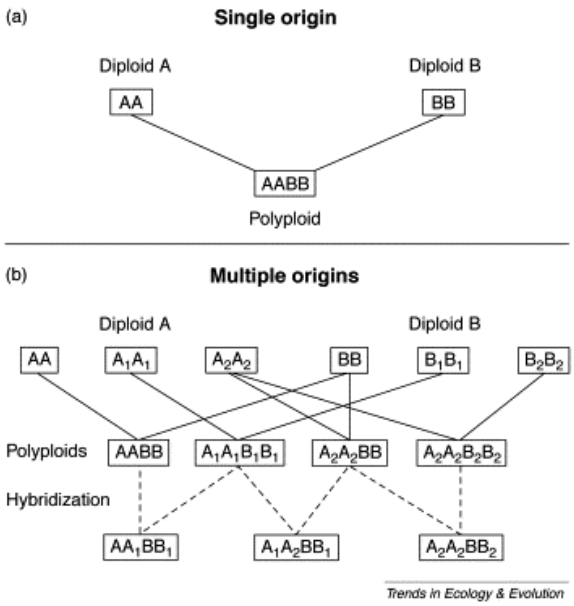


Figure 10. Soltis and Soltis (1999): Fig. 1. Comparison of (a) traditional view of polyploid formation with (b) new or revised view. The traditional view envisioned each polyploid species forming only once, resulting in a new species that was genetically uniform (or nearly so). The new view suggests that each polyploid species forms over and over again from different parental genotypes generating a diverse array of polyploid genotypes. Subsequent hybridization among these polyploid genotypes and recombination result in additional genetic variability.

genetic diversity from multiple populations of their diploid progenitors (Soltis and Soltis, 1993, Tate et al., 2005). Simulation studies of Ramsey and Schemske (1998) suggested an important role of recurrent formation for the success of a new autopolyploid to establish in nature; recurrent formation is critical to counterbalance local extinction of small populations at the initial stages of autopolyploid establishment. And as high rates of unreduced gamete production sometimes occur (up to 26 % 2*n* eggs, De Haan et al., 1992, up to 73 % 2*n* pollen, Maceira et al., 1992, both in *Dactylis glomerata* L.), the odds of recurrence aren't necessarily very low.

If there is a rule, then the allotetraploid *S. rivularis* is probably one of the exceptions. In paper 1 we found little variation in the AFLP dataset, and in paper 2 we demonstrate no variation at all in five different regions of plastid DNA. In comparison, the parental diploid *S. hyperborea* contains many different haplotypes (T. M. Gabrielsen, unpublished results; Fig. 10), particularly in the area where we suggest the hybridisation happened, i.e. Beringia. Thus if *S. rivularis* had recurrent origins, we would expect to see more of the variation found in *S. hyperborea* in *S. rivularis* as well. Or haplotypes inherited from the other ancestor, *S. bracteata*. But this is not the case.

In paper 3 we conclude that polyploidisations have occurred several times in the *C. digitata* aggregate, but we don't really conclude for each of the four documented lineages. Having plotted the known chromosome numbers on a phylogenetic tree (Fig. 11), however, I suggest that hexaploid *C. digitata* (2*n* = 42) has arisen at least three times, as the hexaploid specimens are placed in three different parts of the tree, and closer to tetraploids (2*n* = 28) than to each other. For the other three lineages, we have included too few chromosome vouchers to be able to conclude.

Following the rule, *A. kamchatica* has recurrent origins. In paper 4 we demonstrate recurrent and reciprocal hybridisation events: Japanese *A. kamchatica* has *A. halleri* as mother

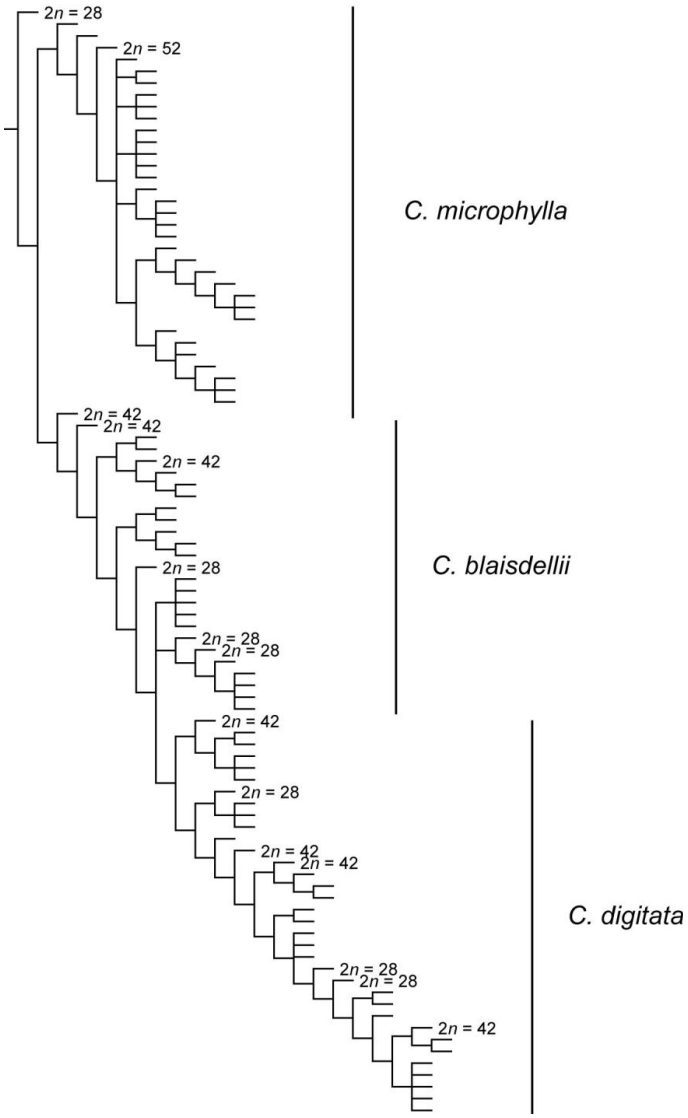


Figure 11. Recurrent formation of polyploids in the *Cardamine digitata* aggregate: Chromosome counts plotted on a strict consensus of six most parsimonious trees from the six loci microsatellite dataset in paper 3, excluding *C. purpurea*.

and *A. lyrata* as father, Chinese *A. kamchatica* has the other way around, and Russian *A. kamchatica* has a third origin. In paper 5 we conclude with "several independent origins" but aren't specific as to the number. The diversity found in both studies is low, thus we cannot conclude with many origins. On the other hand, *A. kamchatica* is almost strictly selfing (e.g.,

Mable et al., 2004), and low diversity could be due to the selfing nature of the species and/or fast distribution expansion following formation (founder effect).

In the last case study, we conclude that both autotetraploid *A. arenosa* and *A. lyrata* have independent origins. In the presented networks, the tetraploids are almost completely intermingled with the diploids. At least in *A. arenosa*, this is probably due to recurrent origins, as the cytotypes are allopatric today. In *A. lyrata*, the cytotypes are sympatric, and ongoing interploidal gene flow may cover any possible signs of recurrence (Schmickl, 2009£ paper 6).

To summarise; of the eight polyploid species studied in papers 1-6, three are probably of recurrent origins (the allopolyploid *A. kamchatica* and the autopolyploid *C. digitata*, and *A. arenosa*), one is of a single origin (the allopolyploid *S. rivularis*), and four are still unknown (the autopolyploid *C. blaisdellii*, *C. microphylla*, *C. purpurea*, and *A. lyrata*). Thus, the case studies presented here seem follow the rule of Soltis and Soltis (1993, 1999) with exceptions. However, could these patterns arise from other mechanisms than recurrent formations? How can we distinguish between recurrent formation and subsequent gene flow from one (or both) of the parents? At least when the cytotypes are sympatric, interploidal gene flow may occur through unreduced gamete formation in the diploid, or through triploid bridges (Husband and Sabara, 2003, Husband, 2004). Harlan and De Wet (1975) list 85 plant genera known to produce unreduced gametes, and all the species listed in a review of Bretagnolle and Thompson (1995) produce 2*n* gametes to some extent, both eggs and pollen. Thus, even though a triploid block may reduce the frequency of triploids among di- and tetraploids, and therefore the probability of a triploid bridge (e.g., Köhler et al., 2009), gene flow from diploids to tetraploids may still occur. In paper 6 we try to make the distinction between recurrence and back-crossing by using isolation with migration analyses (Hey and Nielsen, 2007, Hey, 2010). However, even though these analyses can distinguish between old and recent gene flow, they cannot distinguish between recurrent origins and single origin with

subsequent, but old and ceased, introgression. Soltis et al. (2010) call for criteria and methods to make such distinctions. In paper 6, but we suggest that *Arabidopsis* is a promising study group for that purpose.

Taxonomy in a polyploid world

All taxonomy is based on resemblance among organisms, either in morphology, anatomy, chemistry, or genetics. In 1957, two important works were published on how to measure resemblance: Michener and Sokal (1957) and Sneath (1957), further developed by Sokal and Sneath (1963). They suggest in their Principles of Numerical Taxonomy, that resemblance should be based on phenetics, i.e. overall similarity. In 1966, Hennig published his Phylogenetic Systematics (Hennig, 1966), claiming that systematics should be based on cladistics and shared derived character states, instead of total likeness. Further, he claimed that all species (and higher taxonomic levels) should be monophyletic groups of organisms, i.e. containing a single ancestor and all its descendants. This view has been dominating in the following decades, particularly among zoologists. Some even suggest to leave the Linnean system of ranks and binomials, and propose a new set of rules to name monophyletic groups only, the so-called PhyloCode (Cantino and De Queiroz, 2010). And if evolution is tree-like, the logic is easy to follow.

The Hennigian definition of monophyly contains two components; common ancestry and inclusiveness (Hörandl, 2007). However, the roots of the word monophyly mean a tribe from one, i.e. common ancestry only. Ashlock (1971) therefore suggests that we should recognize two different ways of being monophyletic: 1) holophyletic = inclusive groups (monophyletic sensu Hennig), and 2) paraphyletic = non-inclusive, i.e. not containing *all* descendants of a common ancestor. Forty years have passed since Ashlock's suggestion, and one should think that the biological community would reach an agreement within this time span, but no. Six years ago, Nordal and Stedje made a petition for the acceptance of

paraphyletic taxa (Nordal and Stedje, 2005), and collected 150 signatures from plant systematists world wide. This letter to the editor of the journal Taxon initiated an intense debate in the following volumes (Dias et al., 2005, Potter and Freudenstein, 2005, Williams et al., 2005, Alexander, 2006, Brummitt, 2006, Ebach et al., 2006, Hörandl, 2006, 2007, Van Wyk, 2007, Zander, 2007, Stuessy and König, 2008, Zander, 2008, Stuessy, 2009, Hörandl, 2010). The supporters of paraphyletic taxa claim that when a new set of characters has arisen within a group, according to cladistic principles one must either 1) not give the newly recognisable group any taxonomic rank, since it will leave the remainder of the parental group paraphyletic, or 2) split the remainder of the parental group into different taxa even though there are no characters to recognise them (Nordal and Stedje, 2005). Further they claim that inclusiveness is not necessarily connected with evolution, that in fact the majority of speciation processes are probably *not* inclusive because they are not necessarily connected with extinction (Hörandl, 2006, 2007). The Hennigian school, on the other hand, claims that no information can be retrieved from a paraphyletic group. Taxonomic assemblages that share no unique permutations of characteristics are simply not groups at all (Dias et al., 2005, Williams et al., 2005, Ebach et al., 2006).

However, if we accept monophyly sensu Hennig as a rule, what happens with polyploid taxa and their ancestors? What would be the consequence if all descendants should be included in the parental taxon? In the case of the *S. rivularis* complex (papers 1 and 2), this would mean that *S. rivularis* should be lumped with *S. bracteata* and *S. hyperborea* into a single species. Furthermore, as *S. rivularis* is one of the parents of the local endemics and allopolyploids *S. svalbardensis* and *S. oppdalensis*, with *S. cernua* as the other parent (Brochmann et al., 1998), further lumping is necessary. In effect almost all, if not all, of *Saxifraga* section *Mesogyne* should be treated as a single large and highly polymorph species. Similarly, *A. kamchatica* (papers 4 and 5) should be lumped with its parental taxa *A. hallerii*

and *A. lyrata* into a single species. And as we see signs of hybridisation between *A. lyrata* and *A. arenosa* (paper 6), these should also be lumped to ensure monophyly. Furthermore, the allotetraploid *A. suecica* is a descendant from diploid *A. arenosa* and *A. thaliana* (O'Kane et al., 1996, Lind-Halldén et al., 2002, Jakobsson et al., 2006), thus the giant *Arabidopsis* species must be further expanded to include almost the whole genus. In both cases, i.e. *Saxifraga* section *Mesogyne* and *Arabidopsis*, the resulting large species will be highly polymorphic and consist of populations that are reproductively isolated, morphologically, genetically and cytologically distinct, and ecologically highly different from each other. This will be the case for all polyploid complexes. Even genera will need to be combined; for instance in Poaceae alone there are plenty of examples of hybrid genera (Elven et al., 2010).

Classification is a consequence of man's need to deal with his environment, and the responsibility of taxonomists towards society should not be neglected (Van Wyk, 2007). Our taxonomy must be practical as it has very many users (Brummitt, 2006). The classificatory component of taxonomy cannot itself be made into a science by ill-founded philosophy or essentially arbitrary numerical procedures (Johnson, 1970). Monophyly sensu Hennig will drastically reduce the number of plant taxa, as shown above, and thereby remove names that contain lot of information. If I write about *S. rivularis* the reader will know that it concerns the circum-polar allotetraploid. Without using the name, communication becomes difficult. We are simply merging everything into one amorphous plesiomorphic soup (Brummitt, 2006). Both taxonomic schools use strong words and claim that the opposing school is illogical (e.g., Dias et al., 2005, Nordal and Stedje, 2005). Maybe it's time to "join hands and work together towards best serving the needs of society" (Van Wyk, 2007). Otherwise taxonomy will risk being denoted as "yet another ivory tower science - a pursuit disconnected from practical concerns and everyday life, esoteric, over-specialized, its classifications of little practical use to the majority of end-users" (Van Wyk, 2007).

Literature cited

Alexander PJ. 2006. Descent and modification in evolutionary systematics. *Taxon,* **55**: 4.

Anderson LC. 1972. *Flaveria campestris* (Asteraceae): a case of polyhaploidy or relic ancestral diploidy? *Evolution,* **26**: 671-673.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature,* **408**: 796-815.

Ashlock PD. 1971. Monophyly and associated terms. *Systematic Zoology,* **20**: 63-69.

Barringer BC. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany,* **94**: 1527-1533.

Biesterfeld S, Gerres K, Fischer-Wein G, Böcking A. 1994. Polyploidy in non-neoplastic tissues. *Journal of Clinical Pathology,* **47**: 38-42.

Birky Jr. CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **92**: 11331-11338.

Blackwood CB, Oaks A, Buyer J. 2005. Phylum- and class-specific PCR primers for general microbial community analysis. *Applied and Environmental Microbiology,* **71**: 6193-6198.

Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology,* **13**: 3261–3273.

Bonin A, Ehrich D, Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology,* **16**: 3737-3758.

Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature,* **422**: 433–438.

Bretagnolle F, Thompson JD. 1995. Gametes with somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist,* **129**: 1-22.

Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen A-C, Elven R. 2004. Polyploidy in arctic plants. *Botanical Journal of the Linnean Society,* **82**: 521–536.

Brochmann C, Steen SW. 1999. Sex and genes in the flora of Svalbard - implications for conservation biology and climate change. *Det Norske Vitenskaps-Akademi. I. Matematisk Naturvitenskapelig Klasse. Skrifter. Ny serie,* **38**: 33–72.

Brochmann C, Xiang Q-Y, Brunsfeld SJ, Soltis DE, Soltis PS. 1998. Molecular evidence for polyploid origins in *Saxifraga* (Saxifragaceae): the narrow arctic endemic *S. svalbardensis* and its widespread allies. *American Journal of Botany,* **85**: 135–143.

Brummitt RK. 2006. Am I a bony fish? *Taxon,* **55**: 268-269.

Buggs RJA, Soltis PS, Mavrodiev E, Symonds VV, Soltis DE. 2008. Does phylogenetic distance between parental genomes govern the success of polyploids? *Castanea,* **73**: 74-93.

Cantino PD, De Queiroz K. 2010. PhyloCode: a phylogenetic code of biological nomenclature. Version 4c. http://www.ohio.edu/phylocode/.

Carlsen TA, Bleeker W, Hurka H, Albert VA, Elven R, Brochmann C. 2009. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Annals of the Missouri Botanical Garden,* **96**: 215-236.

Chapman MA, Burke JM. 2007. Genetic divergence and hybrid speciation. *Evolution,* **61**: 1773-1780.

Coyne JA, Orr HA. 2004. *Speciation,* Sunderland, Sinauer Associates, Inc.

Darlington CD. 1937. *Recent advances in cytology,* Philadelphia, P. Blakiston's Son and Co.

Darwin C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life,* London, John Murray.

De Haan AA, Maceira NO, Lumaret R, Delay J. 1992. Production of 2*n* gametes in diploid subspecies of *Dactylis glomerata* L. 2. Occurrence and frequency of 2*n* eggs. *Annals of Botany,* **69**: 345-350.

De Sousa Queiroz C, De Carvalho Batista FR, De Oliveira LO. 2011. Evolution of the 5.8S nrDNA gene and the internal transcribed spacer in *Carapichea ipecacuanha* (Rubiaceae) within a phylogeographic context. *Molecular Phylogenetics and Evolution,* **59**: 293-302.

De Wet JMJ. 1965. Diploid races of tetraploid *Dichantium* species. *The American Naturalist,* **99**: 167-171.

De Wet JMJ. 1968. Diploid-tetraploid-haploid cycles and the origin of variability in *Dichantium* agamospecies. *Evolution,* **22**: 394-397.

De Wet JMJ. 1971. Reversible tetraploidy as an evolutionary mechanism. *Evolution,* **25**: 545-548.

De Wet JMJ, Harlan JR. 1970. Apomixis, polyploidy, and speciation in *Dichanthium*. *Evolution,* **24**.

Dias P, Assis LCS, Udulutsch RG. 2005. Monophyly vs. paraphyly in plant taxonomy. *Taxon,* **54**: 1039-1040.

Duara BN, Stebbins GL. 1952. A polyhaploid obtained from a hybrid derivative of *Sorghum halepense* x *S. vulgare* var. *sudanense*. *Genetics,* **37**: 369-374.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, Sørensen MV, Haddock SHD, Schmidt-Rhaesa A, Okusu A, Kristensen RM, Wheeler WC, Martindale MQ, Giribet G. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature,* **452**: 745-750.

Ebach MC, Williams DM, Morrone JJ. 2006. Paraphyly is bad taxonomy. *Taxon,* **55**: 831-832.

Eckert KA, Kunkel TA. 1991. DNA polymerase fidelity and the polymerase chain reaction. *Genome Research,* **1**: 17-24.

Elven R, Murray DF, Razzhivin VY, Yurtsev BA. 2010. *Panarctic flora checklist. Vascular plants. Draft version,* Oslo, National Centre for Biosystematics, University of Oslo.

Fawcett JA, Maere S, Van de Peer Y. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America,* **106**.

Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology,* **2**: 113-118.

Goldblatt P. 1980. Polyploidy in angiosperms: monocotyledons. In: Lewis WH ed. *Polyploidy: biological relevance.* New York, Plenum Press. 213-239.

Grant V. 1963. *The origin of adaptation,* New York, Colombia University Press.

Grant V. 1971. *Plant speciation,* New York, Colombia University Press.

Guldahl AS, Gabrielsen TM, Scheen A-C, Borgen L, Spjelkavik S, Steen SW, Brochmann C. 2005. The *Saxifraga rivularis*/*hyperborea* species complex in Svalbard: a molecular-morphological approach. *Flora,* **200**: 207-221.

Hagerup O. 1927. *Empetrum hermaphroditum* (Lge.) Hagerup. A new tetraploid, bi-sexual species. *Dansk Botanisk Arkiv,* **5**: 1-17.

Hagerup O. 1932. Über Polyploidie in Beziehung zu Klima, Ökologie und Phylogenie. *Hereditas,* **16**: 19-40.

Harlan JR, De Wet JMJ. 1975. On Ö. Winge and a prayer: the origins of polyploidy. *The Botanical Review,* **41**: 361-390.

Hennig W. 1966. *Phylogenetic systematics,* Urbana, University of Illinois Press.

Hey J. 2010. Isolation with migration models for more than two populations. *Molecular Biology and Evolution,* **27**: 905-920.

Hey J, Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America,* **104**: 2785-2790.

Husband BC. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biological Journal of the Linnean Society,* **82**: 537-546.

Husband BC, Sabara HA. 2003. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist,* **161**: 703-713.

Hörandl E. 2006. Paraphyletic versus monophyletic taxa - evolutionary versus cladistic classification. *Taxon,* **55**: 564-570.

Hörandl E. 2007. Neglecting evolution is bad taxonomy. *Taxon,* **56**: 1-5.

Hörandl E. 2010. Beyond cladistics: Extending evolutionary classifications into deeper time levels. *Taxon,* **59**: 345-350.

Jackson RC. 1976. Evolution and systematic significance of polyploidy. *Annual Review of Ecology and Systematics,* **7**: 209-234.

Jakobsson M, Hagenblad J, Tavaré S, Säll T, Halldén C, Lind-Halldén C, Nordborg M. 2006. A unique recent origin of the allotetraploid species *Arabidopsis suecica*: evidence from nuclear DNA markers. *Molecular Biology and Evolution,* **23**: 1217-1231.

Ji Y-J, Zhang DX, He L-J. 2003. Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. *Molecular Ecology Notes,* **3**: 581-585.

Johnson LAS. 1970. Rainbow's end: the quest for an optimal taxonomy. *Systematic Zoology,* **19**: 203-239.

Kahvejian A, Quackenbush J, Thompson JF. 2008. What would you do if you could sequence everything? *Nature Biotechnology,* **26**: 1125-1133.

Kay KM, Whittall JB, Hodges SA. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology,* **6**: 36.

Kihara H, Ono T. 1926. Chromosomenzahlen und systematische Gruppierung der *Rumex*-Arten. *Zeitschrift für Zellforschung und mikroskopische Anatomie,* **4**: 475-481.

Kloda JM, Dean PDG, Maddren C, MacDonald DW, Mayes S. 2008. Using principal component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example from British Restharrows (*Ononis spinosa* and *Ononis repens*). *Heredity,* **100**: 253-260.

Koopman WJM. 2005. Phylogenetic signal in AFLP data sets. *Systematic Biology,* **54**: 197–217.

Kučera J, Valko I, Marhold K. 2005. On-line database of the chromosome numbers of the genus *Cardamine* (Brassicaceae). *Biologia, Bratislava,* **60**: 473-476.

Köhler C, Scheid OM, Erilova A. 2009. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics,* **26**: 142-148.

Levin DA. 2002. *The role of chromosomal change in plant evolution,* New York, Oxford University Press, Inc.

Lewis WH. 1980. Polyploidy in angiosperms: dicotyledons. In: Lewis WH ed. *Polyploidy: biological relevance.* New York, Plenum Press. 241-268.

Li YC, Korol AB, Fahima T, Beiles A, Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology,* **11**: 2453–2465.

Liao D. 1999. Concerted evolution: molecular mechanism and biological implications. *American Journal of Human Genetics,* **64**: 24-30.

Lind-Halldén C, Halldén C, Säll T. 2002. Genetic variation in *Arabidopsis suecica* and its parental species *A. arenosa* and *A. thaliana*. *Hereditas,* **136**: 45-50.

Mable BK. 2004. Polyploidy and self-compatibility: is there an association? *New Phytologist,* **162**: 803-811.

Mable BK, Beland J, Di Berardo C. 2004. Inheritance and dominance of self-incompatibility alleles in polyploid *Arabidopsis lyrata*. *Heredity,* **93**: 476-486.

Maceira NO, De Haan AA, Lumaret R, Billon M, Delay J. 1992. Production of 2*n* gametes in diploid subspecies of *Dactylis glomerata* L. 1. Occurrence and frequency of 2*n* pollen. *Annals of Botany,* **69**: 335-343.

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y. 2005. Modeling gene and genome duplications in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America,* **102**: 5454-5459.

Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA. 2010. Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. *The Plant Cell,* **22**: 2277-2290.

Marcussen T, Jakobsen KS, Danihelka J, Ballard HE, Blaxland K, Brysting AK, Oxelman B. 2011. Inferring species networks from gene trees in high-polyploid North American and Hawaiian violets (*Viola*, Violaceae). *Systematic Biology,* **accepted**.

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

Matyášek R, Tate JA, Lim YK, Šrubařová H, Koh J, Leitch AR, Soltis DE, Soltis PS, Kovařík A. 2007. Concerted evolution of rDNA in recently formed *Tragopogon* allotetraploids is typically associated with an inverse correlation between gene copy number and expression. *Genetics,* **176**: 2509-2519.

Mayr E. 1963. *Animal species and evolution,* Cambridge, Harvard University Press.

Mayr E. 2000. The biological species concept. In: Wheeler QD, Meier R eds. *Species concepts and phylogenetic theory.* New York, Colombia University Press.

McKone MJ. 1989. Intraspecific variation in pollen yield in Bromegrass (Poaceae: *Bromus*). *American Journal of Botany,* **76**: 231-237.

Michener CD, Sokal RR. 1957. A quantitative approach to a problem in classification. *Evolution,* **11**: 130-162.

Mosquin T. 1967. Evidence for autopolyploidy in *Epilobium angustifolium* (Onagraceae). *Evolution,* **21**: 713-719.

Müntzing A. 1936. The evolutionary significance of autopolyploidy. *Hereditas,* **21**: 263-378.

Nordal I, Stedje B. 2005. Paraphyletic taxa should be accepted. *Taxon,* **54**: 5-6.

O'Kane SL, Schaal BA, Al-Shehbaz IA. 1996. The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. *Systematic Botany,* **21**: 559-566.

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

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

Paun O, Forest F, Fay MF, Chase MW. 2009. Hybrid speciation in angiosperms: parental divergence drives ploidy. *New Phytologist,* **182**: 507-518.

Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wöhreide G, Baurain D. 2011. Resolving difficult phylogenetic questions: Why more sequences are not enough. *PloS Biology,* **9**: e1000602.

Philippe H, Derelle R, Lopez P, Pick K, Borchiellini C, Boury-Esnault N, Vacelet J, Renard E, Houliston E, Quéinnec E, Silva CD, Wincker P, Guyader HL, Leys S, Jackson DJ, Schreiber F, Erpenbeck D, Morgenstern B, Wörheide G, Manuel M. 2009. Phylogenomics revives traditional views on deep animal relationships. *Current Biology,* **19**: 706-712.

Potter D, Freudenstein JV. 2005. Character-based phylogenetic Linnean classification: taxa should be both ranked and monophyletic. *Taxon,* **54**: 1033-1035.

Ramsey J. 2011. Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences of the United States of America,* **108**: 7096-7101.

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

Raven PH, Thompson HJ. 1964. Haploidy and Angiosperm evolution. *The American Naturalist,* **98**: 251-252.

Rokas A, Abbot P. 2009. Harnessing genomics for evolutionary insights. *Trends in Ecology & Evolution,* **24**: 192-200.

Sang T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology,* **37**: 121-147.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* **74**: 5463-5467.

Schierwater B, Eitel M, Jakob W, Osigus H-J, Hadrys H, Dellaporta SL, Kolokotronis S-O, DeSalle R. 2009. Concatenated analysis sheds light on early metazoan evolution and fuels a modern ‘‘Urmetazoon’’ hypothesis. *PloS Biology,* **7**: e1000020.

Schmickl R. 2009. *Reticulate evolution in glacial refuge areas - the genus Arabidopsis in the eastern Austrian Danube Valley (Wachau),* Doctor of Natural Sciences PhD thesis, Ruperto-Carola University of Heidelberg, Heidelberg.

Schmuts H, Meister A, Horres R, Bachmann K. 2004. Genome size variation among accessions of *Arabidopsis thaliana*. *Annals of Botany,* **93**: 317-321.

Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters,* **9**: 615-629.

Shimizu KK, Purugganan MD. 2005. Evolutionary and Ecological genomics of *Arabidopsis*. *Plant Physiology,* **138**: 578-584.

Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y. 2002. The hidden duplication past of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America,* **99**: 13627-13632.

Skrede I, Borgen L, Brochmann C. 2009a. Genetic structuring in three closely related circumpolar plant species: AFLP versus microsatellite markers and high-arctic versus arctic alpine distributions. *Heredity,* **102**: 293-302.

Skrede I, Brochmann C, Borgen L, Rieseberg LH. 2008. Genetics of intrinsic postzygotic isolation in a circumpolar plant species *Draba nivalis* (Brassicaceae). *Evolution,* **62**: 1840-1851.

Skrede I, Carlsen T, Rieseberg LH, Brochmann C. 2009b. Microsatellites for three distantly related genera in the Brassicaceae. *Conservation Genetics,* **10**: 643-648.

Slot JC, Rokas A. 2011. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Current Biology,* **21**: 134-139.

Small RL, Cronn RC, Wendel JF. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany,* **17**: 145-170.

Sneath PHA. 1957. The application of computers to taxonomy. *Journal of General Microbiology,* **17**: 201-226.

Sokal RR, Sneath PHA. 1963. *Principles of numerical taxonomy,* San Fransisco, W. H. Freeman and Company.

Soltis DE, Buggs RJA, Doyle JJ, Soltis PS. 2010. What we still don't know about polyploidy. *Taxon,* **59**: 1387-1403.

Soltis DE, Burleigh JG. 2009. Surviving the K-T mass extinction: new perspectives of polyploidization in angiosperms. *Proceedings of the National Academy of Sciences of the United States of America,* **106**: 5455-5456.

Soltis DE, Soltis PS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences,* **12**: 243-273.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution,* **14**: 348-352.

Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon,* **56**: 13-30.

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

Stebbins GL. 1947. Types of polyploidy: their classification and significance. *Advances in Genetics,* **1**: 403-429.

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

Stebbins GL. 1970. Variation and evolution in plants. *Essays in evolution and genetics in honor of Theodosius Dobzhansky.* New York, Appleton-Century-Croft.

Stebbins GL. 1971. *Chromosomal evolution in higher plants,* London, Edward Arnold (Publishers) Ltd.

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

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

Stuessy TF. 2009. Paradigms in biological classification (1707-2007): Has anything really changed? *Taxon,* **58**: 68-76.

Stuessy TF, König C. 2008. Patrocladistic classification. *Taxon,* **57**: 594-601.

Suda J. 2002. New DNA ploidy level in *Empetrum* (Empetraceae) revealed by flow cytometry. *Annales Botanici Fennici,* **39**: 133-141.

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

Tate JA, Joshi P, Soltis KA, Soltis PS, Soltis DE. 2009. On the road to diploidization? Homoeolog loss in independently formed populations of the allopolyploid *Tragopogon miscellus* (Asteraceae). *BMC Plant Biology,* **9**: 80.

Tate JA, Soltis DE, Soltis PS. 2005. Polyploidy in plants. In: Gregory TR ed. *The evolution of the genome.* Burlington, Elsevier Academic Press. 371-426.

Turcotte EL, Feaster CV. 1963. Haploids: high-frequency production from single-embryo seeds in a line of Pima cotton. *Science,* **140**: 1407-1408.

Van Wyk AE. 2007. The end justifies the means. *Taxon,* **56**: 645-648.

Vision TJ, Brown DG, Tanksley SD. 2000. The origins of genomic duplications in *Arabidopsis*. *Science,* **209**: 2114-2117.

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. Aflp - a new technique for DNA-fingerprinting. *Nucleic Acids Research,* **23**: 4407-4414.

Wang J, Tian L, Madlung A, Lee H-S, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics,* **167**: 1961-1973.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand D, Sninsky JJ, White TJ eds. *PCR protocols: a guide to methods and applications.* San Diego, Academic Press. 315–322.

Williams DM, Ebach MC, Wheeler QD. 2005. 150 reasons for paraphyly: a respons. *Taxon,* **54**: 858.

Wolfe KH. 2001. Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics,* **2**: 333-341.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America,* **106**: 13875-13879.

Zander RH. 2007. Paraphyly and the species concept, a reply to Ebach & al. *Taxon,* **56**: 642-644.

Zander RH. 2008. Evolutionary inferences from non-monophyly on molecular trees. *Taxon,* **57**: 1182-1188.

I

Systematic Botany (2006), 31(4): pp. 702–729

Taxonomy and Evolutionary Relationships in the *Saxifraga rivularis* Complex

MARTE HOLTEN JØRGENSEN,[1,2,4] REIDAR ELVEN,[1] ANDREAS TRIBSCH,[1,3] TOVE MARGETHE GABRIELSEN,[1,2] BRITA STEDJE,[1] and CHRISTIAN BROCHMANN[1]

[1]National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, NO-0318 Oslo, Norway;
[2]Department of Biology, University of Oslo, P.O. Box 1066 Blindern, NO-0316 Oslo, Norway;
[3]Department of Organismic Biology/Ecology and Diversity of Plants, University of Salzburg, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria
[4]Author for correspondence (martej@ulrik.uio.no)

Communicating Editor: James F. Smith

ABSTRACT. In many arctic-alpine plant groups, reticulate evolutionary histories have resulted in problems with species delimitation and phylogenetic reconstruction. In the *Saxifraga rivularis* complex ($2n = 26, 52$), the number of species accepted ranges from a single polymorphic one (*S. rivularis* s.l.) to several (the circumpolar *S. hyperborea*, the amphi-Atlantic *S. rivularis*, the three amphi-Pacific species *S. bracteata*, *S. flexuosa*, and *S. arctolitoralis*, and *S. debilis* in the Rocky Mountains). A combination of molecular (AFLPs), flow cytometrical, and morphological data from samples covering most of the distribution range was used to delimit taxonomic species and to unravel their evolutionary relationships. Four lineages with distinct morphological differences were recognized, representing four species: the diploids *S. bracteata*, *S. hyperborea* (including *S. flexuosa*), and *S. debilis*, and the tetraploid *S. rivularis* (including *S. arctolitoralis*). Based on a synthesis of the available data we provide a taxonomic revision of the complex and propose one rank change (***S. rivularis* subsp. *arctolitoralis*** comb. et stat. nov.). Genome sizes as well as the intermediate position of the *S. rivularis* lineage in the molecular and morphological analyses suggest a single allopolyploid origin from the *S. bracteata* and the *S. hyperborea* lineages, most likely in Beringia.

KEYWORDS: AFLPs, allopolyploidy, biogeography, cytology, morphology, reticulate evolution.

As one of the most polyploid-rich areas, the Arctic is suitable for studying the evolutionary significance of polyploidization (Brochmann et al. 2004). The majority of arctic plants appear to be of hybrid origin, stabilized by allopolyploidy (Stebbins 1985). Reconstructing the evolutionary history of arctic species complexes with several ploidal levels and clarifying their taxonomy are important steps towards a better understanding of the general processes of reticulation and polyploidization.

Saxifraga sect. *Mesogyne* Sternb., which includes the circumpolar *S. rivularis* complex, is almost exclusively arctic-alpine with 10–12 taxa at the species level (Webb and Gornall 1989; Zhmylev 1997; Jintang et al. 2001). Phylogenetic analyses based on chloroplast *matK* and *rbcL*, and nuclear ITS sequences suggest that the section is monophyletic (Soltis et al. 1996; Conti et al. 1999). Preliminary phylogenetic analyses including all species of section *Mesogyne* based on non-coding cpDNA sequences suggest that what is defined here as the *S. rivularis* complex represents a distinct lineage within the section (T. M. Gabrielsen and C. Brochmann, unpubl. data).

The *Saxifraga rivularis* complex consists of small perennial herbs with palmate leaves, rounded leaf lobes, usually cymose inflorescences, small flowers with white to pink petals, and ballistic seed dispersal. The plants grow in moist habitats such as snowbeds, scree slopes, or along rivers and creeks in the Arctic and in northern Pacific and Atlantic alpine regions (Fig. 1; Hultén 1968; Yurtsev 1981). Two main chromosome numbers have been recorded: $2n = 26$ and $2n = 52$ (Table 1; e.g., Zhukova and Tikhonova 1971; Engelskjøn 1979; Löve 1982; Zhukova and Petrovsky 1987), traditionally (and in this paper) designated as diploid and tetraploid. Fixed heterozygosity at an isozyme locus indicates that these chromosome numbers may represent secondary tetraploids and octoploids (Guldahl et al. 2005). Deviating chromosome numbers recorded in some studies (e.g., $2n = 56$, Böcher 1938; $2n = 48$, Zhukova et al. 1973; $2n = 23$, 38, 43, 47, 50, 85, 95, Guldahl et al. 2005), could represent occasional endopolyploidy or aneuploidy (Guldahl et al. 2005).

The Panarctic Flora Project (Elven et al. 2003) has until now tentatively accepted five species in the *Saxifraga rivularis* complex for the Arctic: the amphi-Atlantic tetraploid *S. rivularis* L., the circumpolar diploid *S. hyperborea* R.Br., and three amphi-Pacific taxa: the diploids *S. bracteata* D.Don and *S. flexuosa* Sternb., and the tetraploid *S. arctolitoralis* Jurtz. & V.V.Petrovsky (Elven et al. 2003). Plant size, pigmentation, growth form, presence or absence of rhizomes, number of leaf lobes, shape and number of bracts, length of

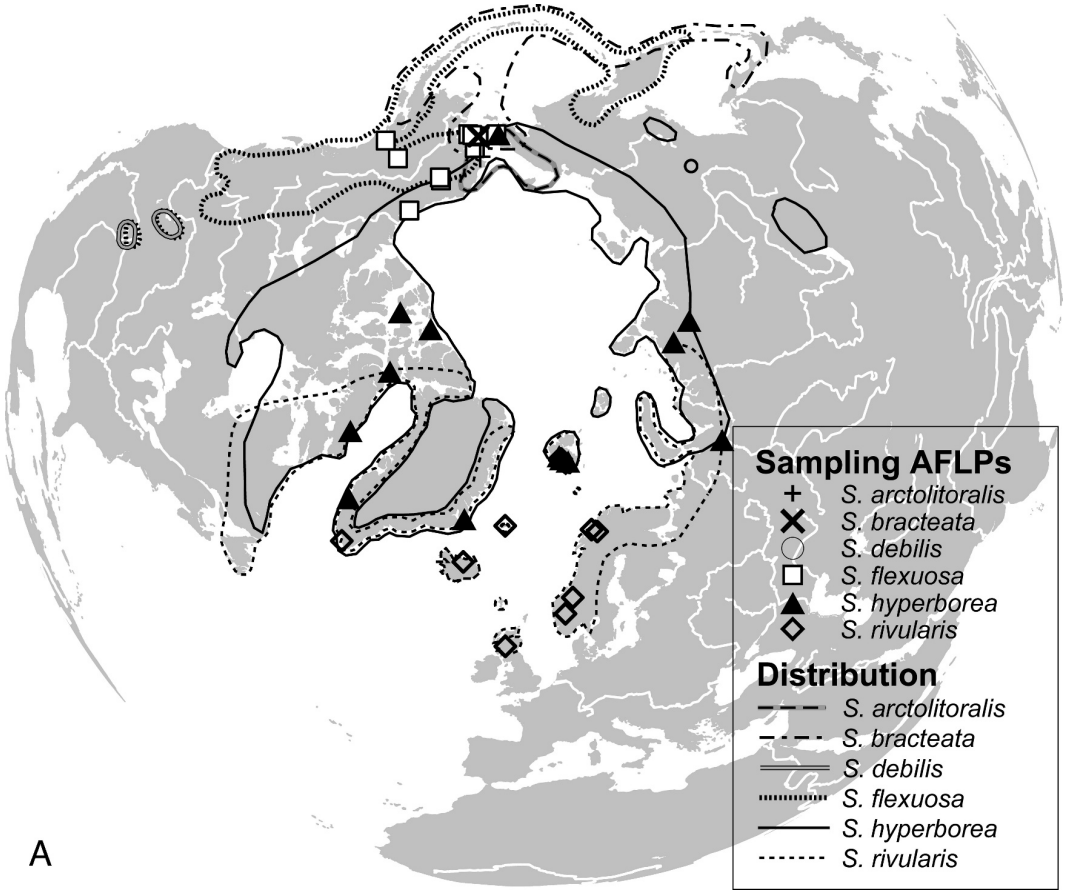


FIG. 1. Sampling and hypothesized distribution of the *Saxifraga rivularis* complex. The distribution areas are modified after Hultén and Fries (1986), who included *S. hyperborea* in *S. rivularis*. The distribution of *S. arctolitoralis* is added according to Rebristaya and Yurtsev (1984) and a revision of herbarium material in the University of Alaska Museum, Fairbanks (ALA) by Reidar Elven (in 2003). The distribution of *S. debilis* is after Weber and Wittmann (2001). The symbols indicate a priori determinations of the plants. A. Silica-dried material used for AFLP analysis. B. Herbarium material used for morphometric analysis, the Rocky Mountain area is inserted.

pedicel, shape of hypanthium, and several trichome characters have been regarded as differential morphological characters (e.g., Weber 1966; Yurtsev 1981; Rebristaya and Yurtsev 1984; Webb and Gornall 1989; Cronquist et al. 1997; Elven in Lid and Lid 2005). Linnaeus (1753) described *S. rivularis* from Lapland, Sweden. The circumpolar arctic *S. hyperborea* was described from Melville Island, Canada, by Brown (1823). *Saxifraga bracteata* was described from the North Pacific region by Don (1822) and is reported from coastal areas between 50° and 70° N on both sides of the Pacific (Fig. 1; Hultén and Fries 1986). The Cordilleran and amphi-Pacific *S. flexuosa* (Fig 1; Hultén and Fries 1986) was described from Lavrentiy Bay, Chukotka Peninsula, by Sternberg (1831). *Saxifraga arctolitoralis* was described from the north-eastern part of the Chukotka Peninsula as a local endemic (Yurtsev 1981; Rebristaya and Yurtsev 1984), but is now reported to be amphi-Beringian (Fig 1; Elven et al. 2003). Outside the Arctic in the Rocky Mountains, an additional taxon, *S. debilis* Engelm., was described by Engelmann in Gray (1863), and has been reported from Colorado, Utah, Wyoming and Montana (Fig. 1; Weber and Wittmann 2001).

The species delimitation is, however, controversial. The high number of synonyms in the *Saxifraga rivularis* complex clearly reflects poor understanding of its taxonomy and evolutionary relationships, and that a revision is needed. Several authors have included different taxa at the subspecific level in *S. rivularis*: *S. bracteata* (Engler 1872), *S. debilis* (Dorn 1988), *S. hyperborea* (Hooker 1834; Lange 1880; Dorn 1988), and *S. flexuosa* (Brown 1819; Engler in Rosendahl 1905; Engler and Irmscher 1916; Gjærevoll 1963). *Saxifraga flexuosa* and *S. hyperborea* are

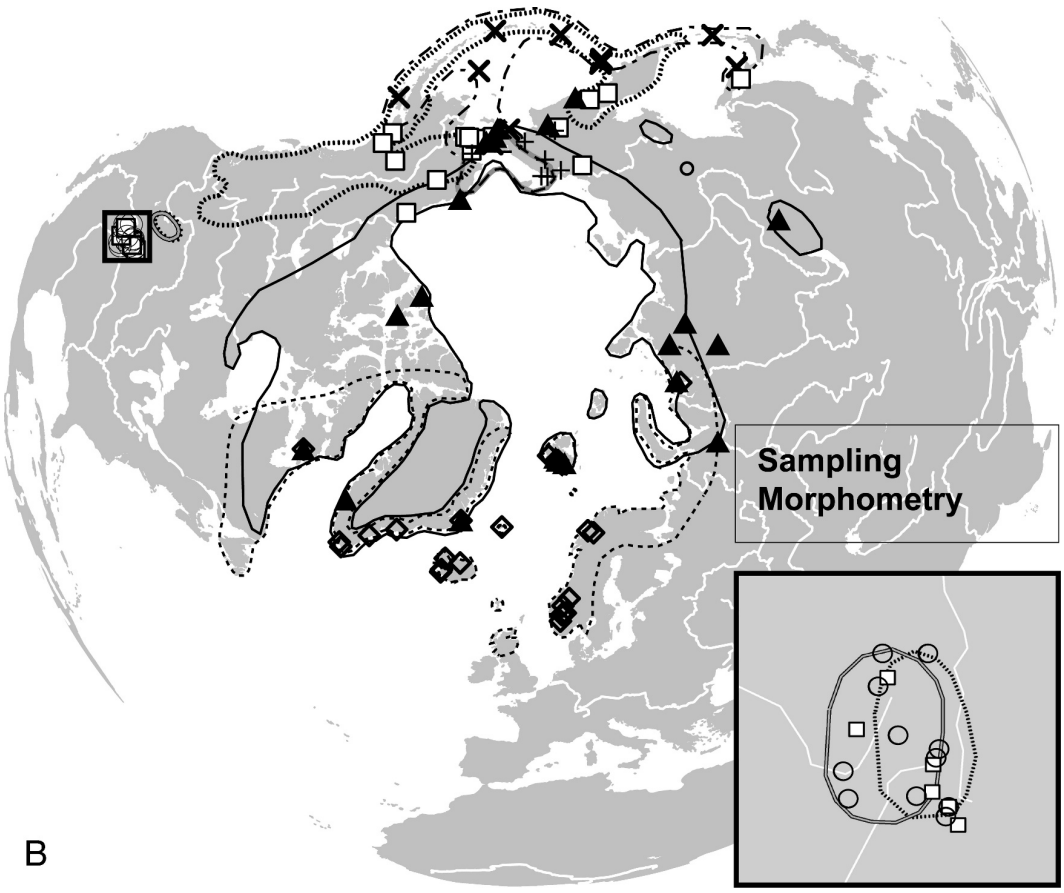


FIG. 1. Continued.

also considered synonymous or conspecific with *S. debilis* by several authors (e.g., Löve et al. 1971; Cronquist et al. 1997; Weber and Wittmann 2001), whereas other authors consider *S. debilis* to be a separate species (Harrington 1954) or more closely related to *S. cernua* L. (Engler 1872).

Several hypotheses on evolutionary relationships have been proposed in the *Saxifraga rivularis* complex. Rebristaya and Yurtsev (1984) suggested that *S. arctolitoralis* is an autopolyploid descendant of *S. hyperborea*, and that *S. rivularis* has an ancestor in common with *S. hyperborea* and *S. bracteata*. Furthermore, they suggested that *S. flexuosa* is conspecific with *S. hyperborea* and only represents plants growing in shaded creek margins. Based on *matK* sequences and ploidal levels, Brochmann et al. (1998) suggested that *S. rivularis* is of hybrid origin with *S. hyperborea* representing the maternal lineage. Guldahl et al. (2005) studied populations of *S. rivularis* and *S. hyperborea* in Svalbard using isozyme and RAPD analyses, flow cytometry, and morphometry. Their data suggest that there are two genetically well-separated entities with different ploidal levels in Svalbard, although no reliable differentiating morphological characters were identified.

Here we provide a taxonomic revision of the *Saxifraga rivularis* complex utilizing Elven et al. (2003) and Weber and Wittmann (2001) as initial frameworks. We also aim to identify evolutionary lineages in the complex based on molecular AFLP analysis, flow cytometry, and morphometry. Furthermore we investigate their phylogenetic and evolutionary relationships; i.e., we test the previous hypotheses suggested by Rebristaya and Yurtsev (1984) and Brochmann et al. (1998).

MATERIAL AND METHODS

Material. Fresh material was collected from most of the distribution area of the *Saxifraga rivularis* complex (except for *S. debilis*; Table 1; Fig. 1) and was either dried in silica gel, cultivated in a phytotron at the University of Oslo, or pressed as vouchers and deposited at the Natural History Museum at the University of Oslo (O). For the AFLP analysis, silica-dried material of usually three plants from each of 41 populations

TABLE 1. Sampled material of the *Saxifraga rivularis* complex. A priori taxonomic determinations of the specimens following Elven et al. (2003) and Weber and Wittmann (2001) are abbreviated as; arc = *S. arctolitoralis* Jurtz. & V.V.Petrovsky, bra = *S. bracteata* D.Don, deb = *S. debilis* Engelm., fle = *S. flexuosa* Sternb., hyp = *S. hyperborea* R.Br. and riv = *S. rivularis* L. A posteriori determinations of the same specimens to taxa are abbreviated as; riv ssp. arc = *S. rivularis* subsp. *arctolitoralis* (Jurtz. & V.V.Petrovsky) Jørgensen & Elven, riv ssp. riv = *S. rivularis* subsp. *rivularis*, bra = *S. bracteata*, deb = *S. debilis*, and hyp = *S. hyperborea*. Pop ID gives the original collection code or herbarium registration number for the populations. Country/area abbreviations: CAN = Canada, GBR = Great Britain, GRL = Greenland, ISL = Iceland, NOR = Norway, RUS = Russia, SJM = Svalbard/Jan Mayen, USA = United States of America. More detailed information about localities, collection dates etc., is available from the author for correspondence. H gives herbarium of deposition of voucher specimens, acronyms according to Index Herbariorum (Holmgren and Holmgren 1998 onwards). #H, #S and #C give the number of herbarium, silica-dried and cultivated plants included in this study, respectively. Previous chromosome counts and estimated ploidal levels from flow cytometry are from (1) Zhukova and Petrovsky (1987), (2) Zhukova and Tikhonova (1971), (3) Zhukova et al. (1973), (4) Zhukova and Petrovsky (1980), and (5) Guldahl et al. (2005). [a] No official codes were registered for these populations.

A priori det.	A posteriori det.	Pop ID	Country	Region	Collector(s)	H	#H	#S	#C	Chromosome number and ploidal level
arc	riv ssp. arc	65-195	RUS	Magadan Oblast, Wrangel Island	Zhukova, Petrovsky	LE	3			4x = 2*n* = 52 (1)
arc	riv ssp. arc	68-09K	RUS	Magadan Oblast, W Chukotka	Levitsev, Korobkov, Yurtsev	LE	3			4x = 2*n* = 52 (1; 2)
arc	riv ssp. arc	69-45K	RUS	Magadan Oblast, E Chukotka	Korobkov	LE	1			4x = 2*n* = 52 (1)
arc	riv ssp. arc	70-160y	RUS	Magadan Oblast, E Chukotka	Kozhevnikov, Nechaev, Yurtsev	LE	3			4x = 2*n* = 52 (1)
arc	riv ssp. arc	71-17-1	RUS	Magadan Oblast, Wrangel Island	Petrovsky, Taraskina, Schteinberg	LE	3			4x = 2*n* = 48 (1; 3)
arc	riv ssp. arc	72-162	RUS	Magadan Oblast, E Chukotka	Zhukova	LE	2			4x = 2*n* = 52 (1)
arc	riv ssp. arc	MJ03.5[a]	RUS	Magadan Oblast, N Chukotka	Sergienko	LE	3			
arc	riv ssp. arc	MJ03.6[a]	RUS	Magadan Oblast, N Chukotka	Sergienko	LE	3			
arc	riv ssp. arc	TG03/46[a]	RUS	Magadan Oblast, Zaliv Kresta	Dorogostayskaya	LE	3			
arc	riv ssp. arc	W79-84	RUS	Magadan Oblast, Wrangel Island	Petrovsky	LE	3			4x = 2*n* = 52 (1)
arc	riv ssp. arc	RE-01-11	USA	Alaska, Seward Peninsula	Elven, Solstad	O	3	3		
bra	bra	MJ03.3[a]	RUS	Sakhalin Oblast, Kuril'skiye Ostrova	Bergman	S	3			
bra	bra	3306	RUS	Sakhalin Oblast, Sakhalin	Egorova, Kolehanova	LE	1			
bra	bra	70-58Pl	RUS	Magadan Oblast, E Chukotka	Plieva	LE	3			2x = 2*n* = 26 (1)
bra	bra	72	RUS	Kamchatskaya Oblast, Komandorskiye Ostrova	Kardakova	LE	3			
bra	bra	804	RUS	Kamchatskaya Oblast, Komandorskiye Ostrova	Hultén	LE	3			
bra	bra	MJ03.7[a]	RUS	Magadan Oblast, E Chukotka	Derviz-Sokolova	LE	3			
bra	bra	MJ03.8[a]	RUS	Magadan Oblast, Beringov Proliv	Tikhomirov	LE	3			
bra	bra	3189	RUS	Kamchatskaya Oblast, Komandorskiye Ostrova	Svensson	S	3			
bra	bra	1069	USA	Alaska, Aleutian islands	Eyerdam	S	3			
bra	bra	8121	USA	Alaska, Aleutian islands	Shacklette	S	3			
bra	bra	574a	USA	Alaska, Kodiak Island	Eyerdam	S	3			
bra	bra	1826	USA	Alaska, Pribilof Islands	Coville, Kearney	S	2			
bra	bra	SUP02-188	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen	O	3	2	4	
bra	bra	SUP02-199	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen			2	1	
deb	deb	00740761	USA	Colorado, San Juan County	Douglass (Douglass 54-420, cf. Weber 1966)	COLO	3			
deb	deb	00740225	USA	Colorado, Boulder County	Weber (Weber 3437, cf. Weber 1966)	COLO	3			
deb	deb	00740720	USA	Colorado, Eagle County	Hogan	COLO	3			

TABLE 1. Continued.

A priori det.	A posteriori det.	Pop ID	Country	Region	Collector(s)	H	#H	#S	#C	Chromosome number and ploidal level
deb	deb	00740332	USA	Colorado, San Miguel County	Siplivinsky	COLO	3			
deb	deb	00740431	USA	Colorado, Garfield County	Douglass	COLO	3			
deb	deb	00740464	USA	Colorado, San Juan County	Hartman	COLO	3			
deb	deb	00740217	USA	Colorado, Clear Creek County	Lederer	COLO	3			
deb	deb	00740373	USA	Colorado, Gunnison County	Weber, Langenheim	COLO	3			
deb	deb	00740167	USA	Colorado, Park County	Hogan, Yeatts	COLO	3			
deb	deb	00740142	USA	Colorado, Alamosa County	Hogan	COLO	3			
deb	deb	00740480	USA	Colorado, Dolores County	Moore	COLO	3			
deb	deb	00740209	USA	Colorado, Park County	Murphy	COLO	3			
fle	hyp	RE2260	CAN	Yukon Territory, Ivvavik National Park	Elven	O	2		1	2x (5)
fle	hyp	RE2263	CAN	Yukon Territory, Ivvavik National Park	Elven			2		
fle	hyp	77-03Pl	RUS	Magadan Oblast, W Chukotka	Petrovsky, Koroleva	LE	3			2x = 2*n* = 26 (1; 4)
fle	hyp	TG03/39[a]	RUS	Sakhalin Oblast, East Sakhalin Mts.	Vyshin, Barkalov	LE	3			
fle	hyp	TG03/40[a]	RUS	Kamchatskaya Oblast, Koryakskiy National District	Gavriljuk	LE	3			
fle	hyp	TG03/42[a]	RUS	Kamchatskaya Oblast, Karaginskiy County	Dorotejev	LE	3			
fle	hyp	TG03/45[a]	RUS	Magadan Oblast, SE Chukotka	Gavriljuk, Talarich	LE	3			
fle	hyp	04-01	RUS	Magadan Oblast, Chutkotka Peninsula	Razzhivin	O	3	3		
fle	hyp	RE-01-10	USA	Alaska, Seward Peninsula	Elven, Solstad	O	3	3		
fle	hyp	RE-C1	USA	Alaska, Kenai Peninsula	Elven, Solstad	O	3			
fle	hyp	SR1	USA	Alaska, Alaska Range NE	Elven, Grundt	O	3	3		
fle	hyp	SUP02-150	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen				2	
fle	hyp	SUP02-154	USA	Alaska, Valdez Area	Elven, Gabrielsen, Jørgensen	O	1	3		
fle	hyp	SUP02-155	USA	Alaska, Valdez Area	Elven, Gabrielsen, Jørgensen	O	3		1	
fle	hyp	SUP02-163	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen	O	3	3		
fle	hyp	SUP02-168	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen				1	
fle	hyp	SUP02-180	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen	O	3	3	2	
fle	hyp	SUP02-181	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen	O	3	3	2	
fle	hyp	SUP02-185	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen				2	
fle	hyp	SUP02-211	USA	Alaska, Seward Peninsula	Elven, Gabrielsen, Jørgensen				2	
fle	hyp	SUP02-254	USA	Alaska, Arctic Slope	Elven			3		
fle	hyp	SUP02-265	USA	Alaska, Brooks Range	Elven	O	3	2		
fle	hyp	00740514	USA	Colorado, Clear Creek County	Lederer	COLO	3			
fle	hyp	00740548	USA	Colorado, Larimer County	Willard	COLO	3			
fle	hyp	00740597	USA	Colorado, Park County	Komarkova	COLO	3			
fle	hyp	00740613	USA	Colorado, Clear Creek County	Weber, Porsild, Holmen	COLO	3			
fle	hyp	00740795	USA	Colorado, Summit County	Weber	COLO	3			
fle	hyp	00740712	USA	Colorado, La Plata County	Michener	COLO	3			
fle	hyp	00740670	USA	Colorado, San Juan County	Michener	COLO	3			
hyp	hyp	04-93a	CAN	Nunavut, Baffin Island	Aiken, LeBlanc	O	3			

TABLE 1. Continued.

A priori det.	A posteriori det.	Pop ID	Country	Region	Collector(s)	H	#H	#S	#C	Chromosome number and ploidal level
hyp	hyp	G-6842	CAN	Northwest Territory, Queen Elizabeth Islands	Gillespie, Consaul	O	3			
hyp	hyp	RE2775	CAN	Northwest Territory, Queen Elizabeth Islands	Elven				1	2x (5)
hyp	hyp	RE2957	CAN	Nunavut, Parry Islands	Elven	O	3			
hyp	hyp	RE2981	CAN	Nunavut, Parry Islands	Elven			2		2x (5)
hyp	hyp	RE3033	CAN	Nunavut, Ellesmere Island	Elven			3		2x (5)
hyp	hyp	RE3425	CAN	Nunavut, Devon Island	Elven			3	2	2x (5)
hyp	hyp	RE3520	CAN	Nunavut, Baffin Island	Elven			3	3	2x (5)
hyp	hyp	RE3525	CAN	Nunavut, Baffin Island	Elven			1	1	2x (5)
hyp	hyp	AK-240	GRL	Kitaa, Kong Frederik IX Land	Bronken	O	1	3	1	
hyp	hyp	CB-99-027	GRL	Tunu, Jameson Land	Brochmann	O	3	3		2x (5)
hyp	hyp	CB-99-038	GRL	Tunu, Jameson Land	Brochmann, Nylehn, Bronken, Kjølner	O	3			2x (5)
hyp	hyp	05-01	RUS	Magadan Oblast, Chutkotka Peninsula	Razzhivin	O	3	3		
hyp	hyp	05-03	RUS	Magadan Oblast, Chutkotka Peninsula	Razzhivin	O	3			
hyp	hyp	70-145	RUS	Magadan Oblast, Wrangel Island	Petrovsky, Zhukova	LE	3			2x = 2*n* = 26 (1)
hyp	hyp	71-06	RUS	Magadan Oblast, Wrangel Island	Petrovsky, Zhukova	LE	3			2x = 2*n* = 26 (1;3)
hyp	hyp	77-32	RUS	Magadan Oblast, Anadyrsky Raion	Petrovsky, Zhukova	LE	3			2x = 2*n* = 26 (1)
hyp	hyp	SUP-3523	RUS	Taimyrsky Autonomous Okrug, North-Siberian Lowland	Schönswetter, Tribsch	O	3	3	6	
hyp	hyp	SUP-4495	RUS	Yamalo-Nenetsky Autonomous Okrug, Polar Ural	Alsos, Kapralov	O	1	3	1	
hyp	hyp	SUP04-35	RUS	Taimyrsky Autonomous Okrug, Taymyr Peninsula	Pospelova	O	1	3		
hyp	hyp	TG03/28[a]	RUS	Krasnoyarskiy Kray, Taymyr Peninsula	Dorogostaiskaya	LE	3			
hyp	hyp	TG03/31[a]	RUS	Krasnoyarskiy Kray, Putoran Mts.	Kozhevnikov	LE	3			
hyp	hyp	TG03/34[a]	RUS	Buryato-Mongolskaya, Stanovoye Mts.	Malychev, Petrochenko	LE	3			
hyp	hyp	TG03/41[a]	RUS	Kamchatskaya Oblast, Korjaskij National District	Gavriljuk	LE	3			
hyp	hyp	TG03/43[a]	RUS	Magadan Oblast, Chukotka Peninsula	Tikhomirov	LE	2			
hyp	hyp	TG03/44[a]	RUS	Magadan Oblast, Beringov Proliv	Tikhomirov	LE	3			
hyp	hyp	TMG-24SH	SJM	Svalbard, Nordenskiöld Land	Brochmann, Gabrielsen, Steen	O	3	4		
hyp	hyp	TMG-26	SJM	Svalbard, Dickson Land	Elven	O	3	1		
hyp	hyp	TMG-27	SJM	Svalbard, Dickson Land	Elven	O	3	3		
hyp	hyp	TMG-34	SJM	Svalbard, Nordenskiöld Land	Gabrielsen	O	3	2		
hyp	hyp	MJ03.2[a]	USA	Alaska, Point Barrow	Thomas	O	3			
riv	riv ssp. riv	SUP04-36	CAN	Nunavut, Baffin Island	Aiken				2	
riv	riv ssp. riv	04-93	CAN	Nunavut, Baffin Island	Aiken, LeBlanc	O	1			
riv	riv ssp. riv	TMG-43	GBR	Scotland, Argyll County	Gabrielsen, Steen, Hansen, Tebbitt, Watson			3		
riv	riv ssp. riv	3189	GRL	Tunu, Kong Frederik VI Kyst	Kliim-Nielsen	O	3			

TABLE 1. Continued.

A priori det.	A posteriori det.	Pop ID	Country	Region	Collector(s)	H	#H	#S	#C	Chromosome number and ploidal level
riv	riv ssp. riv	MJ 03.1[a]	GRL	Tunu, Kong Frederik VI Kyst	Devold, Scholander	O	3			
riv	riv ssp. riv	MJ 03.4[a]	GRL	Tunu, Kong Frederik VI Kyst	Devold, Scholander	O	3			
riv	riv ssp. riv	AK-248	GRL	Kitaa, Kong Frederik IX Land	Bronken, Jacobsen				2	
riv	riv ssp. riv	AK-281	GRL	Tunu, Kong Frederik VI Kyst	Bronken	O	1	3		
riv	riv ssp. riv	CB-99-039	GRL	Tunu, Jameson Land	Brochmann, Nylehn, Bronken, Kjølner	O	3			4x (5)
riv	riv ssp. riv	CB-99-043	GRL	Tunu, Kong Christian IX Land	Alsos, Lund				2	4x (5)
riv	riv ssp. riv	CB-99-045	GRL	Tunu, Kong Christian IX Land	Alsos, Lund	O	3		1	4x (5)
riv	riv ssp. riv	CB-99-047	GRL	Tunu, Kong Christian IX Land	Alsos, Lund	O	3		1	4x (5)
riv	riv ssp. riv	AK-811	ISL	Vestfirðir, Hólmavík	Skrede, Kjølner, Lund	O	1		2	
riv	riv ssp. riv	AK-846	ISL	Vesturland, NE of Akranes	Skrede, Kjølner, Lund	O	1		2	
riv	riv ssp. riv	AK-867	ISL	Norðurland Eystra, Akureyri	Skrede, Kjølner, Lund	O	1	3		
riv	riv ssp. riv	CB-99-049	ISL	Vesturland, NE of Akranes	Alsos, Lund	O	3		3	4x (5)
riv	riv ssp. riv	AK-416a	NOR	Buskerud, Ål	Jørgensen	O	1			
riv	riv ssp. riv	AK-416b	NOR	Buskerud, Ål	Jørgensen				3	
riv	riv ssp. riv	AK-441	NOR	Oppland, Lom	Bronken	O	1		2	
riv	riv ssp. riv	AK-456	NOR	Hordaland, Odda	Bronken	O	1		2	
riv	riv ssp. riv	AK-470	NOR	Sør-Trøndelag, Oppdal	Bronken	O	1	3	1	
riv	riv ssp. riv	AK-516	NOR	Hordaland, Finse	Jørgensen, Skrede, Jacobsen	O	2	3	1	
riv	riv ssp. riv	AK-745	NOR	Troms, Tromsø	Alsos, Westergård	O	2	3		
riv	riv ssp. riv	AK-784	NOR	Troms, Storfjord	Alsos, Westergård	O	1	2		
riv	riv ssp. riv	TG03/30	RUS	Krasnoyarskiy Kray, Taymyria	Matreeva, Zanokha	LE	3			
riv	riv ssp. riv	219876	SJM	Svalbard, Oscar II Land	Emanuelsson	O	3			
riv	riv ssp. riv	AK-109	SJM	Jan Mayen	Wollan	O	1	3	1	
riv	riv ssp. riv	AK-554	SJM	Svalbard, Nordenskiöld Land	Alsos, Westergård	O	1	3		
riv	riv ssp. riv	AK-559	SJM	Svalbard, Nordenskiöld Land	Alsos, Westergård			1		
riv	riv ssp. riv	AK-3135	SJM	Svalbard, Wedel Jarlsberg Land	Jørgensen, Skrede				2	
riv	riv ssp. riv	AK-3245	SJM	Svalbard, Wedel Jarlsberg Land	Jørgensen, Skrede				1	
riv	riv ssp. riv	TMG-21	SJM	Svalbard, Nordenskiöld Land	Brochmann, Gabrielsen, Steen	O	3	2	1	4x (5)
riv	riv ssp. riv	TMG-24	SJM	Svalbard, Nordenskiöld Land	Brochmann, Gabrielsen, Steen				2	4x (5)
riv	riv ssp. riv	TMG-34	SJM	Svalbard, Nordenskiöld Land	Gabrielsen				2	4x (5)
riv	riv ssp. riv	TMG-35	SJM	Svalbard, Nordenskiöld Land	Fjellheim, Scheen				1	4x (5)
riv	riv ssp. riv	TMG-37	SJM	Svalbard, Oscar II Land	Gabrielsen				6	4x = 2n = 43, 52, 85 (5)
riv	riv ssp. riv	TMG-37B	SJM	Svalbard, Oscar II Land	Gabrielsen				1	4x = $2n$ = 52(5)
riv	riv ssp. riv	TMG-39	SJM	Svalbard, Oskar II Land	Gabrielsen				1	4x = 2n = 26, 47, 52, 95 (5)
riv	riv ssp. riv	TMG-61	SJM	Svalbard, Nordenskiöld Land	Scheen, Fjellheim				1	

was included (Table 1; Fig. 1A). Fresh material of 76 cultivated plants from 42 populations was used for flow cytometry (Table 1). 282 herbarium specimens from 108 populations deposited in the University of Colorado Herbarium, University of Colorado, Boulder (COLO), the V. L. Komarov Botanical Institute, Russian Academy of Sciences, St. Petersburg (LE), the Natural History Museum, University of Oslo, Oslo (O), and the Swedish Museum of Natural History, Stockholm (S), were used for the morphometric analyses (Table 1; Fig. 1B). Whenever possible, the same populations were chosen for all three analyses. We were, however, not able to obtain living material of *S. arctolitoralis* and *S. debilis* for flow cytometry, and *S. debilis* had to be excluded from the AFLP analysis because only herbarium material was available for this species.

DNA Isolation and AFLP Analysis. DNA isolation was performed using the Qiagen DNeasy Kit (Qiagen, Hilden) as recommended by the manufacturer, with some modifications: Approximately 1 cm^2 of silica dried leaves was crushed in 2 ml tubes with tungsten carbide beads for 2 × 1 min in a mixer mill (MM301, Retsch GmbH & Co., Haan) at 20 Hz. 400 µl AP1 buffer was added, and the samples were stored overnight at −80°C. After thawing in a 65°C heat block, 3.5 µl RNAse-mix was added, and the mixtures incubated for 20 min at 65°C. The isolated DNA was eluted twice in 50 µl AE buffer and stored at −20°C.

Amplified fragment length polymorphisms (AFLPs; Vos et al. 1995) were obtained using GeneAmp PCR system 9700 (Applied Biosystems, Foster City) at the Natural History Museum, University of Oslo. To 5.5 µl of each DNA extraction we added 0.125 µl *Eco*RI (5 U; Roche, Basel) and 0.020 µl *Mse*I (1 U; New England BioLabs, Beverly) restriction enzymes, 1.0 µl 10 µM *Eco*RI and 1.0 µl 10 µM *Mse*I adapters (MWG, Ebersberg; for the adaptor and primer sequences, see Vos et al. 1995), and 0.200 µl T4 DNA ligase (1 U; Roche, Basel) together with a reaction buffer containing 1.1 µl 10× T4 DNA buffer (with ATP; Roche, Basel), 1.1 µl 0.5 M NaCl, 0.55 µl BSA (1 mg/ml; Roche, Basel), and 0.405 µl autoclaved dH_2O, giving a reaction volume of 11 µl. The mixture was incubated for 3 h at 37°C and diluted 10 times with autoclaved dH_2O.

To 1.5 µl of the diluted restriction ligation product was added: 1.25 µl AmpliTaq buffer (Applied Biosystems, Foster City), 0.75 µl 25 mM $MgCl_2$ (Applied Biosystems, Foster City), 1.0 µl 10 mM dNTP (Applied Biosystems, Foster City), 0.05 µl AmpliTaq (Applied Biosystems, Foster City), 7.45 µl autoclaved dH_2O, and 0.25 µl 10 µM of each of *Eco*RI-A and *Mse*I-C pre-selective primers (MWG, Ebersberg), giving a total reaction volume of 12.5 µl. The fragments were amplified under the following PCR conditions: 2 min at 72°C, 30 cycles of 30 sec at 94°C, 30 sec at 56°C, and 2.5 min at 72°C, and one last hold of 10 min at 72°C. The PCR-products were diluted 10 times with autoclaved dH_2O.

2.5 µl of the diluted preselective products were added to a mixture of 1.25 µl Taq Gold buffer (Applied Biosystems, Foster City), 1.25 µl 25 mM $MgCl_2$ (Applied Biosystems, Foster City), 1.0 µl 10 mM dNTP (Applied Biosystems, Foster City), 0.1 µl AmpliTaq Gold (Applied Biosystems, Foster City), 0.1 µl BSA (Roche, Basel), 5.95 µl autoclaved dH_2O, and 0.1 µl and 0.25 µl 10 µM of the selective primers *Eco*RI-A§§ (MWG, Ebersberg; Applied Biosystems, Foster City) and *Mse*I-C§§ (MWG, Ebersberg), respectively. The PCR profile consisted of 10 min at 95°C, 13 cycles of 30 sec at 94°C, 1 min at 65–56°C (the temperature decreasing 0.7°C after each cycle), and 2.5 min at 72°C, 23 cycles of 30 sec at 94°C, 1 min at 56°C, and 2.5 min at 72°C, and it finished off with a 10 min hold at 72°C.

1.0 µl of each selective PCR product was mixed in 11.8 µl HiDi (formamide) and 0.2 µl GeneScan Rox 500 size standard (Applied Biosystems, Foster City). The fluorescently labeled AFLP fragments were denatured for 5 min at 95°C and analyzed with an ABI 3100 Sequencer (Applied Biosystems, Foster City) with 10 sec injection time and 3 kV injection voltage, otherwise default conditions.

Eight pairs of AFLP selective primers were tested; *Eco*RI-AGA/*Mse*I-CAA, *Eco*RI-AGA/*Mse*I-CAG, *Eco*RI-AGA/*Mse*I-CTG, *Eco*RI-AGA/*Mse*I-CTT, *Eco*RI-ATC/*Mse*I-CAA, *Eco*RI-ATC/*Mse*I-CAG, *Eco*RI-ATC/*Mse*I-CTG and *Eco*RI-ATC/*Mse*I-CTT. Three of these (6-FAM labeled *Eco*RI-AGA/*Mse*I-CAA, 6-FAM labeled *Eco*RI-AGA/*Mse*I-CTG and VIC labeled *Eco*RI-ATC/*Mse*I-CAA) gave clear profiles with appropriate levels of polymorphism and were chosen for the full analysis.

The samples were visualized and sized using GENESCAN Analysis Software version 3.7 (Applied Biosystems, Foster City). Peaks in the range 50–500 base pairs were scored as present (1) or absent (0), the intensity threshold being approximately 5% of the most intensive peak, using GENOGRAPHER 1.6.0 (Benham 2001). Bands not reproduced in duplications of two replicated samples were excluded from further analysis. Most profiles had between 100 and 150 bands, and specimens with profiles outside this range were excluded from the study due to presumedly imperfect amplification. The final dataset consisted of 104 individuals and 250 markers, and the data matrix is available from treeBASE (study number S1564).

The AFLP variation was visualized using principal coordinate analysis (PCO) in NTSYSpc version 2.02 (Rohlf 1999) based on the similarity measure of Dice (1945). PCO analyses were also run for different subsets of the material. The AFLP data were analysed by both hierarchical and non-hierarchical clustering methods. Neighbor-joining analyses were performed in TREECON version 1.3b (Van de Peer and De Wachter 1994), with the Nei-Li distance measure (Nei and Li 1979). Bootstrap analysis was carried out with 1,000 replicates. A Bayesian approach using STRUCTURE version 2 (Pritchard et al. 2000) calculated a logarithmic probability for the data given a number of clusters and assigned the specimens to these clusters probabilistically. The method may be applied to dominant markers such as AFLPs under a no admixture model, assuming no linkage between the loci (Pritchard et al. 2000). Ten replicates of each value of K (= the number of groups) were run for different selections of samples with a burn-in period of 100,000 and 1,000,000 iterations. A parsimony analysis was performed using heuristic search with random trees as starting point, 1,000 replications, saving 10 trees from each replicate, and tree bisection-reconnection (TBR) as branch-swapping algorithm in PAUP* version 4.0b10 (Swofford 2001). Additional TBR-swapping was performed with one tree from each island as starting point, saving 1,000 trees in each analysis. A strict consensus tree was calculated from the resulting trees. Bootstrap support was estimated with heuristic search, 1,000 replicates, using four random addition sequence replications, saving one tree from each replicate, and TBR-swapping.

Flow Cytometry. Flow cytometry was performed by G. Geenen, Plant Cytometry Services (Schijndel, The Netherlands) using DAPI staining and *Lactuca sativa* L. as internal standard. A few cm^2 (20–50 mg) of fresh leaf material was chopped with a razor blade and a DNA buffer (at pH 8) containing 5.0 mM Hepes, 10.0 mM $MgSO_4 \times 7H_2O$, 50.0 mM KCl, 0.2% Triton X-100, 2% DTT (dithiothreitol), and 4 mg/l DAPI, was added. The solution was filtered through a nylon filter of 40 µm mesh size. The flow cytometer PAS II (Partec GmbH, Münster, Germany) was used with a high pressure mercury lamp (OSRAM HBO 100 Long Life; OSRAM GmbH, München), heat protection filter KG-1, excitation filters UG-1 and BG-38, dichroic mirrors TK

TABLE 2. Morphological characters examined in the *Saxifraga rivularis* complex. P/D indicates if the character is primary (P) or derived (D). Character types are given by Q - qualitative, D - quantitative discrete, and C - quantitative continuous. Characters included in multivariate analyses are given in column PCO. All measurements are in mm.

No	P/D	Type	Short name	Character	PCO	Note
1	P	Q	Rhizomes	Presence of rhizomes	X	0 - absent, 1 - present
2	P	C	Stem length	Length of flowering stem	X	
3	P	C	Pedicel length	Length of lowest flower pedicel		In late flowering and fruit stage only
4	D			Char. 3/char. 2	X	
5	P	D	Flowers	No of buds/flowers/fruits per flowering stem	X	
6	P	D	Cauline leaves	No of cauline leaves per flowering stem		Including bracts
7	D			Char. 6/char. 2	X	
8	P	Q	Pigments	Pigments	X	0 - green, 1 - anthocyanin-colored hypanthium and petals, 2 - anthocyanin-colored hypanthium, petals and upper part of stem, 3 - the whole plant anthocyanin-colored
9	P	Q	Density stem	Hair density at the lower 1/3 of stem	X	0 - glabrous, 1 - sparsely hairy, 2 - densely hairy
10	P	Q	Hypanthium shape	Shape of hypanthium, early stage	X	0 - bowl-shaped, 1 - intermediate, 2 - v-shaped
11	P	Q	Hypanthium shape	Shape of hypanthium, late stage	X	0 - bowl-shaped, 1 - intermediate, 2 - v-shaped
12	P	C	Hypanthium length	Length of hypanthium, early stage	X	
13	P	C	Hypanthium length	Length of hypanthium, late stage	X	
14	P	Q	Glandular hairs	Presence of hypanthium glandular hairs	X	0 - absent, 1 - present
15	P	Q	Density hypanthium	Density of hypanthium glandular hairs	X	1 - sparse, 2 - dense
16	P	C	Hair length	Length of hypanthium glandular hairs	X	
17	P	D	Hair cell no.	No. of cells, hypanthium glandular hairs	X	
18	P	Q	Hair color	Color of hair partition-walls, upper pedicel	X	0 - not colored, 1 - light purple, 2 - purple
19	P	C	Sepal length	Length of sepal		Maximum length
20	P	C	Sepal width	Width of sepal		Maximum width
21	D			Char. 19/char. 13	X	
22	D			Char. 20/char. 13	X	
23	D			Char. 20/char. 19	X	
24	P	Q	Sepal shape	Sepal apex shape	X	0 - acute, 1 - intermediate, 2 - obtuse
25	P	C	Petal length	Length of petal		Maximum length
26	P	C	Petal width	Width of petal	X	Maximum width
27	D			Char. 26/char. 25	X	
28	P	Q	Petal shape	Shape of petal	X	0 - oblong, 1 - intermediate, 2 - elliptic, broadly rounded
29	P	C	Gynoecium length	Length of gynoecium		Fruit stage only
30	P	Q	Dist_ubract	Distinction of uppermost bract	X	0 - not distinct, 1 - distinct
31	P	Q	Upper bract petiole	Distinction of uppermost bract petiole	X	0 - not distinct, 1 - distinct
32	P	C	Upper bract length	Length of uppermost bract lamina		Maximum length
33	P	C	Upper bract width	Width of uppermost bract lamina	X	Maximum width
34	D			Char. 33/char. 32	X	
35	P	D	Upper bract lobes	No of lobes, uppermost bract	X	
36	P	Q	Dist_lbract	Distinction of lowest bract		0 - not distinct, 1 - distinct
37	P	Q	Lowest bract petiole	Distinction of lowest bract petiole	X	0 - not distinct, 1 - distinct
38	P	C	Petiole length	Length of lowest bract petiole	X	
39	P	C	Lowest bract length	Length of lowest bract lamina	X	Maximum length
40	P	C	Lowest bract width	Width of lowest bract lamina	X	Maximum width
41	P	D	Lowest bract lobes	No of lobes, lowest bract	X	
42	P	Q	Lowest bract shape	Shape of middle lobe apex, lowest bract	X	0 - rounded, 1 - intermediate, 2 - edged, 3 - acute
43	P	C	Leaf length	Length of basal leaf lamina		Maximum length
44	P	C	Leaf width	Width of basal leaf lamina	X	Maximum width
45	D			Char. 44/char. 43	X	
46	P	C	Dist_inc	Distance from basal leaf incision to basal leaf lamina base	X	

TABLE 2. Continued.

No	P/D	Type	Short name	Character	PCO	Note
47	P	Q	Incision shape	Shape of basal leaf incision	X	0 - v-shaped, 1 - intermediate, 2 - u-shaped
48	P	D	Leaf lobes	No of lobes, basal leaf	X	
49	P	Q	Leaf shape	Shape of middle lobe apex, basal leaf	X	0 - rounded, 1 - intermediate, 2 - edged
50	P	Q	Sheath structure	Structure of basal leaf petiole sheath	X	0 - thin/pale, 1 - intermediate, 2 - firm/brown
51	P	Q	Sheath shape	Shape of basal leaf petiole sheath shoulder	X	0 - rounded, 1 - intermediate, 2 - edged, 3 - acute
52	P	Q	Sheath fringes	Density of basal leaf petiole sheath fringes	X	0 - absent, 1 - sparse, 2 - dense

420 and TK 560, and emission filter GG 435. DNA histograms were obtained using the FLOWS 2.00 Software Package (Partec GmbH, Münster).

Morphometry. Forty-four primary and eight derived morphological characters (Table 2) where chosen based on characters used in floras (e.g., Polunin 1959; Hultén 1968; Rebristaya and Yurtsev 1984; Elven in Lid and Lid 2005) and our own observations. Developmental stage of the plants was recorded as early flowering stage, late flowering stage, and fruit stage, to avoid comparison across stages of stage-dependent characters (flowering stem legth, hypanthium shape and length). Only one replicate was made for each plant due to their small size. Descriptive statistics for the six tentative taxa and all morphological characters were calculated using SPSS 12.0.1 (SPSS Inc., Chicago). A Mann-Whitney *U* test for equality of medians was performed with the same program for all morphological characters and all combinations of taxa to test whether the characters significantly seperate entities. To reduce the probability of type I errors increased by the large number of tests, q-values were calculated from the probabilities using the robust method in the QVALUE software (e.g., Storey and Tibshirani 2003).

Multivariate analyses were run on 43 characters (Table 2). Only characters giving significant differences between most taxa in the Mann-Whitney *U* test were included. If a derived ratio character was included, at least one of the primary characters from which it was calculated was excluded. Each character vector was standardized by subtracting the minimum value and dividing by the range. A similarity matrix was calculated in SPLUS 6.0 (Insightful Corporation, Seattle) using the general similarity coefficient of Gower (1971), applying different similarity measures to dichotomous, qualitative, and quantative characters. The similarity matrix was decentered, and eigenvectors and eigenvalues were calculated using NTSYSpc version 2.02 (Rohlf 1999). Kendall's τ correlation coefficients between the PCO axes and the characters included in the multivariate analyses were calculated using SPSS 12.0.1 (SPSS Inc., Chicago).

RESULTS

AFLP Variation. Except for *Saxifraga bracteata*, no fixed species-specific markers were found. Of the 250 markers scored, 24 were present only in *S. hyperborea*, 13 only in *S. flexuosa*, six only in *S. bracteata*, and four only in *S. rivularis*, whereas no marker was unique to *S. arctolitoralis*. Eighteen markers, most of them fixed, were shared by all taxa except *S. bracteata*. Thirteen markers were exclusively shared by *S. flexuosa* and *S. hyperborea*, ten markers by *S. flexuosa*, *S. hyperborea*, and *S. rivularis*, and eight markers by *S. arctolitoralis*, *S. bracteata*, and *S. rivularis*.

The PCO analysis of the total material grouped the plants into three distinct groups along axes 1 and 2 (46.1% and 7.1% of the variation explained, respectively; Fig. 2): *Saxifraga rivularis* + *S. arctolitoralis*, *S. flexuosa* + *S. hyperborea* and *S. bracteata*. PCO axis 3 spanned the variation within the *S. flexuosa* + *S. hyperborea* group (5.3%; not shown). In a separate PCO analysis of the *S. arctolitoralis* + *S. rivularis* group the first axis (20.7%) separated *S. arctolitoralis* from *S. rivularis* (Fig. 2). A separate PCO analysis of the *S. flexuosa* + *S. hyperborea* group showed little structure corresponding to the tentative species (Fig. 2). However, the Atlantic populations were more or less separated from the Beringian/Pacific populations along PCO axis 1 (12.7%).

In the STRUCTURE analyses of all samples, two was the highest number of groups (K) giving an unambiguous assignment of specimens; one including *Saxifraga flexuosa* + *S. hyperborea*, and one including *S. arctolitoralis*, *S. bracteata*, and *S. rivularis* (LnP(D) = −4,801.6; not shown). When analysing these two groups separately, no further division was obtained in the *S. flexuosa* + *hyperborea* group, whereas the second group was unambiguously divided into three subgroups corresponding to *S. bracteata*, *S. arctolitoralis*, and *S. rivularis* (LnP(D) = −633.1; not shown).

In the neighbor-joining analyses the same two main groups as recognized in the STRUCTURE analyses were obtained with 100% bootstrap support (not shown). *Saxifraga bracteata* formed a strongly supported subcluster with 100% support, whereas a *S. arctolitoralis* subcluster was moderately supported (82%).

The initial parsimony analysis gave 20 trees from two islands, and additional branch-swapping resulted in 2,000 equally parsimonious trees with tree length = 631, CI = 0.300, RI = 0.821, and RC = 0.246. The two main groups found using cluster analyses, were also inferred in the parsimony

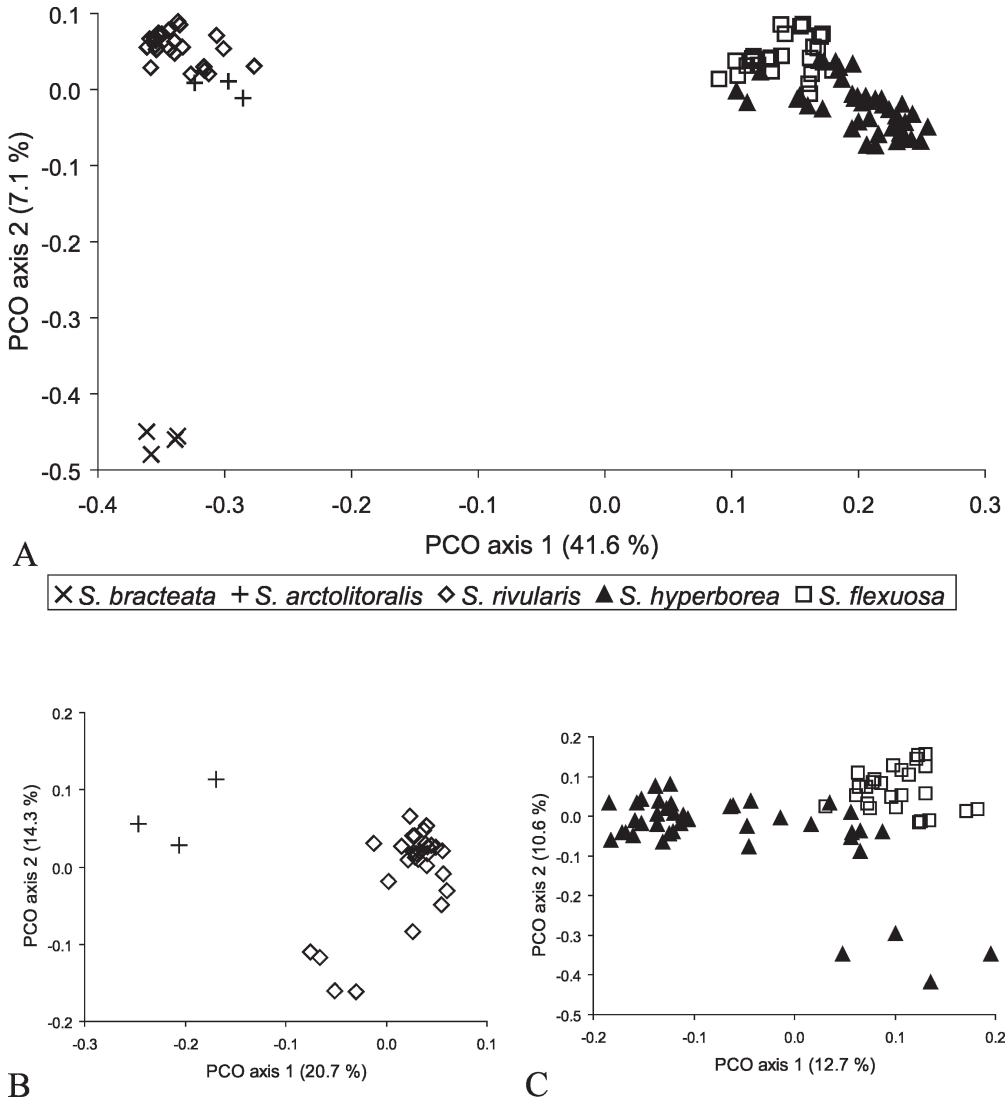


FIG. 2. PCO analyses of individuals of the *Saxifraga rivularis* complex based on Dice's similarity and 250 AFLP markers. A. Analysis of the total material. B. Analysis of 11 populations of *S. rivularis* and one population of *S. arctolitoralis*. C. Analysis of 10 populations of *S. flexuosa* and 15 populations of *S. hyperborea*.

analysis with strong (100%) bootstrap support for the *Saxifraga flexuosa* + *hyperborea* clade (Fig. 3). A *S. bracteata* clade was also strongly supported (100%), whereas a *S. arctolitoralis* + *rivularis* clade was weakly supported (61%). A subclade comprising the *S. arctolitoralis* population was moderately supported (86%).

Variation in Ploidal Levels. Three main levels of DNA content were found in the flow cytometric analysis. The DNA ratios were 0.36–0.38 in *Saxifraga rivularis*, 0.16–0.18 in *S. hyperborea* and *S. flexuosa*, and 0.20–0.21 in *S. bracteata* (Figs. 4, 5). The variation within each level of DNA content was shown by the replicates to be almost exclusively methodological. The first level corresponded to a chromosome number of $2n = 52$ (tetraploid), whereas both latter levels corresponded to a chromosome number of $2n = 26$ (diploid; Table 1). Thus, the diploid *S. bracteata* had higher DNA content than the diploid *S. flexuosa* + *S. hyperborea*, whereas the tetraploid *S. rivularis* had an amount of DNA equaling the sum of the other two.

Morphological Variation. Descriptive statistics for all characters are given in Table 3, and boxplots for selected characters are shown in Fig. 6. The Mann–Whitney *U* test showed significant differ-

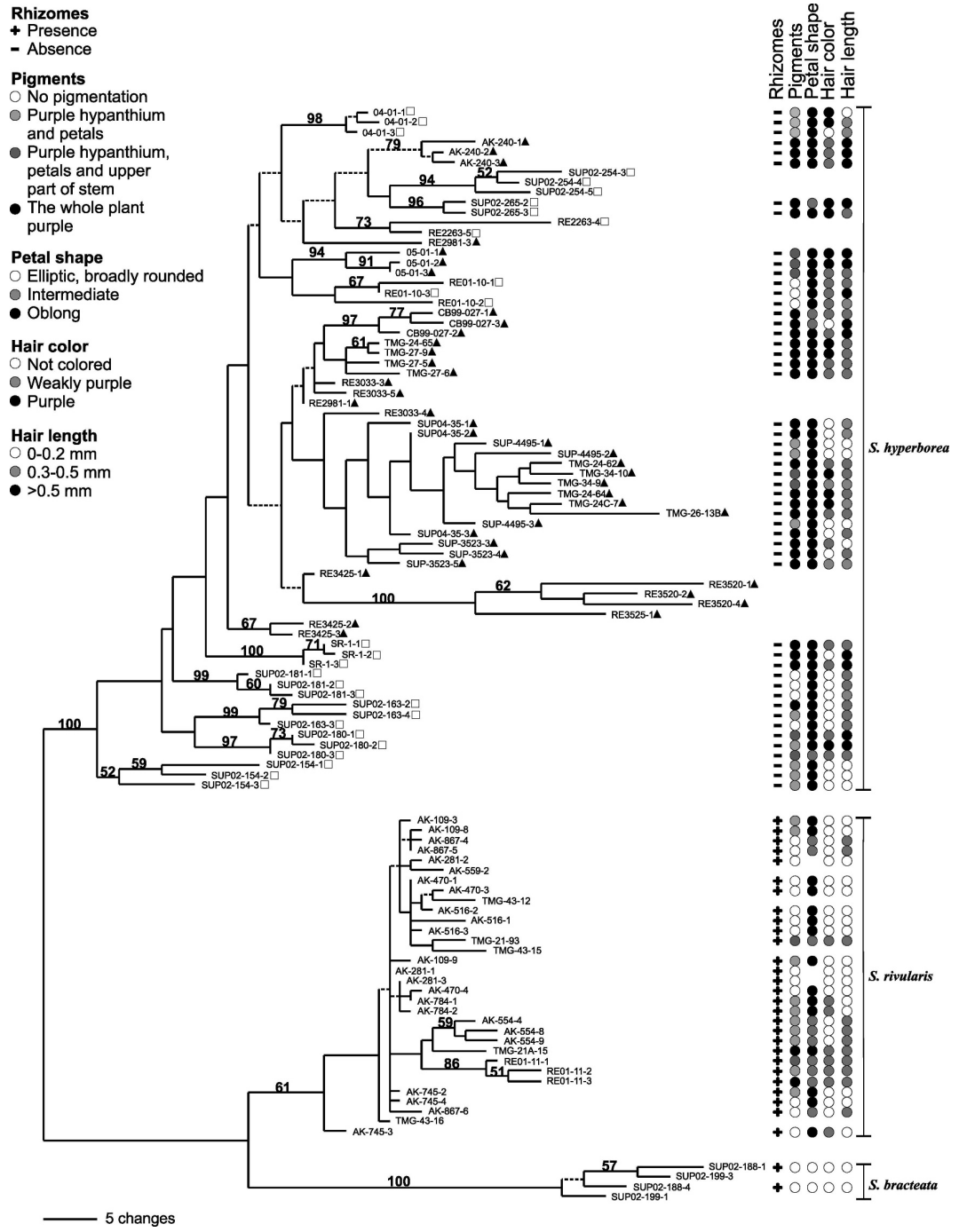


FIG. 3. One of 2,000 most parsimonious trees of the *Saxifraga rivularis* complex inferred from 174 parsimony informative AFLP characters and 104 plants. Midpoint rooted. Bootstrap values higher than 50 are indicated above supported branches. Dashed lines indicate clades collapsing in a strict consensus tree. Tree length = 631 steps, CI = 0.300, RI = 0.821, RC = 0.246. Distribution of five morphological characters among plants from populations included in the analysis, but not necessarily the same specimens, are given to the right. A priori determinations of plants to *S. hyperborea* and *S. flexuosa* are given by triangles and squares, respectively. A posteriori determinations are given by brackets.

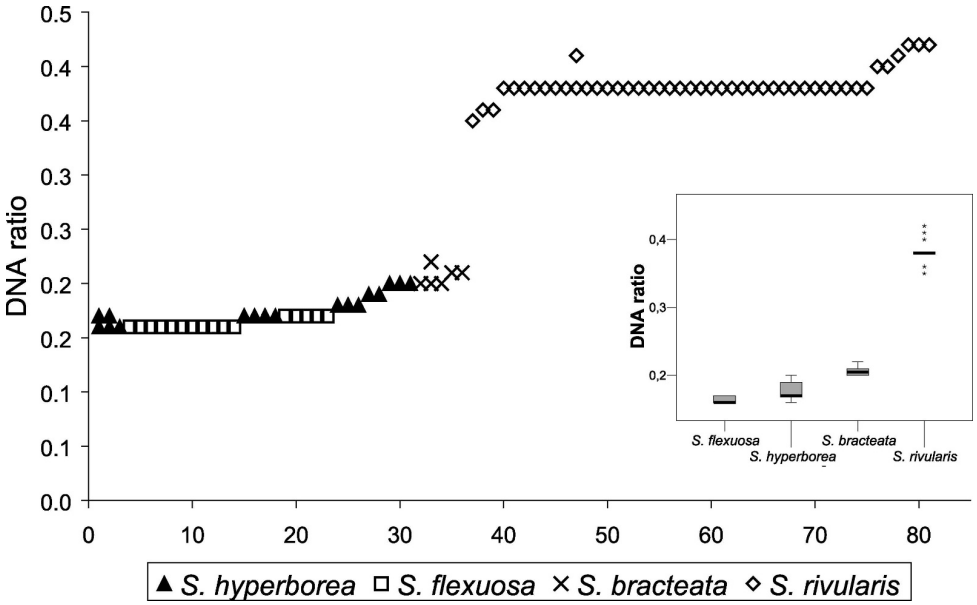


FIG. 4. DNA ratio for 76 plants, determined by flow cytometry using DAPI staining and *Lactuca sativa* as internal standard. *Saxifraga arctolitoralis* and *S. debilis* were not included due to lack of fresh leaf material. Symbols with the same position along the first axis in the scatterplot are duplicates from the same individual.

ences in more than half of the characters when comparing *Saxifraga bracteata* with the other five tentative taxa (i.e., in 34–40 characters), whereas only 16 characters differed significantly between *S. hyperborea* and *S. flexuosa* (Table 4). The other combinations of taxa showed significant differences in 23–36 characters (Table 4).

The material referred to *Saxifraga debilis* lacked rhizomes and had V-shaped hypanthia that were glabrous to sparsely covered with very short glandular hairs. Longer flowering pedicels and narrower sepals distinguished it from the other taxa. Its leaves were larger and with more lobes than in the other taxa, except *S. bracteata*. The material of *S. bracteata* had rhizomes and broadly elliptic petals and the plants were larger than in the other taxa, with longer flowering stems, longer and broader bracts and leaves, and larger flowers, as well as more numerous flowers, bracts, and leaf lobes.

The material of *Saxifraga flexuosa* and *S. hyperborea* could be separated from that of *S. rivularis* and *S. arctolitoralis* by lack of rhizomes and by the darker color of the partition-walls in the hairs. The material referred to *S. flexuosa* differed significantly from that of *S. hyperborea* in the amount of pigmentation, in the color of the hair partition-walls (being more purple in *S. hyperborea*), and in its longer flowering stems and pedicels. The material of *S. arctolitoralis* was shorter than that of *S. rivularis*, more pigmented, and more densely covered by longer and more pigmented hairs (Tables 3, 4; Fig. 6).

Even though a series of characters differed between the taxa, overlap was always observed in the multivariate analyses of the morphological data. In the PCO analysis of all samples, most of the *Saxifraga bracteata* specimens were found at one end of axis 1, *S. flexuosa* and *S. hyperborea* at the other end, and *S. arctolitoralis*, *S. debilis* and *S. rivularis* in between (11.68%; Fig. 7A). The first PCO axis was strongly correlated ($p<0.001$) with characters 1 ($\tau = -0.62$; cf. Table 2), 28 ($\tau = -0.53$), 31 ($\tau = -0.51$), 47 ($\tau = -0.51$), and 40 ($\tau = -0.52$). The Rocky Mountain samples of *S. flexuosa* and *S. debilis* were found at low values along axis 2 (6.48%), partly separated from the other samples. This axis was correlated ($p<0.001$) with characters 14 ($\tau = 0.39$), 16 ($\tau = 0.40$), 18 ($\tau = 0.44$), 20 ($\tau = 0.30$) and 51 ($\tau = 0.30$). *Saxifraga arctolitoralis* was partly separated from the other taxa along axis 3 (5.73%; not shown), which was correlated ($p<0.001$) with characters 18 ($\tau = 0.28$) and 49 ($\tau = 0.27$).

When analysed separately, *Saxifraga arctolitoralis* and *S. rivularis* were almost completely separated along PCO axis 1 (12.83%; Fig. 7B). This axis was strongly correlated ($p<0.001$) with characters 9 ($\tau = 0.53$), 10 ($\tau = 0.67$), 11 ($\tau = 0.55$), 16 ($\tau = 0.57$) and 18 ($\tau = 0.57$).

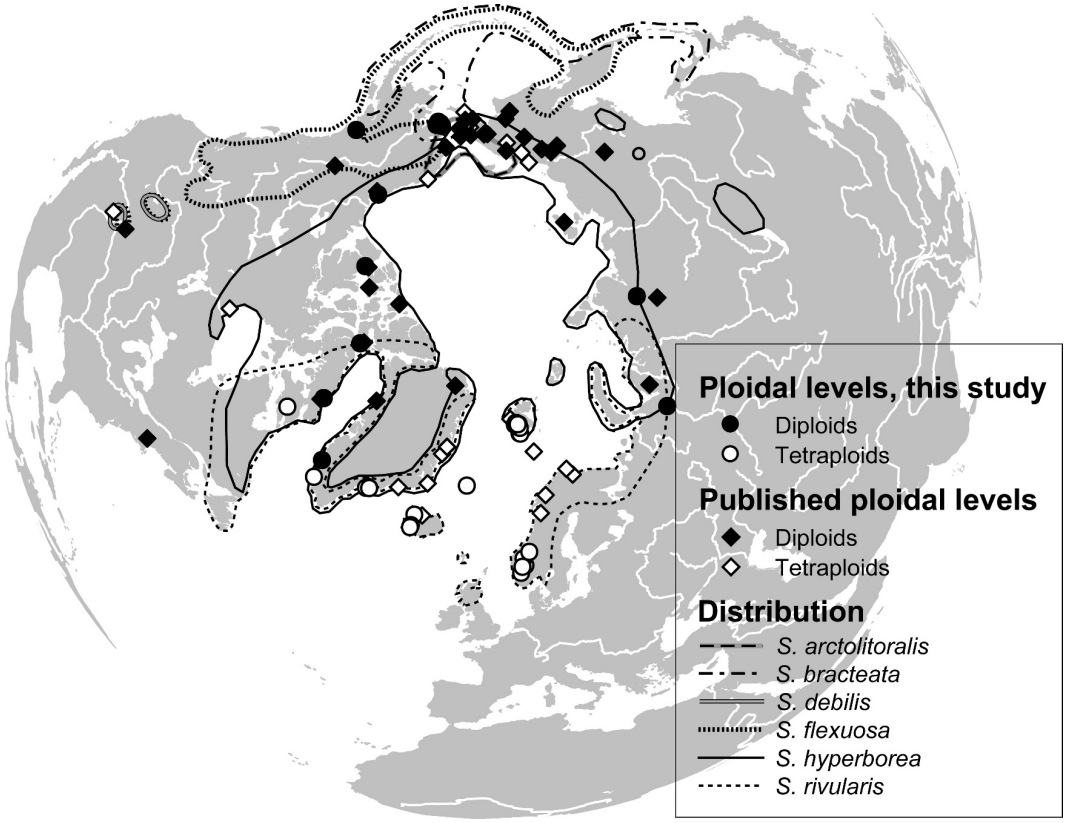


FIG. 5. Recorded ploidal levels in the *Saxifraga rivularis* complex given as diploid (black) and tetraploid (white). The ploidal levels are estimated from previous chromosome counts and ploidal level estimates (diamonds; see references under Taxonomic treatment and Löve et al. 1971), and our flow cytometry results (circles).

A separate PCO analysis of the samples referred to as *Saxifraga debilis*, *S. flexuosa*, and *S. hyperborea* did not result in clear separation of taxa (Fig. 7C). However, *S. debilis* was exclusively found at high values along axes 1 (9.21%) and 2 (8.05%). The Rocky Mountain samples of *S. flexuosa* were found at high values along axis 1, but at low values along axis 2, whereas the other samples of *S. flexuosa* and *S. hyperborea* were found along the lower half of axis 1. Axis 1 was correlated with characters 8 ($\tau = 0.35$, $p<0.001$), 11 ($\tau = -0.27$, $p<0.001$), 18 ($\tau = 0.26$, $p<0.001$), 31 ($\tau = -0.26$, $p<0.001$), and 41 ($\tau = -0.28$, $p=0.004$).

Analyzing *Saxifraga flexuosa* and *S. hyperborea* separately gave no further separation (not shown). Also in this analysis, axis 1 (10.70%) separated the Rocky Mountain samples from the other samples and was correlated ($p<0.001$) with characters 2 ($\tau = -0.35$), 14 ($\tau = -0.51$), 15 ($\tau=0.32$), and 18 ($\tau = -0.40$). When all samples from the Rocky Mountains were analyzed separately, *S. debilis* and *S. flexuosa* were almost completely separated along axis 1 (19.49%; not shown), strongly correlated ($p<0.001$) with characters 6 ($\tau = 0.68$), 12 ($\tau = 0.63$), 14 ($\tau = 0.64$), 22 ($\tau = -0.64$), 23 ($\tau = -0.60$), and 46 ($\tau = 0.63$).

The only three qualitative morphological characters dividing the *Saxifraga rivularis* complex into distinct groups were the presence versus absence of rhizomes (character 1), the shape of the hypanthium (10 and 11), and the shape of petals (28). Rhizomes were almost uniformly present in the material of *S. arctolitoralis*, *S. bracteata*, and *S. rivularis*, and absent in *S. debilis*, *S. hyperborea*, and *S. flexuosa* (Fig. 3; Tables 3, 4). The hypanthium shape distinguished *S. debilis* (V-shaped) from all other taxa (U-shaped). The petal shape distinguished *S. bracteata* (broadly elliptic) from all other taxa (more or less oblong). We also observed quite clear tendencies in the degree of pigmentation (character 8 and 18). Typically, *S. hyperborea* and *S. arctolitoralis* had much pigmentation, *S. bracteata*, *S. debilis*, and *S. rivularis* had little or no pigmentation, and *S. flexuosa* was intermediate (Fig. 3; Tables 3, 4).

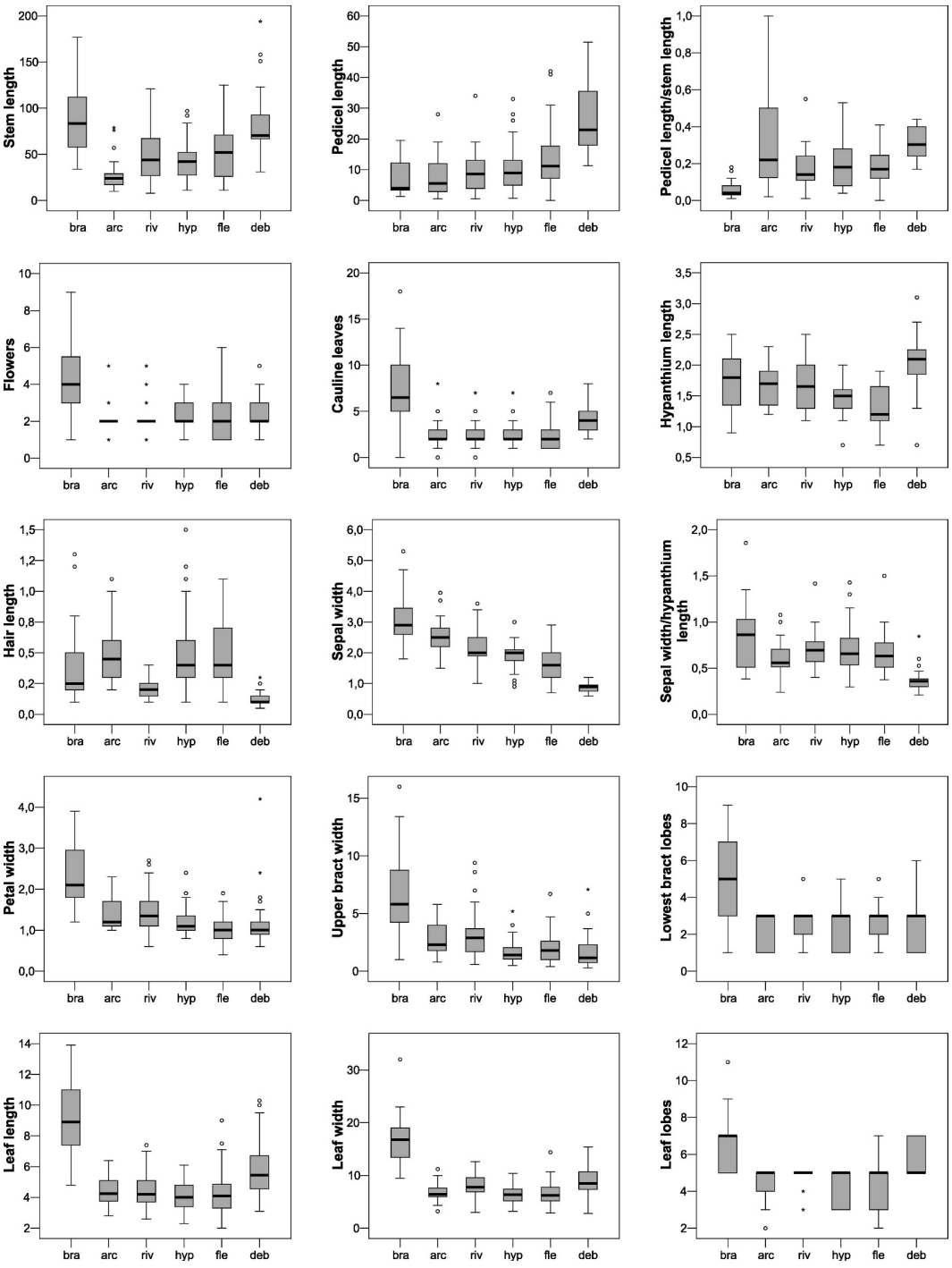


FIG. 6. Variation in 15 morphological characters in the *Saxifraga rivularis* complex. Taxa are given by abbreviations: arc = *S. arctolitoralis*, bra = *S. bracteata*, deb = *S. debilis*, fle = *S. flexuosa*, hyp = *S. hyperborea*, and riv = *S. rivularis*. The character codes are explained in Table 2. One outlier (*S. bracteata*; 804-1) is off scale in the box plots of the characters Flowers and Cauline leaves due to its extreme values (29 and 35, respectively).

TABLE 3. Descriptive statistics for all morphological characters in the *Saxifraga rivularis* complex. Character code is according to Table 2. N is number of plants.

	S. bracteata					*S. arctolitoralis*					*S. rivularis*					*S. hyperborea*					*S. flexuosa*					*S. debilis*				
Character	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Rhizomes	27	0.9	0.27	0	1	23	1.0	0.00	1	1	37	0.9	0.23	0	1	66	0.0	0.00	0	0	63	0.0	0.00	0	0	36	0.0	0.00	0	0
Stem length	34	90.9	40.22	34	177	33	27.1	16.34	10	79	43	48.2	27.73	8	121	68	42.5	20.69	11	97	66	51.5	28.15	11	125	36	82.8	31.82	31	194
Pedicel length	17	7.19	6.266	1.3	19.5	23	8.08	6.813	0.5	28.0	29	9.11	7.354	0.5	34.0	37	10.38	8.082	0.7	33.0	32	13.80	10.455	0.0	42.0	18	26.73	10.690	11.3	51.5
4	16	0.062	0.0514	0.01	0.18	23	0.305	0.2422	0.02	1.00	29	0.171	0.1170	0.01	0.55	37	0.203	0.1396	0.04	0.53	32	0.188	0.1069	0.00	0.41	18	0.309	0.0803	0.17	0.44
Flowers	35	5.0	4.65	1	29	33	2.2	0.74	1	5	43	2.1	0.73	1	5	68	2.4	0.85	1	4	66	2.1	1.08	1	6	36	2.5	1.23	1	5
Cauline leaves	34	8.1	6.00	0	35	33	2.5	1.50	0	8	43	2.3	1.27	0	7	67	2.5	1.32	1	7	66	2.3	1.34	1	7	36	4.1	1.37	2	8
7	34	0.097	0.0543	0.00	0.31	33	0.108	0.0704	0.00	0.29	43	0.057	0.0348	0.00	0.15	67	0.072	0.0468	0.01	0.22	66	0.052	0.0297	0.01	0.17	36	0.053	0.0190	0.02	0.10
Pigments	36	0.1	0.28	0	1	19	2.3	0.95	0	3	43	1.0	1.25	0	3	68	2.6	0.74	0	3	66	1.6	1.21	0	3	36	0.6	1.13	0	3
Density stem	35	1.1	0.48	0	2	33	1.5	0.51	1	2	43	0.8	0.43	0	1	68	1.2	0.41	0	2	66	0.8	0.60	0	2	36	0.9	0.44	0	2
Hypanthium shape (10)	17	0.5	0.51	0	1	7	1.0	0.00	1	1	10	0.1	0.32	0	1	26	0.8	0.43	0	1	19	0.7	0.48	0	1	11	1.7	0.47	1	2
Hypanthium shape (11)	16	0.8	0.45	0	1	26	0.9	0.27	0	1	33	0.3	0.47	0	1	41	0.8	0.40	0	1	47	0.8	0.43	0	1	25	1.6	0.49	1	2
Hypanthium length (12)	15	1.74	0.466	0.9	2.5	7	1.67	0.399	1.2	2.3	10	1.67	0.467	1.1	2.5	26	1.48	0.271	0.7	2.0	19	1.32	0.345	0.7	1.9	11	2.03	0.642	0.7	3.1
Hypanthium length (13)	16	2.82	1.051	1.3	5.2	26	2.59	0.934	1.2	5.0	33	2.24	0.587	1.2	3.8	41	1.96	0.700	1.0	3.8	47	1.95	0.556	0.9	3.6	25	2.55	0.736	1.0	4.0
Glandular hairs	34	1.0	0.00	1	1	32	1.0	0.00	1	1	43	1.0	0.00	1	1	68	1.0	0.00	1	1	66	0.7	0.46	0	1	36	0.8	0.44	0	1
Density hypanthium	34	1.2	0.41	1	2	32	1.5	0.51	1	2	43	1.1	0.32	1	2	68	1.3	0.47	1	2	46	1.3	0.46	1	2	27	1.1	0.32	1	2
Hair length	34	0.393	0.3262	0.10	1.30	32	0.51	0.245	0.20	1.10	43	0.207	0.0803	0.10	0.40	68	0.505	0.2785	0.10	1.50	46	0.483	0.2284	0.10	1.10	27	0.119	0.0638	0.05	0.30
Hair cell no.	34	4.4	1.78	2	8	32	4.0	1.38	3	8	43	3.5	0.70	3	5	68	4.9	1.67	2	10	46	4.8	1.49	2	10	27	3.4	0.75	2	5
Hair color	35	1.2	0.83	0	2	33	1.6	0.49	1	2	43	0.5	0.67	0	2	68	1.4	0.63	0	2	52	1.1	0.80	0	2	31	0.9	0.56	0	2
Sepal length	35	2.03	0.689	1.2	4.9	33	1.41	0.276	0.9	2.1	42	1.46	0.297	1.0	2.3	68	1.32	0.393	0.9	2.8	66	1.38	0.391	0.7	2.2	36	1.90	0.300	1.3	2.9
Sepal width	35	3.06	0.759	1.8	5.3	33	2.50	0.530	1.5	4.0	42	2.17	0.466	1.0	3.6	68	1.91	0.380	0.9	3.0	66	1.65	0.480	0.7	2.9	36	0.87	0.154	0.6	1.2
21	16	1.317	0.6310	0.53	3.21	26	1.072	0.3045	0.60	1.86	32	1.027	0.2736	0.61	1.80	46	1.175	0.3716	0.57	2.57	47	1.039	0.3146	0.58	1.80	25	0.784	0.2900	0.45	1.77
22	16	0.854	0.4058	0.38	1.86	26	0.612	0.1969	0.24	1.08	32	0.698	0.2002	0.40	1.42	46	0.703	0.2335	0.30	1.43	47	0.669	0.2164	0.38	1.50	25	0.376	0.1325	0.21	0.85
23	35	0.679	0.1863	0.36	1.06	33	0.579	0.1290	0.38	1.00	42	0.678	0.1383	0.48	1.18	68	0.610	0.1286	0.41	1.00	66	0.648	0.1460	0.39	1.21	36	0.467	0.0973	0.34	0.69
Sepal shape	35	0.9	0.65	0	2	33	0.5	0.51	0	1	42	1.1	0.77	0	2	68	1.0	0.66	0	2	66	1.0	0.59	0	2	36	1.1	0.67	0	2
Petal length	28	4.42	1.099	2.3	6.5	23	3.68	0.842	2.3	5.5	32	3.34	0.843	1.4	5.4	64	3.20	0.614	1.7	4.5	55	2.85	0.525	1.6	3.8	34	3.66	0.967	1.7	6.2
Petal width	27	2.36	0.813	1.2	3.9	23	1.38	0.390	1.0	2.3	32	1.46	0.504	0.6	2.7	64	1.21	0.310	0.8	2.4	54	1.00	0.291	0.4	1.9	34	1.18	0.638	0.6	4.2
27	27	0.537	0.1459	0.34	1.03	23	0.378	0.0694	0.25	0.49	32	0.443	0.1259	0.24	0.79	64	0.386	0.0897	0.25	0.75	54	0.352	0.0807	0.18	0.55	34	0.321	0.1023	0.20	0.68
Petal shape	29	1.9	0.26	1	2	23	0.8	0.39	0	1	32	0.4	0.61	0	2	64	0.1	0.24	0	1	55	0.2	0.42	0	2	34	0.4	0.65	0	2
Gynoecium length	12	3.62	0.743	2.1	4.5	16	3.89	0.761	2.3	6.0	23	3.70	0.751	2.3	5.0	21	3.50	0.668	2.7	5.2	25	2.72	0.553	1.6	4.1	12	2.84	0.568	1.8	3.8
Dist_ubract	32	1.0	0.00	1	1	29	1.0	0.00	1	1	38	1.0	0.00	1	1	61	1.0	0.13	0	1	45	1.0	0.00	1	1	28	1.0	0.00	1	1
Upper bract petiole	31	0.7	0.44	0	1	29	0.4	0.51	0	1	38	0.5	0.51	0	1	59	0.0	0.13	0	1	45	0.1	0.32	0	1	28	0.2	0.42	0	1
Upper bract length	31	6.70	3.101	2.3	14.7	29	5.31	2.194	2.6	10.5	38	5.02	1.774	2.5	8.9	58	4.61	1.648	1.2	8.7	45	5.05	2.094	1.4	9.5	28	4.35	1.299	2.1	6.8

TABLE 3. Continued.

Character	*S. bracteata*					*S. arctolitoralis*					*S. rivularis*					*S. hyperborea*					*S. flexuosa*					*S. debilis*				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Upper bract width	31	6.78	3.765	1.0	16.0	29	2.76	1.445	0.8	5.8	38	3.19	2.035	0.6	9.4	55	1.67	0.902	0.5	5.2	43	1.97	1.307	0.4	6.7	28	1.74	1.494	0.3	7.1
34	31	1.061	0.4520	0.39	2.03	29	0.585	0.3840	0.19	1.58	38	0.659	0.4218	0.20	2.00	54	0.373	0.1661	0.17	0.95	44	0.400	0.2530	0.16	1.76	28	0.387	0.2772	0.11	1.25
Upper bract lobes	31	2.7	1.42	1	7	29	1.8	1.33	1	5	38	1.6	1.03	1	5	58	1.2	0.63	1	5	45	1.4	0.87	1	5	28	1.3	0.84	1	5
Dist_lbract	32	1.0	0.00	1	1	9	1.0	0.00	1	1	12	0.8	0.45	0	1	31	1.0	0.18	0	1	19	0.9	0.32	0	1	18	1.0	0.00	1	1
Lowest bract petiole	31	0.9	0.25	0	1	9	0.7	0.50	0	1	9	0.9	0.33	0	1	30	0.5	0.51	0	1	16	0.5	0.52	0	1	18	0.8	0.38	0	1
Petiole length	27	5.06	5.741	0.8	29.0	6	10.67	4.792	6.0	19.0	8	8.86	4.513	1.6	17.2	16	5.05	2.428	1.7	8.8	8	5.63	4.774	2.3	14.2	15	4.63	4.486	1.4	19.5
Lowest bract length	30	9.24	3.114	4.8	15.9	8	5.68	3.079	3.5	13.0	8	6.14	2.564	3.1	10.2	30	5.35	1.277	3.6	8.0	16	6.49	2.173	3.0	12.5	17	4.98	1.609	2.5	8.2
Lowest bract width	28	13.21	6.908	3.8	33.6	7	4.17	2.504	1.2	7.5	7	5.69	2.255	1.5	8.3	30	3.72	2.004	0.5	8.0	16	5.49	2.443	1.4	9.2	17	4.24	2.692	0.9	9.8
Lowest bract lobes	30	5.0	1.98	1	9	8	2.3	1.04	1	3	8	2.8	1.28	1	5	30	2.2	1.22	1	5	16	2.8	1.13	1	5	17	2.4	1.50	1	6
Lowest bract shape	26	1.4	0.50	1	2	4	1.0	0.00	1	1	7	1.3	0.76	0	2	22	0.6	0.49	0	1	14	0.5	0.52	0	1	17	0.9	0.66	0	3
Leaf length	33	9.24	2.449	4.8	13.9	33	4.37	0.884	2.8	6.4	43	4.40	1.085	2.6	7.4	66	4.04	0.903	2.3	6.1	64	4.17	1.229	2.0	9.0	36	5.79	1.773	3.1	10.3
Leaf width	33	16.67	4.806	9.5	32.0	33	6.84	1.525	3.2	11.2	42	8.11	2.048	3.0	12.6	66	6.41	1.633	3.2	10.4	63	6.57	2.067	2.9	14.4	36	8.89	2.695	2.8	15.4
45	33	1.817	0.2687	1.36	2.67	33	1.591	0.3206	0.91	2.21	42	1.842	0.2591	1.11	2.50	66	1.591	0.2411	0.95	2.12	63	1.568	0.1959	1.06	2.00	36	1.547	0.2256	0.64	1.92
Dist_inc	33	5.04	1.682	2.1	7.5	33	2.20	0.569	1.1	3.6	43	1.91	0.686	0.9	4.2	66	1.95	0.562	0.8	3.0	63	2.06	0.955	0.7	5.5	36	3.22	1.202	1.4	6.4
Incision shape	33	1.4	0.61	0	2	33	0.8	0.53	0	2	43	1.4	0.73	0	2	66	0.3	0.48	0	1	63	0.4	0.62	0	2	36	1.1	0.75	0	2
Leaf lobes	33	6.8	1.54	5	11	33	4.4	0.94	2	5	43	4.8	0.53	3	5	66	4.2	0.98	3	5	64	4.4	1.01	2	7	36	5.9	1.00	5	7
Leaf shape	33	1.7	0.48	1	2	32	1.1	0.44	0	2	43	0.8	0.60	0	2	66	0.5	0.53	0	2	64	0.5	0.59	0	3	36	1.0	1.00	0	3
Sheath structure	32	1.5	0.72	0	2	33	0.9	0.24	0	1	43	1.3	0.73	0	2	68	0.6	0.60	0	2	62	0.8	0.61	0	2	36	0.9	0.28	0	1
Sheath shape	31	1.1	0.88	0	2	32	1.3	0.60	0	2	43	1.5	0.74	0	2	68	0.8	0.84	0	3	62	0.7	0.76	0	2	36	0.1	0.40	0	2
Sheath fringes	31	1.0	0.18	1	2	33	1.1	0.33	1	2	43	0.9	0.29	0	1	68	1.0	0.12	1	2	62	0.9	0.36	0	2	36	1.1	0.49	0	2

TABLE 4. Mann-Whitney *U* test for equality of medians in measured and derived characters. Character code is according to Table 2. Taxa are abbreviated: arc - *Saxifraga arctolitoralis*, bra - *S. bracteata*, deb - *S. debilis*, fle - *S. flexousa*, hyp - *S. hyperborea* and riv - *S. rivularis*. ss means difference significant at $q < 0.01$, s means difference significant at $q < 0.05$, and blank means no significant difference in medians. The sum of characters with significant difference in median between the two taxa compared is given in the last row.

		arc bra	arc deb	arc fle	arc hyp	arc riv	bra deb	bra fle	bra hyp	bra sriv	deb fle	deb hyp	deb riv	fle hyp	fle riv	hyp riv
1	*Rhizomes*		ss	ss	ss		ss	ss	ss				ss		ss	ss
2	*Stem length*	ss	ss	ss	ss	ss		ss	ss	ss	ss	ss	ss			
3	*Pedicel length*		ss	s			ss	s			ss	ss	ss			
4		ss				s	ss	ss	ss	ss	ss	ss	ss			
5	*Flowers*	ss					ss	ss	ss	ss				s		s
6	*Cauline leaves*	ss	ss				ss	ss	ss	ss	ss	ss	ss			
7			ss	ss	ss	ss	ss	ss	ss	ss				s		
8	*Pigments*	ss	ss	s		ss	s	ss	ss	ss	ss	ss		ss	s	ss
9	*Density stem*	ss	ss	ss	ss	ss		s		s		ss		ss		ss
10	*Hypanthium shape*	s	ss			ss	ss			s	ss	ss	ss		ss	ss
11	*Hypanthium shape*		ss			ss	ss			ss	ss	ss	ss		ss	ss
12	*Hypanthium length*			s				ss			ss	ss			s	
13	*Hypanthium length*			ss	ss			ss	ss	s	ss	ss	s		s	s
14	*Glandular hairs*		ss	ss			ss	ss				ss	ss	ss	ss	
15	*Density hypanthium*	s	ss			ss						s			s	s
16	*Hair length*	ss	ss			ss	ss	ss	ss	ss	ss	ss	ss		ss	ss
17	*Hair cell no.*			ss	ss		s			s	ss	ss			ss	ss
18	*Hair color*	s	ss	ss		ss	s			ss		ss	ss	ss	ss	ss
19	*Sepal length*	ss	ss		s			ss	ss	ss	ss	ss	ss			ss
20	*Sepal width*	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss	ss
21			ss				ss				ss	ss	ss	s		s
22			ss				ss				ss	ss	ss			
23		s	ss	s		ss	ss		s		ss	ss	ss			ss
24	*Sepal shape*	s	ss	ss	ss	ss										
25	*Petal length*	s		ss	s		ss	ss	ss	ss	ss	s		ss	ss	
26	*Petal width*	ss	ss	ss			ss	ss	ss	ss		s	ss	ss	ss	ss
27		ss	ss			s	ss	ss	ss	s	s	ss	ss	s	ss	s
28	*Petal shape*	ss	ss	ss	ss	ss	ss	ss	ss	ss		ss			s	ss
29	*Gynoecium length*		ss	ss	s		ss	ss				s	ss	ss	ss	
30	*Dist_ubract*															
31	*Upper bract petiole*	s	s	ss	ss		ss	ss	ss	s		ss	s	s	ss	ss
32	*Upper bract length*						ss	s	ss	s						
33	*Upper bract width*	ss	ss	s	ss		ss	ss	ss	ss			ss		ss	ss
34		ss	s	s	ss		ss	ss	ss	ss			ss		ss	ss
35	*Upper bract lobes*	ss	s		s		ss	ss	ss	ss			s			ss
36	*Dist_lbract*							s		ss			s			s
37	*Lowest bract petiole*	s						ss	ss		s	s			s	s
38	*Petiole length*	s	ss	s	ss					s			s			s
39	*Lowest bract length*	ss					ss	ss	ss	s	s			s		
40	*Lowest bract width*	ss					ss	ss	ss	ss				s		s
41	*Lowest bract lobes*	ss					ss	ss	ss	ss						
42	*Lowest bract shape*						ss	ss	ss		s				s	s
43	*Leaf length*	ss	ss				ss	ss	ss	ss	ss	ss	ss			
44	*Leaf width*	ss	ss			ss	ss	ss	ss	ss	ss	ss			ss	ss
45		ss				ss	ss	ss	ss				ss		ss	ss
46	*Dist_inc*	ss	ss	s	s	ss	ss	ss	ss	ss	ss	ss	ss			
47	*Incision shape*	ss		ss	ss	ss		ss	ss		ss	ss			ss	ss
48	*Leaf lobes*	ss	ss			s	s	ss	ss	ss	ss	ss	ss		ss	ss
49	*Leaf shape*	ss		ss	ss	s	ss	ss	ss	ss	s	s			s	s
50	*Sheath structure*	ss			ss	s	ss	ss	ss			ss	ss	s	ss	ss
51	*Sheath shape*		ss	ss	ss		ss	s		s	ss	ss	ss		ss	ss
52	*Sheath fringes*			s	s	ss				s	s	s	s			ss
	Total	35	33	27	23	23	39	40	34	36	29	36	32	16	29	35

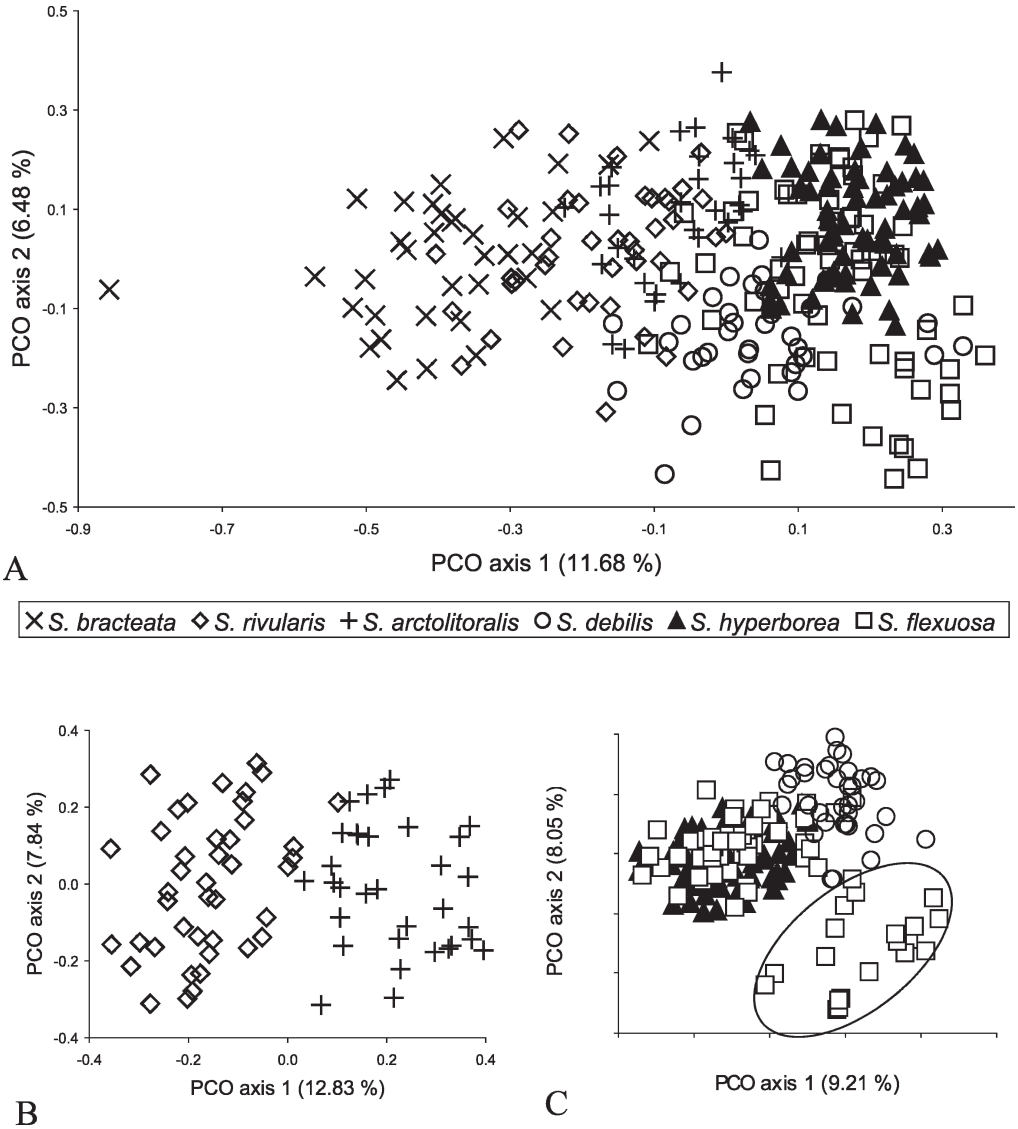


FIG. 7. PCO analyses of the *Saxifraga rivularis* complex based on 43 morphological characters (Table 2) standardized by ranging and using the general similarity coefficient of Gower. A. Analysis of the total material. B. Analysis of samples of *S. rivularis* and *S. arctolitoralis*. C. Analysis of samples of *S. flexuosa*, *S. hyperborea* and *S. debilis*. Encircled are the Rocky Mountain populations of *S. flexuosa*.

DISCUSSION

Based on our morphological, flow cytometrical, and molecular data, we conclude that it is reasonable to recognize four species in the *Saxifraga rivularis* complex, including one with two subspecies: the amphi-Pacific diploid *S. bracteata*, the Cordilleran diploid *S. debilis*, the circumpolar diploid *S. hyperborea*, and the tetraploid *S. rivularis* with the amphi-Atlantic subspecies *rivularis* and the amphi-Beringian subspecies *arctolitoralis* (new combination proposed here). We regard the amphi-Pacific and Cordilleran populations referred to as *S. flexuosa* as conspecific with *S. hyperborea*. In the following discussion, we refer to these two initially hypothesized taxa as '*flexuosa*' and '*hyperborea*'.

Main Evolutionary Lineages. The three very divergent groups obtained in all PCO (Fig. 2), neighbor-joining and parsimony analyses (Fig. 3) of the AFLP data probably represent three evolutionary lineages within the *Saxifraga rivularis*

complex: the diploid *S. bracteata* lineage, the diploid *S. hyperborea* lineage (including '*flexuosa*') and the tetraploid *S. rivularis* lineage. These groups are also supported by differences in morphological characters (Figs. 6, 7) and in genome size (Fig. 4).

Although AFLP data were not obtained for the Cordilleran diploid *S. debilis*, its unique morphology (distinctly different hypanthium shape, shorter glandular hairs and characteristics of cauline leaves; Figs. 6, 7C) indicates that it represents a separate lineage as well. This conclusion is supported by sequence analysis of several noncoding cpDNA regions (T. M. Gabrielsen & C. Brochmann, unpubl. data). The three diploid lineages have quite divergent chloroplast DNA haplotypes, whereas the single cpDNA haplotype of the tetraploid *S. rivularis* lineage is very similar to, but not identical with those of the diploid *S. hyperborea* lineage.

Allopolyploidy as a Single Event. The *Saxifraga rivularis* lineage is most likely an allopolyploid derivative from the *S. bracteata* and *S. hyperborea* lineages, a hypothesis supported by the intermediacy of this lineage between the others both genetically (Figs. 2, 3) and morphologically (Figs. 6, 7B). The two main levels of DNA ratios identified in our flow cytometric analysis correspond to previously published chromosome numbers of $2n = 52$ (tetraploid) for *S. rivularis* and $2n = 26$ for *S. hyperborea* and *S. bracteata* (diploid; cf. Table 1 and the taxonomic treatment). No material of *S. debilis* or *S. rivularis* subsp. *arctolitoralis* was available for our flow cytometric analysis, but according to previous counts, *S. debilis* is diploid ($2n = 26$; Löve et al. 1971) and *S. rivularis* subsp. *arctolitoralis* is tetraploid ($2n = 52$; cf. references in the taxonomic treatment).

It is well-known that the genome size of polyploids does not necessarily represent the sum of their diploid progenitors due to genome size evolution after their formation (Levin 2002), but the finding of close to perfect additivity in *Saxifraga rivularis* (with DNA ratio 0.36–0.38) relative to *S. hyperborea* (0.16–0.18) and *S. bracteata* (0.20–0.21; Figs. 4, 5) provides a compelling argument in favor of allopolyploid origin from these two diploid lineages. Moreover, the low number of specific markers and the lack of fixed specific markers within the *S. rivularis* lineage suggest that polyploid formation involving other taxa is unlikely. Thus, for these reasons and also because of its divergent morphology and its distribution in the southern Rocky Mountains, the diploid *S. debilis* is unlikely a progenitor.

Alloploid origin of the *Saxifraga rivularis* lineage with the circumpolar *S. hyperborea* lineage as one of the diploid progenitors is also consistent with isozyme data (Guldahl et al. 2005), and chloroplast DNA sequences (Brochmann et al. 1998; T. M. Gabrielsen & C. Brochmann, unpubl. data). Both indicate that this lineage represents the maternal progenitor. Thus, the Beringian *S. bracteata* lineage probably represents the paternal progenitor, suggesting a Beringian origin of the tetraploid. The hypothesis of Rebristaya and Yurtsev (1984) that the Beringian tetraploid (*S. rivularis* subsp. *arctolitoralis*) originated as an autopolyploid from *S. hyperborea* can thereby be definitely rejected, but their view that *S. bracteata* and *S. hyperborea* are closely related to *S. rivularis* subsp. *rivularis* gains support.

The low number of specific AFLP-markers, and the shallow genetic structure within the *Saxifraga rivularis* lineage, suggest that *S. rivularis* originated once in Beringia and that the subsequent divergence of the two subspecies happened quite recently. This result is supported by the single cpDNA haplotype found in both subsp. *arctolitoralis* and subsp. *rivularis* (T. M. Gabrielsen and C. Brochmann, unpubl. data). However, the morphological distinction including differences in size, hair length, hair pigmentation and hair density, paired with moderate genetic differentiation (STRUCTURE analyses; Figs. 2, 3), justifies the recognition of two significant units and also suggests that the split between the two is older than post-glacial.

Biogeography. The highest taxonomic and molecular diversity was found in Beringia, suggesting that this area has served as a continuous primary Pleistocene refugium and area of diversification of the *Saxifraga rivularis* complex as shown for many other plant groups (Hultén 1937; Abbott and Brochmann 2003).

The circumpolar *Saxifraga hyperborea* showed some phylogeographic structure in the PCO analysis of the AFLP data (Fig. 2C), indicating that the Pacific and Atlantic populations became isolated in different refugia, at least during the last glaciation. The Rocky Mountain populations, which were not included in the AFLP analysis, differed morphologically from the arctic populations (Fig. 7C), suggesting that a third glacial refugium for *S. hyperborea* was situated in western North America south of the Cordilleran ice sheet, as suggested for several other arctic-alpine species (*Dryas integrifolia* M.Vahl., Tremblay and Schoen 1999; *S. oppositifolia* L., Abbott and Comes 2004; species of *Boechera* Á.Löve & D.Löve, Dobeš et al. 2004; *Vaccinium uliginosum* L., Alsos et al. 2005).

The current, extremely disjunct distribution ranges of the amphi-Atlantic subsp. *rivularis* and

the amphi-Beringian subsp. *arctolitoralis*, separated by approx. 4,000 km in Russia and by almost 3,000 km in North America (Fig. 1), are paralleled by very few non-littoral species with a significant arctic range. The two closest parallels are *Luzula arcuata* (Wahlenb.) Sw., with the amphi-Atlantic subsp. *arcuata* and the amphi-Beringian subsp. *unalaschkensis* (Buchenau) Hultén (see Hultén 1968), and *Ranunculus (Beckwithia) glacialis* L., with the amphi-Atlantic subsp. *glacialis* and the very restricted Beringian subsp. *alaskensis* Jurtz., D.F.Murray & S.Kelso ined. (Lipkin and Murray 1997). In both cases the gaps in the distribution are of the same order, or larger, than in *Saxifraga rivularis*. These three arctic-alpine species have in common that they are confined to the most humid and least winter-cold parts of the Arctic, bordering on the northernmost Atlantic and Pacific oceans. They are also predominantly confined to sites with a thick, insulating and late melting snow cover, protecting against winter and spring frost (Hultén 1968; Lid and Lid 2005).

Such a disjunction as found in the *Saxifraga rivularis* lineage and in *Luzula arcuata* and *Ranunculus glacialis*, could have originated via long-distance dispersal or vicariance. If originating from recent long-distance dispersal from Beringia to the amphi-Atlantic areas, we would expect little or no genetic and morphological divergence between the two ranges. We would also expect no phylogeographic structure within the amphi-Atlantic area and less variation than in Beringia as a result of a genetic bottleneck effect. In case of an origin from vicariance, however, we would expect genetic and morphological differentiation of the two groups, comparable amounts of variation within the two groups, and a phylogeographic pattern also within *S. rivularis* subsp. *rivularis*.

Our results mostly support the vicariance hypothesis. We found weak, but significant, genetic and morphological differentiation of subsp. *arctolitoralis* and subsp. *rivularis*, and regard it as an indication that the separation has been longer than post-glacial. A study of the phylogeography of *Saxifraga rivularis* subsp. *rivularis* also showed some differentiation among populations from Svalbard, Greenland, and the more southern areas in Norway, Iceland, and Scotland (Westergaard 2004). A detailed comparison of levels of variation and phylogeographic structure of subsp. *rivularis* with subsp. *arctolitoralis* was not possible with only one Beringian population included in the molecular analyses. In comparison with the *S. hyperborea* lineage, very little variation was found among populations of subsp. *rivularis* in this study (Fig. 3), which could be the result of a bottleneck effect following long distance dispersal, although other explanations such as autogamous reproduction (Brochmann and Håpnes 2001) are possible.

We want to put forward the hypothesis that these amphi-Atlantic/amphi-Beringian disjunctions were caused by vicariance and survival in separate refugia, probably during the last glaciation, and that the disjunctions subsequently have been retained due to the current continental climates in the major parts of the Arctic bordering on the frozen Polar Sea. An open Polar Sea caused a more oceanic climate along the coasts of arctic Canada and northern Siberia during the last interglacial (e.g., 120,000 BP; Frenzel et al. 1992; Miller et al. 1999; Koerner and Fisher 2002), allowing more continuous circumpolar distributions of species that avoid continental climates, needing snow protection in winter and moist conditions in summer. With a frozen Polar Sea during glaciations and in the current interglacial, more continental climate prevail(ed) in these areas, causing disjunctions and restricting populations to coastal regions with more oceanic climates in Beringia, in eastern North America south of the Laurentian ice sheet (e.g., in the Appalachian Ranges or in ice-free shelf areas), and in western Central Europe.

A very fast colonization of formerly glaciated areas from refugia is supported by the fossil record of *S. rivularis* subsp. *rivularis* from several places in Norway and Scotland already during the deglaciation phase and Allerød, respectively (Birks 1993; Huntley 1994). Colonization of arctic Canada and northern Siberia, however, was and is prevented by contemporary continental climates with dry and cold winters (100–300 mm precipitation p.a.; Moen 1998) causing unsuitable growing conditions due to scant snow cover and little protection against frost.

The phylogeographical analysis of *Saxifraga rivularis* represents the first case study that contributes molecular data to a historical biogeographical interpretation of the amphi-Beringian/amphi-Atlantic disjunctions. Phylogeographical studies in the other taxa listed above would be highly interesting in order to elucidate this disjunction phenomenon.

Delimitation and Classification of Taxa. Two of the species, the diploids *Saxifraga bracteata* and *S. debilis*, are very distinct, differing from all other taxa in several presumably independently inherited morphological characters. We lack genetic evidence for *S. debilis*, but the AFLP data obtained for *S. bracteata* show that this species is also genetically distinct. These two taxa fit the criteria used for the species rank in the Panarctic Flora and

Flora Nordica (Elven et al. 1999; Jonsell 2000). In addition to the different chromosome numbers, *S. hyperborea* and *S. rivularis* can only be unambiguously distinguished from each other by the absence/presence of rhizomes and the length of hairs. However, they are genetically very distinct and can also be distinguished by a combination of other morphological characters (Figs. 2, 3, 6; Tables 3, 4) and are therefore most appropriately recognized at the species level.

We regard the morphological and genetic differentiation between the Beringian and Atlantic tetraploids (Fig. 2C) as weak and insufficient to justify species rank. This conclusion also agrees with the view that the tetrapoids have a single origin. In accordance with the Panarctic Flora and Flora Nordica criteria, we therefore recognize these populations as geographic units at the subspecies level: *Saxifraga rivularis* subsp. *arctolitoralis* and subsp. *rivularis*.

The molecular variation we observed among the populations initially assigned to *'flexuosa'* and *'hyperborea'* did not correspond to the delimitation of the taxa as suggested by Elven et al. (2003). The weak genetic and phylogeographic differentiation found among these populations was not substantially reflected in morphology (STRUCTURE analyses; Figs. 2, 3, 7). Thus, our study does not even support recognition of *'flexuosa'* at an infraspecific level, which is in agreement with the hypothesis of Rebristaya and Yurtsev (1984) that *'flexuosa'* is only a shade-growing *Saxifraga hyperborea*. The most conspicuous morphological differences (length of flowering stem and pigmentation) can probably be attributed to phenotypic plasticity. The southern Rocky Mountain populations of *S. hyperborea*, however, were somewhat morphologically distinct (e.g., absence of hairs). It is possible that these populations should be recognized as a separate subspecies, but due to lack of molecular data from this area and the limited sampling between Alaska and the southern Rocky Mountains, a final conclusion cannot be made here.

Weber (1966, 1967) reported *Saxifraga rivularis* s. str. from the southern Rocky Mountains. His report was based on plants with larger pollen grains and longer stomata than in *S. debilis*, characters that typically distinguish a polyploid from a diploid. However, we included vouchers of both what he called *S. rivularis* (*00470761/Douglass 54–420*; Weber 1966) and what he called *S. hyperborea* subsp. *debilis* (*00470225/Weber 3437*; Weber 1966) in our morphological analyses; they both clearly belong to *S. debilis*. Neither could we identify any other specimens of *S. rivularis* s.str. in the southern Rocky Mountain collections from the Colorado herbarium. Löve et al. (1971), however, reported tetraploids ($2n = 52$) from this area as *S. rivularis* s. str. (Fig. 5). More research is needed to assess whether these tetraploids belong to *S. debilis* or to another unrecognized tetraploid taxon.

Taxonomic Treatment

Key to the Species. The values given are the 25% and 75% percentiles for the measured characters. The complete variation range of each character is found in the species descriptions. For combinations of taxa, the 25% percentile for the taxon with the lowest values, and the 75% percentile for the taxon with the highest values, were used.

1. Rhizomes absent. Growth usually densely tufted (occasionally loosely) 2
   2. Hypanthium V-shaped and sparsely covered with short glandular hairs (0.10–0.15 mm). Sepals narrow, 0.7–1.0 mm. 3–5 cauline leaves, the lowest similar to basal leaves . . . . . 2. *S. debilis*
   2. Hypanthium U-shaped and glabrous to densely covered with long glandular hairs (0.3–0.6 mm). Sepals wide, 1.5–2.1 mm. 1–3 cauline leaves, all smaller and simpler than basal leaves . . . . . 3. *S. hyperborea*
1. Rhizomes present. Growth usually loosely tufted (occasionally densely) 3
   3. Petals broadly elliptic, 1.8–3.1 mm wide. Basal leaves large, 13.1–19.9 mm wide and 7.4–11.2 mm long, with 5 or more angled lobes. Bracts similar to basal leaves, but with 3–7 lobes. Flowering stem 57–114 mm, with 3–6 subsessile flowers and 5–10 cauline leaves. Plant green . . . . . 1. *S. bracteata*
   3. Petals oblong, 1.1–1.7 mm wide. Basal leaves small, 5.9–9.6 mm wide and 3.7–5.2 mm long, with 3–5 rounded lobes. Bracts smaller and simpler than basal leaves, with 1–3 lobes. Flowering stem 17–70 mm, with 2–3 pedicellate flowers and 2–3 cauline leaves. Plant green to purple . . . . . 4. *S. rivularis*

Descriptions. The descriptions of the taxa are based on measurements of morphological characters and examination of 282 specimens (Table 1). Quantitative characters are given as the (0) 25–75 (100) percentiles of the investigated plants.

1. Saxifraga bracteata D.Don, Trans. Linn. Soc. 13: 367. 1822. —TYPE: ''In Siberia orientali''. *D. Merk 9* (holotype BM!). Probably described from Bering Island (Hultén 1945).

S. laurentiana Ser. ex DC., Prodr. IV: 35. 1830. Described from West Alaska: St. Lawrence Island (holotype, G-DC). *S. rivularis* var. *laurentiana* (Ser. ex DC.) Engl., Mon. Gatt. Sax.: 105. 1872.

S. vaginata Presl ex Sternb., Rev. Saxifr. Suppl. II: 39. 1831. —TYPE: U.S.A. Alaska: ''America

arctica: in insula Unalaschka''. *A. Chamisso* (lectotype, left hand plant, designated here by M. H. Jørgensen & R. Elven. PR-954!). The two plants on the sheet correspond with the illustrations of Sternberg (1831). We have designated the best developed and preserved plant.

Description. Perennial with loosely, or occasionally densely tufted growth. Whole plant green. Rhizomes present. Basal leaves large, (4.8) 7.4–11.2 (13.9) mm long and (9.5) 13.1–19.9 (32.0) mm wide, with 5–7 (11) angled lobes. Flowering stems (34) 57–114 (177) mm long, sparsely covered with uniseriate glandular hairs, with (1) 3–6 (29) subsessile flowers and (0) 5–10 (35) bracts. Bracts similar to basal leaves. Lower bracts (4.8) 7.2–11.4 (15.9) mm long and (3.8) 8.0–15.8 (33.6) mm wide, with (1) 3–7 (9) lobes. Upper bracts (2.3) 4.1–8.5 (14.7) mm long and (1.0) 4.2–8.8 (16) mm wide with 1–3 (7) lobes. Hypanthium U-shaped, in late flower and fruit stage (1.3) 2.1–3.7 (5.2) mm long, sparsely covered with uniseriate glandular hairs. Sepals (1.2) 1.5–2.3 (4.9) mm long and (1.8) 2.6–3.5 (5.3) mm wide. Petals white, broadly elliptic, (2.3) 3.7–5.4 (6.5) mm long and (1.2) 1.8–3.1 (3.9) mm wide. Gynoecium reaches (2.1) 3.4–4.2 (4.5) mm above the hypanthium in late flower and fruit stage. Hypanthium hairs (0.10) 0.20–0.50 (1.30) mm long with (2) 3–5 (8) cells. Hair partition-walls weakly purple-colored.

Chromosome Number. $2n = 26$ (Zhukova and Tikhonova 1973, Russian Far East, Chukotka, two counts; Zhukova 1982, Russian Far East, Chukotka). Diploid in flow cytometry (this study, U.S.A., Alaska, Seward Peninsula, five measures).

Habitat. Coastal cliffs, often in or immediately above the spray zone. Bird-manured slopes and cliffs. Margins of brackish gravel flats.

Distribution. Amphi-Pacific. Coastal areas of Alaska from Seward Peninsula south to the Aleutian Islands including islands of the Bering Sea, and Russian Far East; South and East Chukotka, Kamchatka, the Commodore Islands, the Kuril Islands and Sakhalin.

2. SAXIFRAGA DEBILIS Engelm. ex A.Gray, Proc. Acad. Nat. Sci. Philadelphia 15: 62. 1864. *S. cernua* var. *debilis* (Engelm. ex A.Gray) Engl., Mon. Gatt. Sax.: 170. 1872. *S. hyperborea* subsp. *debilis* (Engelm. ex A.Gray) Á.Löve, D.Löve & Kapoor, Arctic Alpine Res. 3: 151. 1971. *S. rivularis* var. *debilis* (Engelm. ex A.Gray) Dorn, Vasc. Pl. Wyoming: 300. 1988. —TYPE: U.S.A. Rocky Mountain alpine flora, lat. 39°–41°, Colorado. 1862. *E. Hall & J. P. Harbour 198* (holotype: GH).

Description. Perennial forming compact or occasionally loose tufts. Whole plant green. Rhizomes absent. Basal leaves (3.1) 4.5–6.7 (10.3) mm long and (2.8) 7.3–10.6 (15.4) mm wide, with 5–7 partly angled lobes. Flowering stems (31) 67–93 (194) mm long, sparsely covered with uniseriate glandular hairs, with (1) 2–3 (5) subsessile flowers and (2) 3–5 (8) cauline leaves. Cauline leaves beneath inflorescence similar to basal leaves. Bracts smaller and simpler than basal leaves. Lower bracts (2.5) 3.6–5.7 (8.2) mm long and (0.9) 1.9–6.5 (9.8) mm wide, with 1–3 (6) lobes. Upper bracts (2.1) 3.5–5.4 (6.8) mm long and (0.3) 0.7–2.3 (7.1) mm wide, with 1 (5) lobes. Hypanthium V-shaped, in late flower and fruit stage (1.0) 2.1–3.0 (4.0) mm long, glabrous to sparsely covered with uniseriate glandular hairs. Sepals narrow, (1.3) 1.8–2.0 (2.9) mm long and (0.6) 0.7–1.0 (1.2) mm wide. Petals white to pale purple, oblong, (1.7) 2.9–4.4 (6.2) mm long and (0.6) 0.9–1.2 (4.2) mm wide. Gynoecium reaches (1.8) 2.4–3.3 (3.8) mm above the hypanthium in late flower and fruit stage. Hypanthium hairs (0.05) 0.10–0.15 (0.30) mm long with (2) 3–4 (5) cells. Hair partition-walls colorless to weakly purple-colored.

Chromosome Number. $2n = 26$ (Löve et al. 1971, U.S.A., Rocky Mountains, Colorado).

Habitat. Alpine meadows in shade of boulders and rocks, open gravel and silt, seepage areas, brook and lake margins, snowbeds, shady taluses and ravines, cliffs in shade, rock crevices.

Distribution. Cordilleran. Rocky Mountains in Colorado, Wyoming, Utah and Montana.

3. SAXIFRAGA HYPERBOREA R.Br., Chlor. Melvill.: 16. 1823. *S. rivularis* var. *hyperborea* (R.Br.) Hook., Fl. Bor.-Amer. 1: 246. 1834. *S. rivularis* subsp. *hyperborea* (R.Br.) Dorn, Vasc. Pl. Wyoming: 300. 1988. —TYPE: CANADA. Melville Island. *Saxifraga hyperborea* Flor. Melv. *J. Ross 25* (holotype: BM!).

S. petiolaris R.Br. in John Ross nom. nud., Explor. Baffin's Bay. ii: 192, 1819, nom. nud. Reported from East Greenland and/or North Baffin Island.

S. flexuosa Sternb., Rev. Saxifrag. Suppl. II: 38. 1831. *S. rivularis* var. *flexuosa* (Sternb.) Engl. & Irmsch. in Engler, Pflanzenr. 4, Fam. 117 1: 282, fig. 64. 1916. *S. rivularis* subsp. *flexuosa* (Sternb.) Gjærev., Kongel. Norske Vidensk. Selsk. Skr. 4: 61. 1963. —TYPE: RUSSIA. Russian Far East: East Chukotka: Lavrentiy Bay (lectotype, left hand plant, designated here by M. H. Jørgensen & R. Elven. PR-463!). The sheet is annotated as *Saxifraga flexuosa* Sternb. in Sternberg's handwriting. The left hand plant

is the original of the illustration in Sternberg (1831).

S. rivularis var. *purpurascens* Lange, Consp. Fl. Groenland.: 62. 1880. —TYPE: GREENLAND. Groenlandia borealis: Disco, Quannesvit, 22.06.1871. *Th. M. Fries s.n.* (lectotype, designated here by M. H. Jørgensen & R. Elven. C!). Lange (1880) referred to two specimens, both in C. The Fries specimen designated here conforms with the diagnosis in all aspects. The other specimen (leg. Rink) deviates in, e.g., pigmentation. Neither specimen was annotated as var. *purpurascens* by Lange who only annotated post-1880 specimens as such.

Description. Perennial forming compact or occasionally loose tufts. Pigmentation varies from whole plant green to whole plant purple pigmented, but mostly the latter. Rhizomes absent. Basal leaves small, (2.0) 3.4–4.8 (9.0) mm long and (2.9) 5.2–7.6 (14.4) mm wide, with (2) 3–5 (7) rounded lobes. Flowering stems (11) 27–64 (125) mm long, sparsely to densely covered with uniseriate glandular hairs, with (1) 2–3 (6) pedunculate flowers and 1–3 (7) bracts. Bracts smaller and simpler than basal leaves. Lower bracts (3.0) 4.4–6.9 (12.5) mm long and (0.5) 2.1–6.2 (9.2) mm wide with 1–3 (5) lobes. Upper bracts (1.2) 3.5–6.0 (9.5) mm long and (0.4) 1.0–2.2 (6.7) mm wide, with 1 (5) lobes. Hypanthium U-shaped, in late flower and fruit stage (0.9) 1.5–2.2 (3.8) mm long, sparsely to densely covered with uniseriate glandular hairs. Sepals (0.7) 1.0–1.6 (2.8) mm long and (0.7) 1.5–2.1 (3.0) mm wide. Petals white to purple, oblong, (1.6) 2.6–3.4 (4.5) mm long and (0.4) 0.9–1.3 (2.4) mm wide. Gynoecium reaches (1.6) 2.6–3.5 (5.2) mm above the hypanthium in late flower and fruit stage. Hypanthium hairs (0.10) 0.30–0.60 (1.50) mm long with (2) 4–6 (10) cells. Hair partition-walls purple-colored.

Chromosome Number. $2n$ = (23) 26 (38) (Flovik 1940, Svalbard, as *S. rivularis*; Holmen 1952, Greenland, Peary Land; Jørgensen et al. 1958, Greenland; Packer in Löve and Löve 1961; Johnson and Packer 1968, U.S.A., Alaska, Point Hope; Mulligan and Porsild 1968, Canada, Central Yukon Plateau as *S. rivularis* var. *flexuosa*; Zhukova 1968, Russian Far East, Chukotka, two counts; Zhukova 1969, Russian Far East, Chukotka; Zhukova and Petrovsky 1971, Russian Far East, Wrangel Island; Zhukova and Tikhonova 1971, Russian Far East, Chukotka; Zhukova and Petrovsky 1972, Russian Far East, Wrangel Island, two counts; Zhukova et al. 1973, Russian Far East, Wrangel Island, two counts, and Chukotka, two counts; Zhukova and Petrovsky 1977, Russian Far East, Chukotka; Engelskjøn 1979, Svalbard, two counts; Zhukova 1980, Russian Far East, South Chukotka, three counts; Zhukova and Petrovsky 1980, Russian Far East, Anyui Mountains, two counts; Petrovsky and Zhukova 1981, Russian Far East, Wrangel Island; Yurtsev and Zhukova 1982, Siberia, northern Yakutia; Borgen and Elven 1983, Svalbard, four counts; Zhukova and Petrovsky 1987, Russian Far East, Chukotka and Siberia, Yamal, 13 counts; Devyatov et al. 1997, Siberia, Yamal, two counts, and Novosibirskiye Islands; $2n$ = (23) 26 (38), Guldahl et al. 2005, Svalbard). Diploid in flow cytometry (Guldahl et al. 2005, Greenland, Liverpool Land, five measures, Svalbard, Canada, Northwest Territory, six measures; this study, Canada, Yukon, Northwest Territory, and Nunavut, six measures, U.S.A., Alaska, Seward Peninsula, 11 measures, and Valdez area, Greenland, Kitaa, Siberia, Taymyr, six counts, Yamal).

Diploid chromosome counts from White Mountains, New Hampshire (Löve and Löve in Löve and Solbrig 1964; Löve and Löve 1966) published under the name *S. hyperborea* need confirmation by voucher.

Habitat. Damp tundra, open gravel and silt, brook and lake margins, snowbeds, shady ravines, cliffs in shade. Silty and gravelly seashores.

Distribution. Circumpolar, mainly arctic. Greenland, arctic Canada south to northern Labrador and southern Hudson Bay, Rocky Mountains south to Colorado, Alaska, Russian Far East; Kamchatka, the Kuril Islands, Sakhalin, Stanovoye Mountains, Russian arctic coast from Chukotka to Ural Mountains, Novaya Zemlya, Zemlya Frantsa Josifa and Svalbard. *Saxifraga hyperborea* is reported from New Hampshire, USA, by Löve and Löve (Löve and Löve in Löve and Solbrig 1964; Löve and Löve 1966), but the determinations need confirmation.

4. SAXIFRAGA RIVULARIS L., Sp. Pl.: 404. 1753. *Lobaria rivularis* (L.) Haw., Enum. Saxifr.: 19. 1821. —TYPE: SWEDEN. Linnaeus, Flora Lapponica: t. 2, f. 7 (1737) (lectotype; designated by Jonsell and Jarvis 2002 p. 73. LAPP).

Description. Perennial with loosely or occasionally densely tufted growth. Rhizomes present. Basal leaves of medium size, (2.6) 3.7–5.1 (7.4) mm long and (3.0) 6.3–8.4 (12.6) mm wide, with (2) 5 somewhat angled lobes. Flowering stems (8) 20–56 (121) mm long with (1) 2 (5) pedunculate flowers and (0) 2–3 (8) bracts. Bracts smaller and simpler than basal leaves. Lower bracts (3.1) 3.8–6.4 (13.0) mm long and (1.2) 3.2–6.9 (8.3) mm wide, with 1–3 (5) lobes. Upper bracts (2.5) 3.7–6.8 (10.5) mm long and (0.6) 1.7–3.7 (9.4) mm wide, with 1–2 (5) lobes. Hypanthium U-shaped, in late flower

and fruit stage (1.2) 1.8–2.8 (5.0) mm long. Sepals (0.9) 1.2–1.6 (2.1) mm long and (1.0) 2.0–2.6 (4.0) mm wide. Petals white to purple, oblong to elliptic, (1.4) 3.0–4.2 (5.5) mm long and (0.6) 1.1–1.7 (2.7) mm wide. Gynoecium reaches (2.3) 3.3–4.2 (6.0) mm above the hypanthium in late flower and fruit stage.

Key to the Subspecies

1. Hypanthium sparsely covered by short glandular hairs (0.15–0.25 mm) with non-colored or weakly-colored partition-walls. The whole plant green or with some purple pigmentation in the inflorescence. Flowering stem long, 27–70 mm, glabrous or sparsely hairy . . . . . subsp. *rivularis*
1. Hypanthium densely covered by long glandular hairs (0.30–0.60 mm) with purple partition-walls. The whole plant or at least the inflorescence mostly purple-pigmented. Flowering stem short, 17–30 mm, sparsely to densely hairy . . . . . subsp. *arctolitoralis*

Saxifraga rivularis subsp. rivularis

Description. Pigmentation varies from whole plant green to purple pigmented only in the inflorescence. Stem glabrous to sparsely covered with uniseriate glandular hairs. Hypanthium U-shaped, and sparsely covered with uniseriate glandular hairs. Hypanthium hairs (0.10) 0.15–0.25 (0.40) mm long with 3–4 (5) cells. Hair partition-walls without color or (rarely) pale purple.

Chromosome Number. $2n$ = (26, 43, 47, 50) 52 (56, 85, 95) (Böcher 1938, Greenland, Jacobsen Fjord, $2n$ = 56; Sørensen and Westergaard in Löve and Löve 1948, Greenland; Löve and Löve 1951, Iceland; Jørgensen et al. 1958, Greenland; Engelskjøn and Knaben 1971, Norway, $2n$ = 50–53, four counts; Engelskjøn 1979, Bear Island, three counts; Löve and Löve in Löve 1982, Canada, Manitoba; Borgen and Elven 1983, Svalbard, three counts, and Norway, five counts; $2n$ = (26, 43, 47) 50–52 (85, 95), Guldahl et al. 2005, Svalbard). Tetraploid in flow cytometry (Guldahl et al. 2005, Svalbard, 28 measures, Greenland, Liverpool Land, 10 measures, and Tunu, 11 measures, Iceland, Vesturland, four measures; this study, Canada, Nunavut, two measures, Greenland, Kitaa, two measures, and Tunu, four measures, Iceland, seven measures, Norway, nine measures, Svalbard, 18 measures, Jan Mayen).

Habitat. Snowbeds, damp tundra, bird-manured cliffs, springs, seepage slopes, brook margins. Silty and gravelly seashores.

Distribution. Amphi-Atlantic. The Russian arctic coast from West Taymyr to Kola Peninsula, Novaya Zemlya, Zemlya Frantsa Josifa, Svalbard, Bear Island, Jan Mayen Island, Norway, Scotland, Faroe Islands, Iceland, Greenland, the Atlantic coast of Canada from Baffin Island south to Gaspé Peninsula, White Mountains in New Hampshire.

Saxifraga rivularis subsp. arctolitoralis (Jurtz. & V.V.Petrovsky) Jørgensen & Elven comb. et stat. nov. Basionym: *S. arctolitoralis* Jurtz. & V.V.Petrovsky, Bot. Zhurn. 66(7): 1045. 1981. —TYPE: Russia. Russian Far East, Chukchi Peninsula. ''Paeninsula Tschuktschorum pars australi-orientalis, litus boreali-occidentalis freti Senjavini, inter ostia fl. Chutenreczchen-veem et fl. Kurgyveem'' 19.08.1978. *B. A. Yurtsev S–78–1* (holotype: LE!).

Description. Whole plant mostly purple pigmented. Stem and U-shaped hypanthium sparsely to densely covered with uniseriate glandular hairs. Hypanthium hairs (0.20) 0.30–0.60 (1.10) mm long with 3–4 (8) cells. Hair partition-walls purple.

Chromosome Number. $2n$ = (48) 52 (Zhukova 1968, Russian Far East, Wrangel Island as *S. rivularis*; Zhukova and Tikhonova 1971, Russian Far East, Chukotka as *S. rivularis*; Zhukova et al. 1973, Russian Far East, Wrangel Island, 2n = 48 as *S. rivularis*; Zhukova and Tikhonova 1973, Russian Far East, Chukotka as *S. rivularis*; Packer and McPherson 1974, U.S.A., Alaska, Barrow as *S. rivularis*; Zhukova and Petrovsky 1977, Russian Far East, Chukotka as *S. rivularis*; Zhukova and Petrovsky 1987, Russian Far East, Wrangel Island, two counts, and Russian Far East, Chukotka three counts as *S. arctolitoralis*).

Habitat. Arctic seashores on silt and clay, sloping soil banks.

Distribution. Amphi-Beringian. The arctic Pacific coast of Alaska from Barrow south to Seward Peninsula, the eastern and northern coast of Chukotka and Wrangel Island.

Acknowledgements. We are grateful to all the field workers who collected material for us. Gerard Geenen is acknowledged for performing the flow cytometry analyses, Wenche Johansen for finding all the 19th Century literature, Dorothee Ehrich for programming the Gower index, Volodya Razzhivin for his help at the Komarov institute, and Dave F. Murray for his thorough reading and commenting of the manuscript. We are also grateful for the thorough and constructive comments from the reviewers and editors of Systematic Botany. The work with this manuscript was supported by the Strategic University Program 146515/420 to C. Brochmann under the botany subprogram *Migration and evolution of arctic plants in response to Quaternary climate*

changes. The program is jointly financed by the Research Council of Norway and the Natural History Museum, University of Oslo. Some of the material was obtained by Reidar Elven's participation in the Tundra North-West (TNW) '99 expedition, funded by the Polar Research Secretariate at the Royal Swedish Academy of Sciences.

Literature Cited

Abbott, R. J. and C. Brochmann. 2003. History and evolution of the arctic flora: in the footsteps of Eric Hultén. *Molecular Ecology* 12: 299–313.

——— and H. P. Comes. 2004. Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (Purple saxifrage). *New Phytologist* 161: 211–224.

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. *Molecular Ecology* 14: 2739–2753.

Benham, J. J. 2001. *Genographer*. 1.6.0. Montana State University: Montana.

Birks, H. H. 1993. The importance of plant macrofossils in late-glacial climatic reconstructions: an example from Western Norway. *Quaternary Science Reviews* 12: 719–726.

Böcher, T. W. 1938. Zur Zytologie einiger arktischen und borealen Blütenpflanzen. *Svensk Botanisk Tidsskrift* 32: 346–361.

Borgen, L. and R. Elven. 1983. Chromosome numbers of flowering plants from Northern Norway and Svalbard. *Nordic Journal of Botany* 3: 301–306.

Brochmann, C. and A. Håpnes. 2001. Reproductive strategies in some arctic *Saxifraga* (Saxifragaceae), with emphasis on the narrow endemic *S. svalbardensis* and its parental species. *Botanical Journal of the Linnean Society* 137: 31–49.

———, A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. *Botanical Journal of the Linnean Society* 82: 521–536.

———, Q.-Y. Xiang, S. J. Brunsfeld, D. E. Soltis, and P. S. Soltis. 1998. Molecular evidence for polyploid origins in *Saxifraga* (Saxifragaceae): the narrow arctic endemic *S. svalbardensis* and its widespread allies. *American Journal of Botany* 85: 135–143.

Brown, R. 1819. Appendix no. VI. List of plants in *A voyage of discovery, made under the orders of the admiralty, in His Majesty's ships, for the purpose of exploring Baffin's Bay, and enquiring into the probability of a North-West Passage*, ed. J. Ross. London: Longman, Hurst, Rees, Orme and Brown.

———. 1823. *Chloris Melvilliana. A list of plants collected in Melville Island in the year 1820 by the officers of the voyage of discovery under the orders of Captain Parry*. London: John Murray.

Conti, E., D. E. Soltis, T. M. Hardig, and J. Schneider. 1999. Phylogenetic relationships of the silver saxifrages (*Saxifraga, sect. Ligulatae Haworth*): implications for the evolution of substrate specificity, life histories, and biogeography. *Molecular Phylogenetics and Evolution* 13: 536–555.

Cronquist, A., N. H. Holmgren, and P. K. Holmgren. 1997. *Intermountain flora. Vascular plants of the Intermountain West, U.S.A. Volume three, part A. Subclass Rosidae (except Fabales)*. Bronx: The New York Botanical Garden.

Devyatov, A. G., P. Y. Zhmylev, and A. D. Kozhevnikova. 1997. Chromosome numbers of some arctic species of the genus *Saxifraga* (*Saxifragaceae*). *Botanicheskiy Zhurnal* 82(9): 122. [In Russian].

Dice, L. R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26: 297–302.

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). *Molecular Ecology* 13: 349–370.

Don, D. 1822. A monograph of the genus Saxifraga. *The Transactions of the Linnean Society of London* 13: 341–452.

Dorn, R. D. 1988. *Vascular plants of Wyoming*. Cheyenne: Mountain West Publishing.

Elven, R., B. Jonsell, D. F. Murray, and B. A. Yurtsev. 1999. An operational species concept for the Panarctic Flora. *Det Norske Vitenskaps-Akademi I Matematisk Naturvitenskapelig Klasse Skrifter Ny serie* 38: 23–32.

———, D. F. Murray, V. Y. Razzhivin, and B. A. Yurtsev. 2003. *Panarctic flora checklist. Vascular plants. Draft version*. Oslo: National Centre for Biosystematics, University of Oslo.

Engelskjøn, T. 1979. Chromosome number in vascular plants from Norway, including Svalbard. *Opera Botanica* 52: 1–38.

——— and G. Knaben. 1971. Chromosome numbers of Scandinavian arctic-alpine plant species III. *Acta Borealia A. Scientia* 28: 1–30.

Engler, A. 1872. *Monographie der Gattung* Saxifraga *L. mit besonderer Berücksichtigung der geographischen Verhältnisse*. Breslau: J. U. Kern's Verlag.

——— and E. Irmscher. 1916. *Das Pflanzenreich IV. 117 Saxifragaceae* -Saxifraga. Leipzig: Wilhelm Engelmann.

Flovik, K. 1940. Chromosome numbers and polyploidy within the flora of Spitzbergen. *Hereditas* 26: 430–440.

Frenzel, B., M. Pécsi, and A. A. Velichko Eds. 1992. *Atlas of paleoclimates and paleoenvironments of the northern hemisphere. Late Pleistocene – Holocene*. Budapest: Geographical Research Institute, Hungarian Academy of Sciences; Gustav Fischer Verlag.

Gjærevoll, O. 1963. Botanical investigations in Central Alaska, especially in White Mts. Part II Dicotyledones Salicaceae – Umbelliferae. *Det Kgl Norske Videnskabers Selskabs Skrifter* 4: 1–116.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27: 857–874.

Gray, A. 1863. Enumeration of the species of plants collected by Dr. C. C. Parry, and Messrs. Elihu Hall and J. P. Harbour, during the Summer and Autumn of 1862, on and near the Rocky Mountains, in Colorado Territory, lat. 39°–41°. *Proceedings of the Academy of Natural Sciences of Philadelphia* 2: 55–80.

Guldahl, A. S., T. M. Gabrielsen, A.-C. Scheen, L. Borgen, S. Spjelkavik, S. W. Steen, and C. Brochmann. 2005. The *Saxifraga rivularis/hyperborea* species complex in Svalbard: a molecular-morphological approach. *Flora* 200: 207–221.

Harrington, H. D. 1954. *Manual of the plants of Colorado*. Denver: Sage Books.

Holmen, K. 1952. Cytological studies in the flora of Peary Land, North Greenland. *Meddelelser om Grønland* 128: 1–40.

Holmgren, P. K. and N. H. Holmgren. 1998 onwards. *Index Herbariorum*. New York Botanical Garden. http://sciweb.nybg.org/science2/IndexHerbariorum.asp

Hooker, W. J. 1834. *Flora boreali-americana*. London: Henry G. Bohn.

Hultén, E. 1937. *Outline of the history of arctic and boreal biota during the Quaternary period*. Stockholm: Thule Förlag.

———. 1945. Flora of Alaska and Yukon V. Dicotyledoneae. Rhoedales, Sarraceniales, Rosales I (*Crassulaceae*, *Saxi-*

fragaceae). *Lunds Universitets Årsskrift Ny Följd Andra Avdelningen* 41: 900–942.

———. 1968. *Flora of Alaska and neighboring territories. A manual of the vascular plants*. Stanford: Stanford University Press.

——— and M. FRIES. 1986. *Atlas of North European vascular plants north of the Tropic of Cancer I–III*. Königstein: Koeltz Scientific Books.

HUNTLEY, B. 1994. Late Devensian and Holocene paleoecology and paleoenvironments of the Morrone Birkwoods, Aberdeenshire, Scotland. *Journal of Quaternary Science* 9: 311–336.

JINTANG, P., R. J. GORNALL, and H. OHBA. 2001. *Saxifraga* L. in *Flora of China. Vol. 8 (Brassicaceae through Saxifragaceae)*, eds. Wu, Z. and P. H. Raven. Beijing & St. Louis: Science Press & Missouri Botanical Garden Press.

JOHNSON, A. W. and J. G. PACKER. 1968. Chromosome numbers in the flora of Ogotoruk Creek, N.W. Alaska. *Botaniska Notiser* 121: 403–456.

JONSELL, B., Ed. 2000. *Flora Nordica 1. Lycopodiaceae – Polygonaceae*. Stockholm: The Bergius Foundation, The Royal Swedish Academy of Sciences.

——— and C. E. JARVIS. 2002. Lectotypification of Linnean names for Flora Nordica (Brassicaceae - Apiaceae). *Nordic Journal of Botany* 22: 67–86.

JØRGENSEN, C. A., T. SØRENSEN, and M. WESTERGAARD. 1958. The flowering plants of Greenland. A taxonomical and cytological survey. *Det Kongelige Danske Videnskabernes Selskab Biologiske Skrifter* 9: 1–172.

KOERNER, R. M. and D. A. FISHER. 2002. Ice-core evidence for widespread Arctic glacier retreat in the Last Interglacial and the early Holocene. *Annals of Glaciology* 35: 19–24.

LANGE, J. 1880. Conspectus florae groenlandicae. *Meddelelser om Grønland* 3.

LEVIN, D. A. 2002. *The role of chromosomal change in plant evolution*. New York: Oxford University Press, Inc.

LID, J. and D. T. LID. 2005. *Norsk flora*. 7th ed. by ELVEN, R. Oslo: Det Norske Samlaget.

LINNAEUS, C. 1753. *Species plantarum: exhibentes plantas rite cognitas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas*. Holmiae: Impensis Laurentii Salvii.

LIPKIN, R. and D. F. MURRAY. 1997. *Alaska rare plant field guide*. Washington, D.C.: U. S. Department of the Interior, Office of Equal Opportunity.

LÖVE, Á. 1982. IOPB Chromosome number reports LXXV. *Taxon* 31: 342–368.

——— and D. LÖVE. 1948. *Chromosome numbers of northern plant species*. Reykjavík: Ingólfsprent.

——— and ———. 1951. Studies on the origin of the Icelandic flora. II. Saxifragaceae. *Svensk Botanisk Tidsskrift* 45: 368–399.

——— and ———. 1961. Chromosome numbers of Central and Northwest European plant species. *Opera Botanica* 5: 1–581.

——— and ———. 1966. Cytotaxonomy of the alpine vascular plants of Mount Washington. *University of Colorado Studies, Series in Biology* 24: 1–74.

——— and O. T. SOLBRIG. 1964. IOPB chromosome number reports II. *Taxon* 13: 201–209.

———, D. LÖVE, and B. M. KAPOOR. 1971. Cytotaxonomy of a century of Rocky Mountain orophytes. *Arctic and Alpine Research* 3: 139–165.

MILLER, G. H., W. N. MODE, A. P. WOLFE, P. E. SAUER, O. BENNIKE, S. L. FORMAN, S. K. SHORT, and T. W. STAFFORD. 1999. Stratified interglacial lacustrine sediments from Baffin Island, Arctic Canada: chronology and palaeoenvironmental implications. *Quaternary Science Reviews* 18: 789–810.

MOEN, A. 1998. *Nasjonalatlas for Norge: Vegetasjon*. Hønefoss: Statens kartverk.

MULLIGAN, G. A. and A. E. PORSILD. 1968. Chromosome numbers of some plants from the unglaciated central Yukon plateau, Canada. *Canadian Journal of Botany* 47: 655–662.

NEI, M. and W. H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* 76: 5269–5273.

PACKER, J. G. and G. D. MCPHERSON. 1974. Chromosome numbers in some vascular plants from northern Alaska. *Canadian Journal of Botany* 52: 1095–1099.

PETROVSKY, V. V. and P. G. ZHUKOVA. 1981. Chromosome numbers and taxonomy of some plant species of Wrangel Island. *Botanicheskiy Zhurnal* 66: 380–387. [In Russian].

POLUNIN, N. 1959. *Circumpolar arctic flora*. London: Oxford University Press.

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

REBRISTAYA, O. V. and B. A. YURTSEV. 1984. *Saxifraga* L. Pp. 25–88 in *Arkticheskaya flora SSSR. IX. Droseraceae-Leguminosae. 1. Droseraceae-Rosaceae*, ed. Yurtsev, B. A. Leningrad: Izdatelstvo ''Nauka''. [In Russian].

ROHLF, F. 1999. *NTSYS-pc. Numerical taxonomy and multivariate analysis system*. 2.02. New York: Exeter Software.

ROSENDAHL, C. O. 1905. Die nordamerikanischen Saxifraginae und ihre Verwandtschafts-Verhältnisse in Beziehung zu ihrer geographischen Verbreitung. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 37. Beibl. 83: 1–87.

SOLTIS, D. E., R. K. KUZOFF, E. CONTI, R. GORNALL, and K. FERGUSON. 1996. *matK* and *rbcL* gene sequence data indicate that *Saxifraga* (Saxifragaceae) is polyphyletic. *American Journal of Botany* 83: 371–382.

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

STERNBERG, K. M. 1831. *Revisio Saxifragarum, Supplementum II*. Prague.

STOREY, J. D. and R. TIBSHIRANI. 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* 100: 9440–9445.

SWOFFORD, D. L. 2001. PAUP*. Phylogenetic analysis using parsimony (* and other methods) 4.0b10. Sunderland: Sinauer Associates.

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

VAN DE PEER, Y. and R. DE WACHTER. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows Environment. *Computer Applications in the Biosciences* 10: 569–570.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VANDELEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, and M. ZABEAU. 1995. Aflp - a new technique for DNA-fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

WEBB, D. A. and R. J. GORNALL. 1989. *A manual of saxifrages and their cultivation*. Portland: Timber Press.

WEBER, W. A. 1966. Additions to the flora of Colorado - IV. *University of Colorado Studies, Series in Biology* 23: 1–24.

———. 1967. *Rocky Mountain flora*. Boulder: University of Colorado Press.

——— and R. C. WITTMANN. 2001. *Colorado flora: eastern slope.* 3th ed. Boulder: University Press of Colorado.

WESTERGAARD, K. 2004. *Phylogeography of the high arctic Saxifraga rivularis L. inferred by amplified fragment length polymorphisms (AFLPs)*. Candidata scientiarum thesis. Tromsø: University of Tromsø.

YURTSEV, B. A. 1981. Seven new taxa of flowering plants from northeastern Asia and adjacent territories. *Botanicheskiy Zhurnal* 66: 1041–1046. [In Russian].

——— and P. G. ZHUKOVA. 1982. Chromosome numbers of some plants of the northeastern Yakutia (the drainage of the Indigirka River in its middle reaches). *Botanicheskiy Zhurnal* 67: 778–787. [In Russian].

ZHMYLEV, P. Y. 1997. Systematical review of rockfoils (*Saxifraga* L.) of Russia and contiguous territories: subgenera *Porphyrion* and *Saxifraga*. *Bulletin Moskovskogo obscestva ispytateley prirody, Otdel biologiceskiy* 102: 42–48. [In Russian].

ZHUKOVA, P. G. 1968. Chromosome numbers in some plant species from the North-East of the USSR III. *Botanicheskiy Zhurnal* 53: 365–368. [In Russian].

———. 1969. Chromosome number in certain plant species indigenous to the North-East of the USSR IV. *Botanicheskiy Zhurnal* 54: 1985–1990. [In Russian].

———. 1980. Chromosome number of some Southern Chukotka plant species. *Botanicheskiy Zhurnal* 65: 51–59. [In Russian].

———. 1982. Chromosome numbers of some plant species of north-eastern Asia. *Botanicheskiy Zhurnal* 67: 360–365. [In Russian].

——— and V. V. PETROVSKY. 1971. Chromosome numbers of certain flowering plants of the Wrangel Island. *Botanicheskiy Zhurnal* 56: 294–305. [In Russian].

——— and ———. 1972. Chromosome numbers of some flowering plants of the Wrangel Island. II. *Botanicheskiy Zhurnal* 57: 554–563. [In Russian].

——— and ———. 1977. Chromosome numbers of some western Chukotka plant species. III. *Botanicheskiy Zhurnal* 62: 1215–1223. [In Russian].

——— and ———. 1980. Chromosome numbers and taxonomy of some species of the Anyui Mts. *Botanicheskiy Zhurnal* 65: 651–659. [In Russian].

——— and ———. 1987. Karyotaxonomic study of some species of the genus *Saxifraga* (Saxifragaceae) from northern Asia. *Botanicheskiy Zhurnal* 72: 632–640. [In Russian].

——— and A. D. TIKHONOVA. 1971. Chromosome numbers of certain plant species indigenous to the Chukotsky province. *Botanicheskiy Zhurnal* 56: 868–875. [In Russian].

——— and ———. 1973. The chromosome numbers of some Chukotka plant species. II. *Botanicheskiy Zhurnal* 58: 395–402. [In Russian].

———, V. V. PETROVSKY, and T. V. PLIEVA. 1973. The chromosome numbers and taxonomy of some plant species from Siberia and Far East. *Botanicheskiy Zhurnal* 58: 1331–1342. [In Russian].

II

III

SPECIATION

Microsatellites resolve the taxonomy of the polyploid *Cardamine digitata* aggregate (Brassicaceae)

Marte Holten Jørgensen[1], Tor Carlsen[2], Inger Skrede[2] & Reidar Elven[2]

[1] *Department of Biology, University of Oslo, P.O. Box 1066 Blindern, 0316 Oslo, Norway. martej@ulrik.uio.no (author for correspondence)*
[2] *National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, 0318 Oslo, Norway*

Polyploidy, probably the single most important mode of sympatric speciation in plants, tends to result in complicated evolutionary patterns. The *Cardamine digitata* aggregate is a species complex where polyploidy has resulted in taxonomic and nomenclatural controversies. Two basic chromosome numbers are found (x = 7 and x = 8), and all plants studied so far are tetra- to dodecaploids. We used six microsatellite loci originally developed for the *Arabidopsis* genome to identify evolutionary and taxonomic units within the *C. digitata* aggregate, obtaining 102 polymorphic markers. Using different analysis methods (PCO, CVA, STRUCTURE, parsimony), we recognised four approximately equidistant units corresponding in morphology with the four described species: *C. blaisdellii*, *C. digitata*, *C. microphylla*, and *C. purpurea*. All taxa include at least two ploidal levels; thus recurrent taxonomic autopolyploidy is indicated.

KEYWORDS: Arctic, *Cardamine*, microsatellites, nomenclature, polyploidy, taxonomy

INTRODUCTION

Polyploidy may be the single most important mode of sympatric speciation in the plant kingdom (Otto & Whitton, 2000). With consequences such as rapid genomic rearrangements, genomic downsizing, movement of genetic elements across genomes, and movement of foreign genetic material into the polyploid genome (reviewed by Soltis & al., 2003), polyploidy is an evolutionary trigger. Evidence of these genomic processes induced by polyploidy is mostly collected from studies on crop plants, and the direct consequences of polyploidisation in natural populations are still essentially unknown (Soltis & al., 2003). Identifying polyploid units and their origins is a necessary prerequisite before studies of these consequences can begin.

Cardamine L. is a nearly cosmopolitan genus with 160–200 species (Sjöstedt, 1975; Hewson, 1982; Al-Shehbaz, 1988; Webb & al., 1988; Al-Shehbaz & al., 2006; Lihová & Marhold, 2006). The taxonomy of many species complexes is unexplored and remains controversial. An exception is the *C. pratensis* group and many of its close relatives, which has been studied extensively using nuclear and plastid sequences and fingerprinting methods such as amplified fragment length polymorphism (AFLP; Franzke & Hurka, 2000; Marhold & al., 2002, 2004; Lihová & al., 2004; Marhold & Lihová, 2006).

Schulz (1903, 1936) considered *Cardamine* sect. *Cardaminella* Prantl to be one of the main sections in the genus *Cardamine*. However, a recent study of the phylogeny of *Cardamine* has found support for the long-time suspicion that *C.* sect. *Cardaminella* is polyphyletic (Carlsen & al., in press). The circumpolar and alpine *C. bellidifolia* L. and some of its European alpine relatives constitute a separate and distinct branch, whereas other species of the section appear in several parts of the tree. There is, however, a consistent and monophyletic Pacific–Beringian branch of *C.* sect. *Cardaminella*. In this branch, we find the *C. digitata* Richardson aggregate and the Asian Beringian *C. victoris* N. Busch and *C. sphenophylla* Jurtzev, see Petrovsky in Tolmachev (1975). From morphological evidence *C. sphenophylla* and *C. victoris* appear as two distinct species but are probably more closely related to each other than to their next closest relative, the *C. digitata* aggregate. The nomenclature and circumscription of some of the species in the *C. digitata* aggregate have been disputed (see discussion for details).

Most species of *Cardamine* are polyploid, and up to five basic chromosome numbers have been suggested (Al-Shehbaz, 1988). The most probable basic number for the majority of species is x = 8 (Kučera & al., 2005). For some species, such as the Beringian taxa in *C.* sect. *Cardaminella*, the most probable basic number is x = 7 (Elven & al., 2006). Diploids are known only with $2n$ = 16, and the highest recorded number is $2n = 32x = 256$ (*C. concatenata* O. Schwarz and *C. diphylla* Wood; Kučera & al., 2005). The reliable reports for *Cardamine microphylla* Adams are of tetraploid and hexaploid ($2n$ = 28, 42) plants from S Chukotka (Zhukova,

1980). There are other reports, but all are dubious and must be checked against vouchers before acceptance: Zhukova & Tikhonova (1973) and Krogulevich (1976) from Asia, and Mulligan (1965) from North America. Reports for *C. digitata* are considered reliable; the species is tetraploid and hexaploid ($2n = 28, 42$). Tetraploids are reported from Alaska (Rollins, 1966; Johnson & Packer, 1968) and Chukotka (Zhukova, 1969; Zhukova & Petrovsky, 1971, 1972; Zhukova & al., 1973; Zhukova & Petrovsky, 1984), while hexaploids are reported from Chukotka (Zhukova, 1965 also one count as $2n = 40$, 1966; Zhukova & Petrovsky, 1972; Zhukova & al., 1973; Petrovsky & Zhukova, 1981). The identity of some of the plants counted as *C. blaisdellii* Eastw. is problematic, but reports show the taxon to be tetraploid and hexaploid. Tetraploids are reported from Alaska (Murray & Kelso, 1997) as *C. microphylla* subsp. *blaisdellii* and from Chukotka (Zhukova & al., 1973 as *C. hyperborea*; Zhukova & Tikhonova, 1973 as *C. hyperborea*; Zhukova & Petrovsky, 1975, 1976, 1977 all as *C. hyperborea*, 1980, 1984, 1987), whereas hexaploids are only reported from Chukotka (Zhukova, 1966, 1969; Zhukova & al., 1973; Zhukova & Petrovsky, 1984). Reports of ploidy level for *C. purpurea* Cham. & Schltdl. are considered reliable. The single American report, from NW Alaska, is of a decaploid number ($2n =$ c. 80, $x = 8$; Johnson & Packer, 1968) whereas several counts of a dodecaploid number are reported from Wrangel Island ($2n = 96$, $x = 8$; Zhukova & Petrovsky, 1972; Petrovsky & Zhukova, 1981).

Microsatellites are widely used molecular markers for population genetic studies and have also been used to infer evolutionary relationships among closely related species (Harr & al., 1998; Petren & al., 1999; Alvarez & al., 2001; Chirhart & al., 2005). The evolutionary rate of microsatellites has been suggested to be too fast for phylogenetic studies. However, for several studies of closely related species, where sequence variation is difficult to obtain, microsatellites have been a useful marker system (Goldstein & Pollock, 1997; Schlötterer, 2001). *Cardamine* is a young genus with little sequence variation (Koch & al., 2000, Haubold & Wiehe, 2001; Carlsen & al., in press), microsatellites were therefore used in order to obtain enough variation to delimit species in the genus.

We tested the number and the circumscription of taxa in the *Cardamine digitata* aggregate using microsatellites. We conclude that there are four distinct taxonomic units at equal rank, preferrably as species: *C. blaisdellii* Eastw., *C. digitata* Richardson, *C. microphylla* Adams, and *C. purpurea* Cham. & Schltdl. Furthermore, there might be an additional unit we informally indicate as *C.* "*hyperborealis*". We use these names as a framework.

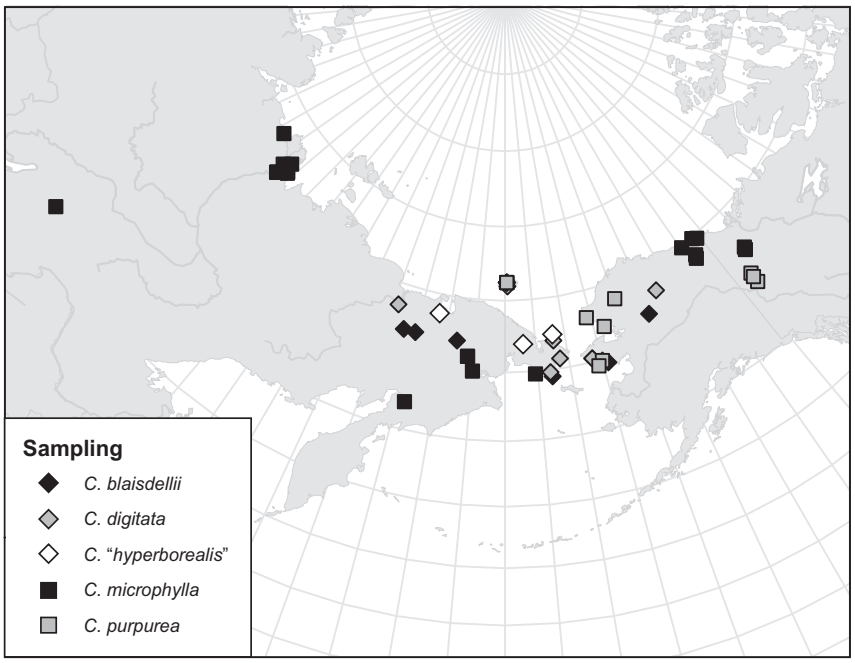


Fig. 1. Sampling of the *Cardamine digitata* aggregate included in this study. The map shows approximately the known ranges of the complex. See Appendix for further details.

MATERIAL AND METHODS

Material. — We included 131 specimens from 54 populations, a sampling covering most of the known ranges (Fig. 1; Appendix). Both herbarium (denoted TC) and silica-dried (denoted BE or SUP) specimens were included (Appendix). The herbarium specimens are deposited at the V.L. Komarov Botanical Institute, Russian Academy of Sciences, St. Petersburg (LE; denoted TC03) and the University of Alaska Museum of the North Herbarium, Fairbanks (ALA; denoted TC06; Appendix). Vouchers for the remaining populations are deposited at the Natural History Museum, University of Oslo, Oslo (O).

Microsatellite analysis. — DNA was extracted using the DNeasy™ Plant Mini Kit or DNeasy™ Plant 96 Kit (Qiagen) following the manufacturer's protocol. Microsatellites were amplified using marker specific primers (Table 1; MWG Biotech AG) and M13-primers (5′-CACGACGTTG TAAAACGAC-3′; Schuelke 2000) dyed with FAM (MWG Biotech AG), VIC (Applied Biosystems), and NED (Applied Biosystems). A PCR reaction volume of 10 µL contained 0.1 mM dNTP, 1.5–3.0 mM $MgCl_2$ (Table 1; Applied Biosystems), 0.01% BSA (Roche), 10 µM TMA (Sigma), 0.2 µM reverse primer (Table 1), 0.06 µM forward primer (Table 1), 0.2 µM M13 labeled primer (Table 1), 0.15 U AmpliTaq Gold (Applied Biosystems), 1× PCR Gold Buffer (Applied Biosystems), mqH_2O, and 1.0 µL diluted DNA template. The reactions were run with GeneAmp PCR system 9700 (Applied Biosystems) at the Natural History Museum, University of Oslo, using one of three different PCR programs (depending on marker; Table 1). The PCR products of different dyes were coloaded, and 1 µL of product mixture (FAM : NED : VIC = 2 : 3 : 2) was added 8.8 µL HiDi (formamide) and 0.2 µL GeneScan Rox 500 size standard (Applied Biosystems). The products were denatured for 5 min at 95°C and analyzed with an ABI 3100 Sequencer (Applied Biosystems) with 10 s injection time and 3 kV injection voltage, otherwise default conditions.

Numerous microsatellite markers previously developed for *Arabidopsis thaliana* were tested (Skrede & al., 2008), and six markers with an appropriate level of variation were chosen (Table 1). Three replicates were made for every sample included. The resulting profiles were visualized, sized and scored using GeneMapper vs. 3.7 (Applied Biosystems). Due to polyploidy, the markers were treated as dominant, and peaks in the range of 50–500 base pairs (Table 1) were scored as present (1) or absent (0). Variation in ploidy level is problematic when scoring microsatellite loci due to partial heterozygous individuals (e.g., AAAB, ABCC, ABBB, etc.). Studies analysing microsatellites for polyploid species often score dosage differences. As the plants in the present study have ploidy levels from diploid to dodecaploid, and as the ploidy level of most of the investigated specimens is not known, scoring dosage differences is an impossible task. We therefore decided to score the microsatellites as phenotypes rather than genotypes. The final dataset consisted of 131 individuals and 102 bands.

Data analysis. — The variation in the dataset was visualized using principal coordinate analysis (PCO) in NTSYSpc version 2.02 (Rohlf, 1999) based on the similarity measure of Dice (1945). The first 96 of the resulting 131 eigenvectors explained 100% of the variation in the dataset, and these were analyzed with canonical variate analysis

Table 1. Microsatellite markers used in this study.

Marker	Forward	Reverse	Dye	[$MgCl_2$]	PCR program	App. range
ICE14	5′-TCGAGGTGCTTTCTGA GGTT-3′	5′-TACCTCACCCTTTTGA CCCA-3′	FAM	2.5 mM	51	220–280
MR187	5′-GAGTTTTGGTTCCACC ATTA-3′	5′-CCCTTCAGCCTTTGAT AAAT-3′	NED	3.0 mM	51	145–195
Atts0191	5′-GACTGATGTTGATGGA GATGGTCA-3′	5′-CTCCACCAATCATGCA AATG-3′	VIC	1.5 mM	TD48	190–205
Atts0392	5′-GACGTTGATCGCAGCT TGATAAGC-3′	5′-TTGGAGTTAGACACGG ATCTG-3′	FAM	2.5 mM	TD50	145–225
nga1145	5′-GACCCTTCACATCCAA AACCCAC-3′	5′-GCACATACCCACAACC AGAA-3′	VIC	2.0 mM	TD50	245–275
AthCTRI	5′-GACTATCAACAGAAAC GCACCGAG-3′	5′-CCACTTGTTTCTCTCTC TAG-3′	NED	2.5 mM	TD50	145–160

All forward primers also contain a M13-tail (5′-CACGACGTTGTAAAACGAC-3′) at the 5′ end of the sequence. The PCR programs contained the following steps: 51 (5 min at 95°C, 35 cycles of the three steps 30 s at 94°C, 30 s at 51°C, and 45 s at 72°C, and a final hold of 20 min at 72°C), TD50 (5 min at 95°C, 16 cycles of the three steps 30 s at 94°C, 30 s at 58–50.5°C [decreasing 0.5°C every cycle], and 45 s at 72°C, 35 cycles of the three steps 30 s at 94°C, 30 s at 50°C, and 45 s at 72°C, and a final hold of 20 min at 72°C), or TD48 (5 min at 95°C, 10 cycles of the three steps 30 s at 94°C, 30 s at 53–48.5°C [decreasing 0.5°C every cycle], and 45 s at 72°C, 28 cycles of the three steps 30 s at 94°C, 30 s at 48°C, and 45 s at 72°C, and a final hold of 20 min at 72°C).

(CVA) and multivariate analysis of variance (MANOVA) in PAST version 1.29 (Hammer & Harper, 2004), following the procedure of Anderson & Willis (2003), of a priori dividing the specimens into groups according to the results of other analyses (PCO, STRUCTURE, parsimony).

A Bayesian approach using STRUCTURE version 2 (Pritchard & al., 2000) calculated a logarithmic probability for the data being assigned to a given number of clusters. The method was originally designed for codominant markers but may be applied to dominant markers under a no-admixture model, assuming no linkage among loci (Pritchard & al., 2000). Ten replicates of each value of K (= the number of groups) were run for different selections of samples with a burn-in period of 100,000 and 1,000,000 iterations. Similarity coefficients comparing the resulting assignments were calculated using StructureSum (Rosenberg & al., 2002; Ehrich, 2006).

Parsimony analyses were performed in TNT (Goloboff & al., 2000) with bands coded as present or absent. Heuristic searches were performed with 10,000 random additional sequences and TBR branch swapping, saving ten trees per replication. The resulting trees were swapped with TBR saving up to 100,000 trees altogether. Collapsing rule was set to minimum length = 0. Random seed was set to "time". Goodness of fit was calculated using consistency index (CI), retention index (RI), and rescaled consistency index (RC) (Kluge & Farris, 1969; Farris, 1989). Bremer supports (Bremer, 1994) were calculated producing 60,000 trees, of which 10,000 were one step longer, 10,000 were two steps longer, etc., up to six steps longer. Jackknife (Farris & al., 1996) and traditional bootstrap (Felsenstein, 1985) resampling studies were performed with 1,000 replicates (10 random entry orders and 10 trees saved for each repetition). Jackknifing was performed with 36% deletion. Both bootstrap and jackknife were performed with cut-off value of 50% and absolute frequencies as output.

RESULTS

Scoring. — The six microsatellite loci provided a matrix of 102 scored variable alleles. AthCTRI was the most conservative locus providing only one or two alleles even for the high-ploid individuals. Atts0392 was the most variable and allele-rich locus providing from two to nine alleles for each individual (Table 2). All isolates had a number of alleles that corresponded well to known or expected ploidy level, except for some accessions of *C. microphylla* from the Sakha Republic (Table 2) that had seven to nine alleles at the locus Atts0392.

Ordination analyses. — PCO analysis of all samples separated *Cardamine purpurea* and *C. microphylla* from *C. digitata* and *C. blaisdellii* along axis 1 (spanning 17.6% of the variation; Fig. 2A), although a few specimens of *C. digitata* were placed closer to *C. microphylla* than the majority of *C. digitata*. Axis 2 (10.7%) separated *C. purpurea* from *C. microphylla* and *C. digitata* from *C. blaisdellii*. The *C.* "*hyperborealis*" specimens grouped with *C. digitata* (Fig. 2A). Axis 3 (8.2%) gave no further information (not shown).

In an analysis excluding the *C. purpurea* samples, *C. microphylla* was found at high values along axis 1 (19.4%), *C. blaisdellii* was found at low values, and *C. digitata* was intermediate and partially overlapping with *C. microphylla* (not shown). The *C.* "*hyperborealis*" specimens grouped with *C. digitata* along the first axis, whereas one (TC03-28) grouped with *C. blaisdellii* along the second axis (11.1%; not shown).

A PCO analysis excluding both *C. purpurea* and *C. microphylla* separated *C. blaisdellii* and *C. digitata* along the first axis (20.8%; not shown). One of the *C.* "*hyperborealis*" plants (TC03-28) grouped with *C. blaisdellii*, and the other two with *C. digitata*. The second axis gave no additional taxonomic information (10.2%; not shown).

The CVA analysis separated the five a priori defined groups completely and significantly (Wilk's lambda = 3.485E–11, df1 = 384, df2 = 122.8, F = 131.8, P(same) = 5.636E–102; Pillai trace = 3.986, df1 = 384, df2 = 132, F = 94.65, P(same) = 8.102E–100; Fig. 3). The first axis (61.9%) separated *C. microphylla* from the other groups, whereas the second axis (26.2%) separated the remaining groups. *Cardamine* "*hyperborealis*" was intermediate between *C. blaisdellii* and *C. digitata* along the second axis (Fig. 3).

STRUCTURE analyses. — The STRUCTURE analysis including all specimens, separated the plants into

Table 2. Number of alleles per plant found for each taxon and locus.

	2*n*	ICE14	nga1145	Atts0392	MR187	AthCTRI	Atts0191
C. blaisdellii	28	1–2	2–4	2–4	1–4	1	1–2
C. digitata	28, 42	1–3	2–4	3–6	1–4	1–2	2–5
C. "*hyperborealis*"	42	2–3	3–4	4–5	3–4	1	2–4
C. microphylla	28, 42, 52	1–5	2–4	2–9	1–5	1–2	1–4
C. purpurea	96	1–4	1–4	2–6	2–7	1–2	1–2

2*n* gives chromosome counts of plants included in this study, cf. Appendix.

two groups; one comprising *C. microphylla* and *C. purpurea*, and one comprising *C. blaisdellii* and *C. digitata*. An increase in number of groups resulted in ambiguous division of the dataset (similarity coefficients < 1). Further analyses of the two groups separately gave no further division of the first group, while the second was unambiguously divided in two: one group comprising *C. blaisdellii*, one specimen of *C. digitata* (SUP02-177-3), and one of *C.* "*hyperborealis*" (TC03-28), and one group comprising *C. digitata* and the other two *C.* "*hyperborealis*" specimens.

Parsimony analyses. — Heuristic search and subsequent TBR swapping gave 6,336 most parsimonious trees of length 568, from two different "islands" in tree-space. Goodness of fit values were CI 0.180, RI 0.721, and RC 0.130. A strict consensus tree is presented in Fig. 4. Resampling analyses gave support only for internal nodes within species (result not shown). Bremer support value for the branch separating *C. purpurea* from the other species is 2 (Fig. 4). Bremer support value for the branch separating *C. microphylla* from *C. blaisdellii* and *C. digitata* is 1 (Fig. 4; Bremer support for all other branches not shown).

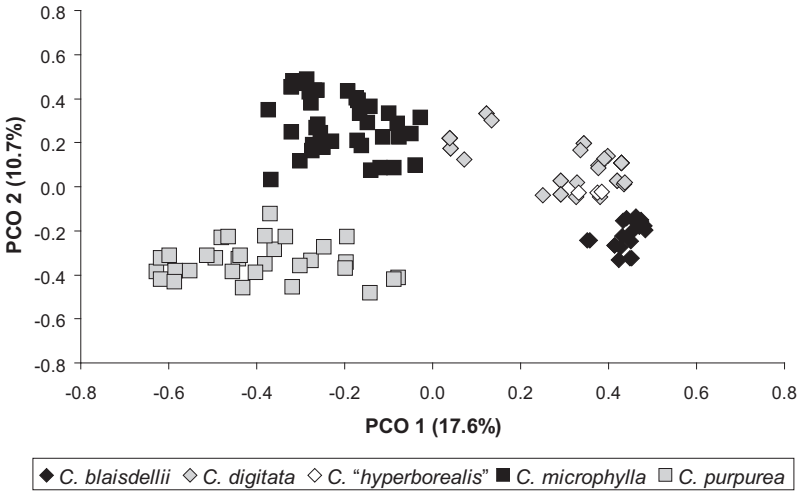


Fig. 2. PCO analysis of 131 plants in the *Cardamine digitata* aggregate based on six microsatellite loci and Dice's similarity.

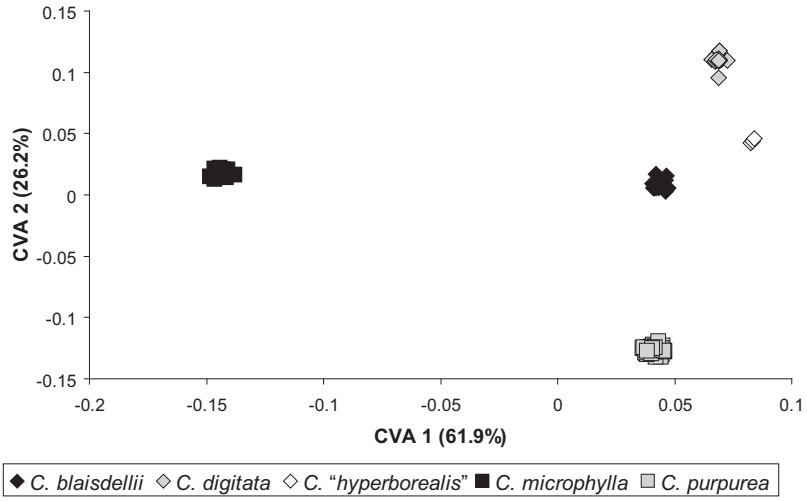


Fig. 3. CVA analysis of 96 eigenvectors from the PCO analysis of the *Cardamine digitata* aggregate, a priori grouped according to taxa. See text for further details.

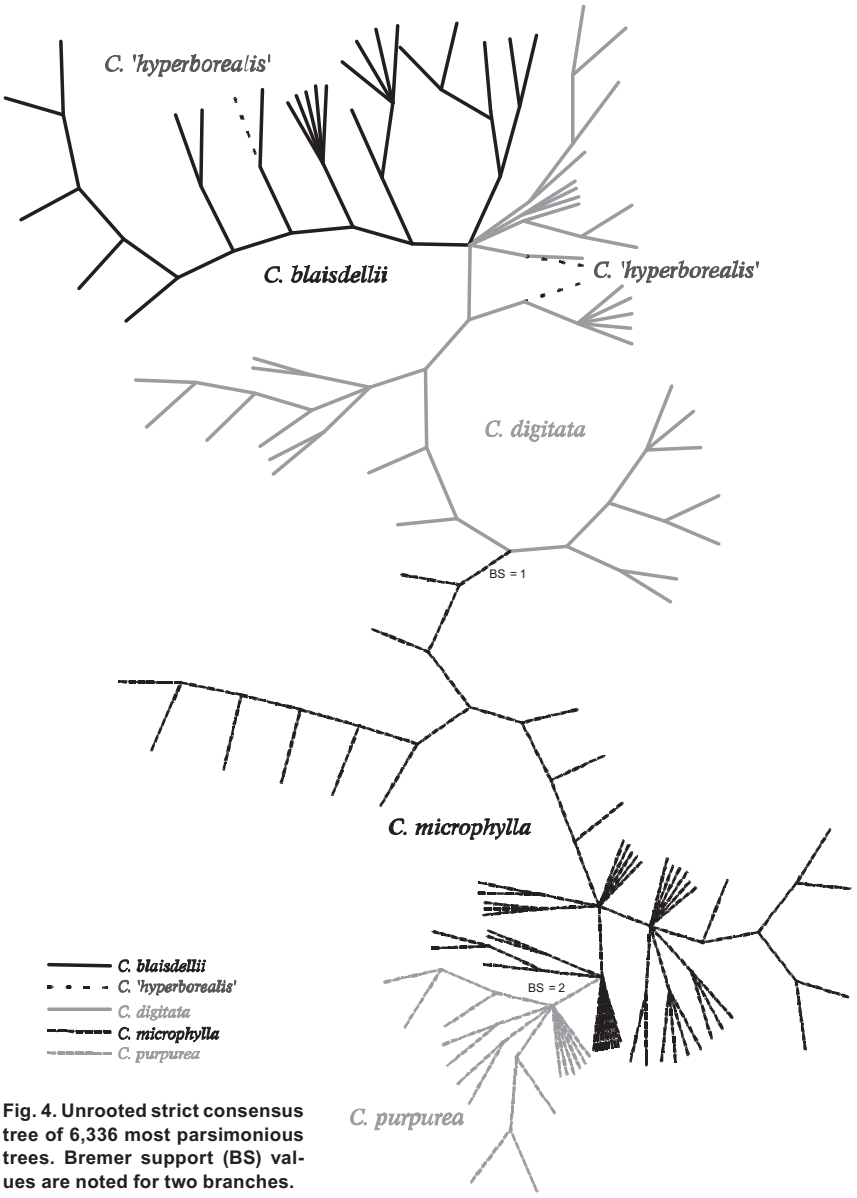


Fig. 4. Unrooted strict consensus tree of 6,336 most parsimonious trees. Bremer support (BS) values are noted for two branches.

DISCUSSION

The combination of analyses presented here supports the recognition of four evolutionary units in the *Cardamine digitata* aggregate. Although they yield partly different results and resolutions, the analyses are not conflicting. The separation of *C. purpurea* and *C. microphylla* from the remaining units is supported in all three analyses (PCO, STRUCTURE, parsimony analyses), and the isolation of the groups even received Bremer support of 2 and 1, respectively. *Cardamine digitata* and *C. blaisdellii* were separated by both STRUCTURE and PCO analyses. The resolution in the parsimony analyses gave no support to, but neither contradicted the separation of the two groups. As the four groups correspond to morphologically defined and comparatively distinct units, we suggest to acknowledge the groups as four taxa at the rank of species: *C. blaisdellii*, *C. digitata*, *C. microphylla*, and *C. purpurea*.

Murray & Kelso (1997) suggested reducing the rank of *C. blaisdellii* to that of subspecies, as *C. microphylla* subsp. *blaisdellii* (Eastw.) D.F. Murray & S. Kelso. We found no transition between *C. blaisdellii* and *C. microphylla* in our study. On the contrary, we found all four units to be approximately equidistant, suggesting similar ranking for all units. The closest relationship of *C. blaisdellii* is to *C. digitata* rather than to *C. microphylla*. Furthermore, *C. blaisdellii* and *C. microphylla* are parapatric with *C. blaisdellii* interspaced between the two major parts of the range of *C. microphylla*, and the genetic and morphological differences observed are therefore probably not resulting from the geographical isolation often defining subspecies.

The three specimens we have named *C.* "*hyperborealis*" are morphologically intermediates between *C. digitata* and *C. blaisdellii* (pers. obs.), and they are genetically grouped with both taxa. Deciding whether these represent an additional stabilised allopolyploid taxon requires further studying. Checking the vouchers of these plants revealed that they all had pollen of irregular size, indicating sterility and hybrid origin (pers. obs.). They were all, however, collected at an early developmental stage, making it impossible to examine fruits and seeds.

Chromosome numbers and ploidy levels. — The basic chromosome number of *C. purpurea* is $x = 8$, but $x = 7$ for the other study species. All numbers reported from the relevant units are polyploid, from tetraploid to dodecaploid. Assignment of chromosome number reports to taxa is made difficult by the different application of names and circumscriptions of species.

Two or more ploidy levels are documented for each of the four species. We have studied the numerous vouchers for the Russian chromosome counts (in LE) and find no evident morphological difference among plants at different ploidy levels in *C. blaisdellii*, *C. digitata* or *C. microphylla*. Taxonomic autopolyploidy is therefore indicated. In *C. purpurea* there is a slight difference in leaf shape between the isolated Wrangel Island plants (counted with $2n = 96$, chromosome number vouchers studied) and the NW Alaskan plants (counted with $2n =$ c. 80, but voucher for the chromosome count not studied by us) but neither of them approach any of the other species morphologically. The several NE Asian chromosome number vouchers annotated as *C. hyperborea* are another matter. The vouchers for tetraploid counts ($2n = 28$) correspond closely morphologically (and also in microsatellite markers) with tetraploid *C. blaisdellii* and we assign them there. The vouchers for hexaploid counts ($2n = 42$) are intermediate in shoot and leaf morphology between *C. blaisdellii* and *C. digitata*, and we informally name them as *C.* "*hyperborealis*".

Chromosome counts above hexaploid level (Mulligan, 1965; Zhukova & Tikhonova, 1973; Krogulevich, 1976) are dubious in *C. microphylla*, but, as we have found up to nine alleles in the locus Atts0392 in plants from the Sakha Republic (cf. Table 2), this could indicate that cytotypes exist at higher ploidy levels than currently documented in this species.

Cytotypes, not autopolyploid speciation. — The *Cardamine digitata* aggregate can be listed as one of the many plant groups showing that autopolyploids are much more common than traditionally maintained (reviewed by Soltis & al., 2003). Soltis & al. (2003) suggested that recurrent formation of polyploids is the rule, not the exception, and the different levels of polyploidy in at least three of four units in our study support this statement. But even though our knowledge on autopolyploid dynamics and the frequency of formation has dramatically increased during the past decade, the recognition of autopolyploidy as a major mode of speciation has not (Soltis & al., 2007). Soltis & al. (2007) claimed that a failure to name autopolyploids as separate species is caused by the adherence of plant systematists to a strict taxonomic species concept stressing morphological features, resulting in a serious underestimate of the role of polyploidy in plant speciation. The lack of both genetic and morphological distinction among cytotypes within our four units, however, indicates frequent gene flow, a condition considered by most systematists (including Soltis & al., 2007) to characterise conspecificity. Thus, we choose the conservative approach and suggest the cytotypes to be conspecific.

Concluding notes on nomenclature and taxonomy. — *Cardamine purpurea* Cham. & Schltdl. (Chamisso & Schlechtendal, 1826) was described from western Alaska: St. Lawrence Island in the northern Bering Sea: "Ins. St. Laurentii", leg. L.K.A. von Chamisso (HAL 85360) holotype (Hoffmann 2000). The meaning and application of the name has been unambiguous since its description. The species is mainly Beringian American (Yukon Territory and Alaska, widespread) and the only Asian occurrence reported by Petrovsky in Tolmachev (1975) is on Wrangel Island.

Cardamine microphylla Adams (Adams, 1817) was described from the estuary of Lena River in northern Siberia (Sakha Republic): "Promontorio Bykovský Mys, ora fl. Lena", leg. M.F. Adams (MW, lectotype). The name was until fairly recently applied collectively to include all plants with broad leaf lobes and comparatively large, white flowers (e.g., Hultén, 1968; also tentatively Petrovsky in Tolmachev, 1975; Porsild & Cody, 1980; Berkutenko, 1988), i.e., including *C. blaisdellii* (published 1902) which these authors have considered as a later synonym. With the late recognition of the more narrowly amphi-Beringian *C. blaisdellii*, by Khatri (1990) as a variety of *C. microphylla*, and by Murray & Kelso (1997) as a subspecies, the questions arise how to circumscribe *C. microphylla* s.str. and where it occurs. Plants identified morphologically as *C. microphylla* s.str. are found in three separate regions: in a restricted area around the Lena River estuary

in N Siberia (Petrovsky in Tolmachev, 1975), in a wider area in W and E Chukotka in NE Russian Far East (about half of what Petrovsky in Tolmachev (1975) mapped as *C. microphylla*, the other half is part of *C. blaisdellii*), and in a significant area in NE Alaska, N Yukon Territory, and NW Mackenzie District. The gap between the N Siberian and Chukotkan parts of the range is c. 1,900 km, that between the Chukotkan and NW American parts c. 1,500 km.

Cardamine digitata Richardson (Richardson, 1823) was described from NW North America with a type from NW Canada, Mackenzie District: "Barren Grounds from lat. 64° to Arctic Sea, in lat. 69°", leg. N. Richardson (BM, holotype). There has been much confusion concerning the name. Trautvetter (1879) applied Richardson's name to plants with broad leaf lobes and coined the name *C. digitata* var. *oxyphylla* Trautv. for the plants with narrow leaf lobes (i.e., those of the current-day opinion of *C. digitata* s.str.). In the only global revision of *Cardamine*, Schulz (1903) rejected the name *C. digitata* Richardson as he assumed that it was predated by *C. digitata* Lam. (Lamarck, 1786) and thereby a later homonym. He coined *C. hyperborea* O.E. Schulz as a nomen novum for *C. digitata* in the Richardson meaning (Schulz, 1903). He also followed Trautvetter (1879) in assuming this to be a plant with broad leaf lobes and made the combination *C. hyperborea* var. *oxyphylla* (Trautv.) O.E. Schulz for the plant with narrow leaf lobes (Schulz, 1903). Both propositions were erroneous. Lamarck (1786) described his species as a *Dentaria*, not as a *Cardamine*, and the name therefore does not make illegitimate Richardson's *C. digitata* which is the correct name for the species. Richardson's plant is that with narrow leaf lobes (as seen in the type specimen). Only such plants occur in the region from where it was described. Trautvetter's and Schulz's *C. digitata* var. *oxyphylla* and *C. hyperborea* var. *oxyphylla* is the true *C. digitata* whereas their application of these species names excluding the var. *oxyphylla* refers to another species, probably *C. blaisdellii*. That *C. hyperborea* is necessarily homotypic with the legitimate *C. digitata* and is, therefore, a superfluous synonym was already pointed out by Shetler (1961) and Rollins (1993). The next step in the confusion of names and applications is the introduction of the name *C. richardsonii* Hultén. Hultén (1945) coined this name as a nomen novum for *C. digitata* Richardson as he also, on the authority of Schulz, erroneously assumed that Richardson's name was predated by Lamarck's name and in addition that the name *C. hyperborea* as applied by Schulz belonged to the plants with broad leaf lobes. Hultén's name is therefore a superfluous, full synonym of *C. digitata*. It has been applied by, e.g., Löve & Löve (1975), but was already rejected by Hultén (1968) in favour of *C. hyperborea*. *Cardamine digitata* is a widespread northern Beringian and North American plant with a nearly continuous range from W Chukotka eastwards to Hudson Bay (Petrovsky in Tolmachev, 1975; Porsild & Cody, 1980).

The meaning of the name *C. hyperborea* is now clear, as a homotypic synonym of *C. digitata*. Its previous application is problematic, however, especially in Siberia. Petrovsky in Tolmachev (1975, Russian Arctic) entered *C. digitata* and *C. hyperborea* as two species and only tentatively indicated *C. blaisdellii* as possibly synonymous with the latter. Petrovsky still (pers. comm.) suggests that there may be a third entity present in NE Asia, besides *C. digitata* and *C. blaisdellii* (which he now accepts as a species and as a part of his 1975 concept of *C. hyperborea*). Berkutenko in Charkevicz (1988, Russian Far East) also entered *C. digitata* and *C. hyperborea* but without reference to *C. blaisdellii*. Doronkin in Malyschev & Peschkova (1994, Siberia, i.e., excl. Russian Far East) synonymized *C. hyperborea* in the Schulz meaning (excluding the plants with narrow leaf lobes) with *C. microphylla*.

Cardamine blaisdellii Eastw. (Eastwood, 1902) was described from W Alaska with a holotype from Seward Peninsula: Cape Nome, summer 1900, leg. F.E. Blaisdell (CAS). Porsild (1938) synonymized *C. hyperborea* with *C. blaisdellii*, but the name was otherwise largely forgotten until Khatri's revision *C.* sect. *Cardaminella* (Khatri, 1990) and especially a closer study of the Alaskan plants by Murray & Kelso (1997). Since then *C. blaisdellii* or *C. microphylla* subsp. *blaisdellii* has been accepted as the correct name for at least the NW North American parts of Schulz's *C. hyperborea* with broad leaf lobes, recently also for at least a major part of the NE Asian plants (Petrovsky, pers. comm.). As currently understood *C. blaisdellii* is amphi-Beringian with a small part area in E Chukotka and a larger one in Alaska, Yukon Territory, and probably reaching Mackenzie District in the mountains west of Mackenzie River.

ACKNOWLEDGEMENTS

Thanks are due to Loren Rieseberg and the Rieseberg lab at Indiana University for lab facilities and giving advice to I. Skrede during primer-testing, to the Botanical Garden, Natural History Museum, University of Oslo for the grant to T. Carlsen covering all lab costs, to Volodya Razzhivin and Galina Gusarova for assistance in the V.L. Komarov Botanical Institute, to Liv Guro Kvernstuen for advices on microsatellites, to Øyvind Hammer for assistance with ordination analysis, and to David F. Murray, Christian Brochmann, and Anne K. Brysting for comments and suggestions on the manuscript. Thanks to everyone collecting material for us. Some material was obtained via R. Elven's participation in Beringia 2005 expedition funded by the Polar Research Secretariat at the Royal Swedish Academy of Sciences, and from expeditions funded by the Strategic University Programme grant 146515/420 from the Research Council of Norway to C. Brochmann and R. Elven (subprogram "Migration and evolution of arctic plants in response to Quaternary climate changes").

LITERATURE CITED

Adams, M.F. 1817. Descriptiones plantarum minus cognitarum Sibiriae, praesertim orientales quas in itinere ann. 1805 et 1806 observavit. *Mém. Soc. Imp. Naturalistes Moscou* 5: 89–116.

Al-Shehbaz, I.A. 1988. The genera of Arabideae (Cruciferae; Brassicaceae) in the southeastern United States. *J. Arnold Arbor.* 69: 85–166.

Al-Shehbaz, I.A., Beilstein, M.A. & Kellogg, E.A. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Pl. Syst. Evol.* 259: 89–120.

Alvarez, A.E., Van de Wiel, C.C.M., Smulders, M.J.M. & Vosman, B. 2001. Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. *Theor. Appl. Genet.* 103: 1283–1292.

Anderson, M.J. & Willis, T.J. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84: 511–525.

Berkutenko, A.N. 1988. Brassicaceae. Pp. 38–115 in: Charkevicz, S. (eds.), *Plantae vasculares orientis extremi sovietici*, vol. 3. Nauka, Leningrad.

Bremer, K. 1994. Branch support and tree stability. *Cladistics* 10: 295–304.

Carlsen, T., Bleeker, W., Hurka, H., Elven, R. & Brochmann, C. In press. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Ann. Missouri Bot. Gard.*

Chamisso, L.A. v. & Schlechtendal, F.L. v. 1826. De plantis in expeditione speculatoria Romanzoffiana observatis rationem dicunt. 1. *Linnaea* 1: 1–73.

Chirhart, S.E., Honeycutt, R.L. & Greenbaum, I.F. 2005. Microsatellite variation and evolution in the *Peromyscus maniculatus* species group. *Molec. Phylog. Evol.* 34: 408–415.

Dice, L.R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26: 297–302.

Eastwood, A. 1902. A descriptive list of the plants collected by Dr. F. E. Blaisdell at Nome City, Alaska. *Bot. Gaz.* 33: 126–149.

Ehrich, D. 2006. AFLPDAT: a collection of R functions for convenient handling of AFLP data. *Molec. Ecol. Notes* 6: 603–604.

Elven, R., Murray, D.F., Razzhivin, V.Y. & Yurtsev, B.A. 2006. *Panarctic Flora Checklist. Vascular Plants. Draft version (2006).* National Centre for Biosystematics, University of Oslo, Oslo. http://www.nhm.uio.no/panarctflora/checklist/pafchckl.htm.

Farris, J.S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

Farris, J.S., Albert, V.A., Källersjö, M., Lipscomb, D. & Kluge, A.G. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 4: 783–791.

Franzke, A. & Hurka, H. 2000. Molecular systematics and biogeography of the *Cardamine pratensis* complex (Brassicaceae). *Pl. Syst. Evol.* 224: 213–234.

Goldstein, D.B. & Pollock, D.D. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Heredity* 88: 335–342.

Goloboff, P.A., Farris, J.S. & Nixon, K.C. 2000. *TNT (Tree Analysis Using New Technology)*. Published by the authors. Tucumán.

Hammer, Ø. & Harper, D.A.T. 2004. *PAST. Palaeontological statistics.* http://folk.uio.no/ohammer/past/.

Harr, B., Weiss, S., David, J.R., Brem, G. & Schlötterer, C. 1998. A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Curr. Biol.* 8: 1183–1186.

Haubold, B. & Wiehe, T. 2001. Statistics of divergence times. *Molec. Biol. Evol.* 18: 1157–1160.

Hewson, H.J. 1982. *Flora of Australia*. Australian Government Publishing Service, Canberra.

Hoffmann, M.H. 2000. Type specimens of the Brassicaceae in the Herbarium of the Martin-Luther-University Halle Wittenberg (HAL). *Schlechtendalia* 4: 35–40.

Hultén, E. 1945. Flora of Alaska and Yukon V. Dicotyledoneae. Rhoedales, Sarraceniales, Rosales I (*Crassulaceae*, *Saxifragaceae*). *Acta Univ. Lund., n.s.,* 41: 797–978.

Hultén, E. 1968. *Flora of Alaska and Neighboring Territories. A Manual of the Vascular Plants.* Stanford University Press, Stanford.

Johnson, A.W. & Packer, J.G. 1968. Chromosome numbers in the flora of Ogotoruk Creek, N.W. Alaska. *Bot. Not.* 121: 403–456.

Khatri, K.S. 1990. *Cardamine* L. sect. *Cardaminella* Prantl (Cruciferae) in the Asian territory of Soviet Union. *Feddes Repert.* 101: 435–448.

Kluge, A.G. & Farris, J.S. 1969. Quantitative phyletics and evolution of Anurans. *Syst. Zool.* 18: 1–32.

Koch, M.A., Haubold, B. & Mitchell-Olds, T. 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Molec. Biol. Evol.* 17: 1483–1498.

Krogulevich, R.E. 1976. Rol poliploidii v genezise flory Putorana. Pp. 217–235 in: Malyshev, L.I. (ed.), *Flora Putorana. Materialy k poznaniiu osobennostei sostava i genezica gornykh sybarkticheskikh flor Sibiri*. Izdatelstvo "Nauka", Novosibirsk.

Kučera, J., Valko, I. & Marhold, K. 2005. On-line database of the chromosome numbers of the genus *Cardamine* (Brassicaceae). *Biologia (Bratislava)* 60: 473–476.

Lamarck, J.B.A.P.M. de. 1786. *Encyclopédie méthodique. Botanique*, vol. 2, *Cic–Gor.* Panckoucke, Paris.

Lihová, J. & Marhold, K. 2006. Phylogenetic and diversity patterns in *Cardamine* (Brassicaceae)—a genus with conspicuous polyploid and reticulate evolution. Pp. 149–186 in: Sharma, A.K. & Sharma, A. (eds.), *Plant Genome: Biodiversity and Evolution*. Science Publishers, Enfield.

Lihová, J., Marhold, K., Tribsch, A. & Stuessy, T.F. 2004. Morphometric and AFLP re-evaluation of tetraploid *Cardamine amara* (Brassicaceae) in the Mediterranean. *Syst. Bot.* 29: 134–146.

Löve, Á. & Löve, D. 1975. *Cytotaxonomical Atlas of the Arctic Flora*. Cramer, Vaduz.

Malyschev, L.I. & Peschkova, G.A. (eds.). 1994. *Flora Sibiri*, vol. 7, *Berberidaceae–Grossulariaceae.* Nauka, Novosibirsk.

Marhold, K. & Lihová, J. 2006. Polyploidy, hybridization and reticulate evolution: lessons from Brassicaceae. *Pl. Syst. Evol.* 259: 143–174.

Marhold, K., Lihová, J., Perný, M. & Bleeker, W. 2004. Comparative ITS and AFLP analysis of diploid *Cardamine* (Brassicaceae) taxa from closely related polyploid complexes. *Ann. Bot.* 93: 507–520.

Marhold, K., Lihová, J., Perný, M., Grupe, R. & Neuffer, B. 2002. Natural hybridization in *Cardamine* (Brassicaceae) in the Pyrenees: evidence from morphological and molecular data. *Bot. J. Linn. Soc.* 139: 275–294.

Mulligan, G.A. 1965. Chromosome numbers of the family Cruciferae. II. *Canad. J. Bot.* 43: 657–668.

Murray, D.F. & Kelso, S. 1997. Chromosome numbers and notes on the taxonomy of selected Alaskan vascular plants. *Rhodora* 99: 33–55.

Otto, S.P. & Whitton, J. 2000. Polyploid incidence and evolution. *Annual Rev. Genet.* 34: 401–437.

Petren, K., Grant, B.R. & Grant, P.R. 1999. A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 266: 321–329.

Petrovsky, V.V. & Zhukova, P.G. 1981. Chromosome numbers and taxonomy of some plant species of Wrangel Island. *Bot. Zhurn. (Moscow & Leningrad)* 66: 380–387. [In Russian].

Porsild, A.E. 1938. Flora of Little Diomede Island in Bering Strait. *Trans. Roy. Soc. Canada, Sect. 5, Biol. Sci.* 32: 21–28.

Porsild, A.E. & Cody, W.J. 1980. *Vascular Plants of Continental Northwest Territories, Canada*. National Museum of Natural Sciences, Ottawa.

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

Richardson, J. 1823. Botanical appendix. Pp. 729–768 in: Franklin, J. (ed.), *Narrative of a Journey to the Shores of the Polar Sea in the Years 1819, 20, 21, and 22 With an Appendix on Various Subjects Relating to Science and Natural History, Illustrated by Numerous Plates and Maps*. John Murray, London.

Rohlf, F. 1999. *NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System.* Exeter Software. New York.

Rollins, R.C. 1966. Chromosome numbers of Cruciferae. *Contr. Gray Herb.* 197: 43–65.

Rollins, R.C. 1993. *The Cruciferae of Continental North America. Systematics of the Mustard Family from the Arctic to Panama.* Stanford University Press, Stanford.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A. & Feldman, M.W. 2002. Genetic structure of human populations. *Science* 298: 2381–2385.

Schlötterer, C. 2001. Genealogical inference of closely related species based on microsatellites. *Genet. Res.* 78: 209–212.

Schuelke, M. 2000. An economic method for the flourescent labeling of PCR fragments. *Nat. Biotechnol.* 18: 233–234.

Schulz, O.E. 1903. Monographie der Gattung *Cardamine*. *Bot. Jahrb. Syst.* 32: 280–623.

Schulz, O.E. 1936. Cruciferae. Pp. 227–658 in: Engler, A. & Harms, H. (eds.), *Die natürlichen Pflanzenfamilien*, ed. 2, vol. 17B. Engelmann, Leipzig.

Shetler, S.G. 1961. Application of the names *Cardamine digitata*, *Dentaria pentaphylla*, and *D. heptaphylla*. *Taxon* 10: 264–267.

Sjöstedt, B. 1975. Revision of genus *Cardamine* L. (Cruciferae) in South America and Central America. *Bot. Not.* 128: 8–19.

Skrede, I. & Carlsen, T., Rieseberg, L.H. & Brochmann, C. 2008. Microsatellites for three distantly genera in the Brassicaceae. *Cons. Genet.* DOI 10.1007/s10592-008-9598-x

Soltis, D.E., Soltis, P.S., Schemske, D.W., Hancock, J.F., Thompson, J.N., Husband, B.C. & Judd, W.S. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56: 13–30.

Soltis, D.E., Soltis, P.S. & Tate, J.A. 2003. Advances in the study of polyploidy since *Plant Speciation*. *New Phytol.* 161: 173–191.

Tolmachev, A.I. 1975. *Flora Arctica URSS*, vol. 7, *Papaveraceae–Cruciferae*. Nauka, sectio Leninopoli, Leninopoli.

Trautvetter, E.R. v. 1879. Flora terrae tschuktschorum. *Trudy Imp. S.-Peterburgsk. Bot. Sada* 6: 5–40.

Webb, C.J., Sykes, W.R. & Garnock-Jones, P.J. 1988. *Flora of New Zealand.* Botany division, D.S.I.R, Christchurch.

Zhukova, P.G. 1965. Caryological characteristics of some plant species of Wrangel Island. *Bot. Zhurn. (Moscow & Leningrad)* 50: 1320–1322. [In Russian]

Zhukova, P.G. 1966. Chromosome numbers in some species of plants of the north-eastern part of the U.S.S.R. *Bot. Zhurn. (Moscow & Leningrad)* 51: 1511–1516. [In Russian]

Zhukova, P.G. 1969. Chromosome number in certain plant species indigenous to the North-East of the USSR IV. *Bot. Zhurn. (Moscow & Leningrad)* 54: 1985–1990. [In Russian]

Zhukova, P.G. 1980. Chromosome number of some Southern Chukotka plant species. *Bot. Zhurn. (Moscow & Leningrad)* 65: 51–59. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1971. Chromosome numbers of certain flowering plants of the Wrangel Island. *Bot. Zhurn. (Moscow & Leningrad)* 56: 294–305. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1972. Chromosome numbers of some flowering plants of the Wrangel Island. II. *Bot. Zhurn. (Moscow & Leningrad)* 57: 554–563. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1975. Chromosome numbers of some western Chukotka plant species. *Bot. Zhurn. (Moscow & Leningrad)* 60: 395–401. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1976. Chromosome numbers of some western Chukotka plant species, II. *Bot. Zhurn. (Moscow & Leningrad)* 61: 963–969. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1977. Chromosome numbers of some western Chukotka plant species. III. *Bot. Zhurn. (Moscow & Leningrad)* 62: 1215–1223. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1980. Chromosome numbers and taxonomy of some species of the Anyui Mts. *Bot. Zhurn. (Moscow & Leningrad)* 65: 651–659. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1984. A cytotaxonomical study of some species of the family Brassicaceae in northern Asia. *Bot. Zhurn. (Moscow & Leningrad)* 69: 236–240. [In Russian]

Zhukova, P.G. & Petrovsky, V.V. 1987. Chromosome numbers and taxonomy of some plant species from the northern Asia regions. *Bot. Zhurn. (Moscow & Leningrad)* 72: 1617–1624. [In Russian]

Zhukova, P.G., Petrovsky, V.V. & Plieva, T.V. 1973. The chromosome numbers and taxonomy of some plant species from Siberia and Far East. *Bot. Zhurn. (Moscow & Leningrad)* 58: 1331–1342. [In Russian]

Zhukova, P.G. & Tikhonova, A.D. 1973. The chromosome numbers of some Chukotka plant species. II. *Bot. Zhurn. (Moscow & Leningrad)* 58: 395–402. [In Russian]

Appendix. Sampled material of the *Cardamine digitata* aggregate. The states are abbreviated as follows: ALA, Alaska; CAN, Canada; DFO, Far Eastern Federal District; SFO, Siberian Federal District. Positive and negative values for longitudes and latitudes give N/E and S/W, respectively. Chromosome counts give the original collection numbers, chromosome numbers, and publications: 1, 2, 3, 6, 7, 8: Zhukova & Petrovsky (1976, 1980, 1977, 1972, 1987, 1984, resp.); 4, Petrovsky & Zhukova (1981); 5, Zhukova & al. (1973); 9, 11: Zhukova (1969, 1980, resp.); 10, Zhukova & Tikhonova (1973).

Species: population/herbarium number (no. of specimens, state, region, locality, lat., long., year, collectors, chromosome counts)

Cardamine blaisdellii*:** BE05-41 (5, DFO, Chukotka Autonomous Okrug, Chukchi Peninsula, Novo Chaplino, 64.50, –172.83, 2005, *Solstad, Elven*), SUP02-166 (5, ALA, Nome Census Area, Seward Peninsula, Solomon River East Fork, 64.70, –164.18, 2002, *Elven, Gabrielsen, Jørgensen*), SUP02-193 (5, ALA, Nome Census Area, Seward Peninsula, Teller, coast/cliffs W of Teller, 65.25, –166.41, 2002, *Elven, Gabrielsen, Jørgensen*), SUP02-243 (5, ALA, Northwest Arctic Borough, Kobuk River Area, Anayucham Mts, Fritts Mt, 66.92, –155.52, 2002, *Elven*), TC03-30 (1, DFO, Chukotka Autonomous Okrug, W Chukotka, Bilibino settlement, 67.00, 172.00, 1974, *Tikhonova*, 74-216 T, $2n = 28$[1]), TC03-31 (1, DFO, Chukotka Autonomous Okrug, W Chukotka, Anyuy Mts, Pogynden River basin, Alarmagtyn River, 67.00, 165.00, 1974, *Tikhonova*, 74-10 T, $2n = 28$[2]), TC03-32 (1, DFO, Chukotka Autonomous Okrug, W Chukotka, Anyuy Mts, Mainghy-Pauktuvaam River, 67.00, 163.00, 1976, *Petrovsky, Koroleva*, 76-86, $2n = 28$[3]); ***C. digitata*:** BE05-555 (5, DFO, Chukotka Autonomous Okrug, Chukchi Peninsula, Lavrentiya Bay, bay 19 km NW of Lavrentiya, 65.70, –171.32, 2005, *Solstad, Elven*), BE05-260 (3, DFO, Chukotka Autonomous Okrug, Chukchi Peninsula, Penkigney Bay, Pestolova River, 64.83, –173.09, 2005, *Solstad, Elven*), BE05-979 (5, DFO, Chukotka Autonomous Okrug, Wrangel Isl, Somnitelnaya settlement and River, 70.98, –179.55, 2005, *Solstad, Elven*), SUP02-177 (4, ALA, Nome Census Area, Seward Peninsula, Kigluaik Mts, Mt W of Shaw Creek, 64.93, –164.99, 2002, *Elven, Gabrielsen, Jørgensen*), SUP02-189 (5, ALA, Nome Census Area, Seward Peninsula, Teller, coast/cliffs W of Teller, 65.25, –166.41, 2002, *Elven, Gabrielsen, Jørgensen*), SUP02-274 (5, ALA, North Slope Borough, Brooks Range, Endicott Mts, river south of Chandler Lake, 68.20, –152.74, 2002, *Elven*), TC03-33 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl, Somnitelnaya harbour, 71.25, –179.67, 1979, *Petrovsky*, 79-224, $2n = 42$[4]), TC03-34 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl, Somnitelnaya harbour, 71.25, –179.67, 1971, Petrovsky, *Steinberg*, 71-47, $2n = 42$[5]), TC03-35 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl, Mamontoraya River, 71.25, –179.67, 1970, Zhukova, *Petrovsky*, 70-213, $2n = 42$[6]), TC03-36 (1, DFO, Chukotka Autonomous Okrug, W Chukotka, Anyuisky Mts, by Anyuy River basin, Rybnaya River, 68.50, 160.82, 1982, Plieva, *Petrovsky*, 82-143/82-144, $2n = 28$[7]), TC03-37 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl, Gusinaya River, 71.25, –179.67, 1970, *Zhukova, Petrovsky*, 70-66, $2n = 28$[6]), TC03-38 (1, DFO, Chukotka Autonomous Okrug, E Chukotka, Lavrentiya settlement, 67.00, –172.00, 1972, *Zhukova*, 72-20 9?, $2n = 28$[8]); ***C. "hyperborealis": TC03-27 (1, DFO, Chukotka Autonomous Okrug, E Chukotka, Lavrentiya settlement, 67.50, –172.00, 1972, *Zhukova*, 72-38, $2n = 42$[8]), TC03-28 (1, DFO, Chukotka Autonomous Okrug, W Chukotka, Baranikha settlement, 68.85, 168.25, 1967, *Korobkov*, 67-18 K, $2n = 42$[9]), TC03-29 (1, DFO, Chukotka Autonomous Okrug, E Chukotka, Amguema R., 115 km road Egvekinot—Iultin, 67.00, –177.00, 1970, *Kozlova, Tikhonova*, 70-52 T, $2n = 42$[10]); ***C. microphylla*:** SUP-3912 (5, DFO, Sakha Republic, Lena River west bank, Chekurovka village, valley 1–2 km N of settlement, 71.06, 127.51, 2004, *Solstad, Elven*), SUP-3946 (5, DFO, Sakha Republic, Lena River west bank, plateau mountain 3–5 km W of Chekurovka village, 71.06, 127.47, 2004, *Solstad, Elven*), SUP-4093 (5, DFO, Sakha Republic, Lena River estuary, Area of Lena-Nordenskiöld Research Station, NE-most Kharaulakh Mts and Lena River delta flat, 72.20, 128.06, 2004, *Solstad, Elven*), SUP-4128 (5, DFO, Sakha Republic, Tiksi S, valley and small mountain E of town, 71.64, 128.86, 2004, *Solstad, Elven*), SUP03-372 (5, CAN, Yukon Territory, Richardson Mts, Wright Pass W side, Dempster hwy, 463–465 km, 67.05, –136.25, 2003, *Solstad, Elven*), TC03-130 (1, DFO, Chukotka Autonomous Okrug, Beringovsky, W border of Pekulnejskoe Lake, 65.00, 175.00, 1984, *Korobkov*), TC03-132 (1, DFO, Koryak Autonomous Okrug, North Korjakia, 10 (…), coast of Majnip [Majnik?] Lake, 62.00, 166.00, 1984, *Razzhivin*), TC03-24 (1, DFO, Chukotka Autonomous Okrug, E Chukotka, Nunligran settlement, 64.80, –175.40, 1970, *Korobkov,* 70-25 K, $2n = 52$[10]), TC03-25 (1, DFO, Chukotka Autonomous Okrug, S Chukotka, Pekulney Ridge, Bychya River, 66.00, 174.00, 1977, *Zhukova*, 77-214, $2n = 28$[11]), TC03-26 (1, DFO, Chukotka Autonomous Okrug, S Chukotka, Pekulney Ridge, Bychya River, 66.00, 174.00, 1977, *Zhukova*, 77-162, $2n = 42$[11]), TC03-3 (1, DFO, Sakha Republic, Lena River, Olenek Gulf, Stannakh-Khocho settlement, 72.95, 121.67, 1956, *Tolmachev, Polozova*), TC03-4 (1, DFO, Sakha Republic, Lena River, Tas-Azy Isl, 71.75, 127.00, 1956, *Norin, Petrovsky, Shtepa*), TC03-5 (1, DFO, Sakha Republic, Lena River, Sietachar River mouth, 71.08, 127.50, 1956, *Norin, Petrovsky, Shtepa*), TC03-6 (1, DFO, Sakha Republic, Tiksi harbour, Kengdey River basin, 71.58, 129.00, 1956, *Tolmachev, Yurtsev*), TC03-7 (1, SFO, Buryat Republic, Stanovoe Mts, South Muysky Ridge, Kindikan River source, 56.00, 115.00, 1965, *Petrochenko*), TC06-100 (1, ALA, North Slope Borough, Table Mountain Quad, Ambresvajun Lake, Last Lake, 68.60, –143.75, 1975, *Batten, Batten*), TC06-101 (1, ALA, North Slope Borough, Demarcation Point Quad, Beaufort Lagoon, Nuvagapak Point, 69.88, –142.30, 1974, *Murray, Batten*), TC06-102 (1, ALA, North Slope Borough, Demarcation Point Quad, Arctic National Wildlife Range, Pingokraluk Lagoon, Raluk, 69.70, –141.52, 1970, *Murray*), TC06-103 (1, ALA, North Slope Borough, Mount Michelsen Quad, Marsh Creek, app. 15 mi. inlan, 69.79, –144.82, 1985, *Lipkin*), TC06-104 (1, ALA, North Slope Borough, Table Mountain Quad, 32 km N of Ambresvajun Lake (Last Lake), Sheenjek River floodplain, 68.83, –143.50, 1975, Batten, Batten), TC06-105 (1, CAN, Yukon Territory, Eagle River Quad, Rock River, 66.87, –136.38, 1978, *Russell*); ***C. purpurea*:** SUP02-212 (4, ALA, Nome Census Area, Seward Peninsula, Kigluaik Mts, Mt E of Shaw Creek, 64.92, –164.97, 2002, *Elven, Gabrielsen, Jørgensen*), SUP02-225 (5, ALA, Nome Census Area, Seward Peninsula, Teller Road, Mt W of Penny River, 64.63, –165.68, 2002, *Elven, Gabrielsen, Jørgensen*), SUP03-129 (5, CAN, Yukon Territory, Ogilvie Mts C, Dempster hwy, km 91. Seepage on alpine mountain slope, 64.63, –138.37, 2003, *Elven, Solstad*), SUP03-16 (5, ALA, Northwest Arctic Borough, Noatak Quad, Igichuk Hills, Kaksurok Mt, N side of mountain, 67.21, –163.22, 2003, *Parker, Elven, Solstad*), SUP03-373 (5, CAN, Yukon Territory, N Ogilvie Mts, Ogilvie River at confluence with Engineer Creek, Dempster hwy 196 km, 65.39, –138.27, 2003, *Elven, Solstad*), SUP03-382 (2, CAN, Yukon Territory, N Ogilvie Mts, steep limestone mountain S of W end of Windy Pass, Dempster hwy 157 km, 65.07, –138.33, 2003, *Solstad, Elven*), TC03-12 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl., Draga harbour, 71.25, –179.67, 1954, *Sey*), TC03-16 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl., Draga harbour, 71.25, –179.67, 1970, *Petrovsky*, 70-300, $2n = 96$[6]), TC03-17 (1, DFO, Chukotka Autonomous Okrug, Wrangel Isl., Red Flag R., 71.25, –179.67, 1979, *Petrovsky*, 79-116, $2n = 96$[4]), TC06-106 (1, ALA, North Slope Borough, Point Hope Quad., Ogotoruk Creek, Headwaters of Snowbank Creek, 68.12, –165.78, 1980, *Murray, Johnson*), TC06-107 (1, ALA, North Slope Borough, Misheguk Mountain Quad., Noluck Lake, Storm Creek, 68.80, –160.00, 1972, *Parker*)

IV

Schmickl et al. BMC Evolutionary Biology 2010, 10:98
http://www.biomedcentral.com/1471-2148/10/98

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier

Roswitha Schmickl[1], Marte H Jørgensen[2], Anne K Brysting[2], Marcus A Koch[1*]

Abstract

Background: The genomes of higher plants are, on the majority, polyploid, and hybridisation is more frequent in plants than in animals. Both polyploidisation and hybridisation contribute to increased variability within species, and may transfer adaptations between species in a changing environment. Studying these aspects of evolution within a diversified species complex could help to clarify overall spatial and temporal patterns of plant speciation. The *Arabidopsis lyrata* complex, which is closely related to the model plant *Arabidopsis thaliana*, is a perennial, outcrossing, herbaceous species complex with a circumpolar distribution in the Northern Hemisphere as well as a disjunct Central European distribution in relictual habitats. This species complex comprises three species and four subspecies, mainly diploids but also several tetraploids, including one natural hybrid. The complex is ecologically, but not fully geographically, separated from members of the closely related species complex of *Arabidopsis halleri*, and the evolutionary histories of both species compexes have largely been influenced by Pleistocene climate oscillations.

Results: Using DNA sequence data from the nuclear encoded cytosolic phosphoglucoisomerase and Internal Transcribed Spacers 1 and 2 of the ribosomal DNA, as well as the *trn*L/F region from the chloroplast genome, we unravelled the phylogeography of the various taxonomic units of the *A. lyrata* complex. We demonstrate the existence of two major gene pools in Central Europe and Northern America. These two major gene pools are constructed from different taxonomic units. We also confirmed that *A. kamchatica* is the allotetraploid hybrid between *A. lyrata* and *A. halleri*, occupying the amphi-Beringian area in Eastern Asia and Northern America. This species closes the large distribution gap of the various other *A. lyrata* segregates. Furthermore, we revealed a threefold independent allopolyploid origin of this hybrid species in Japan, China, and Kamchatka.

Conclusions: Unglaciated parts of the Eastern Austrian Alps and arctic Eurasia, including Beringia, served as major glacial refugia of the Eurasian *A. lyrata* lineage, whereas *A. halleri* and its various subspecies probably survived in refuges in Central Europe and Eastern Asia with a large distribution gap in between. The North American *A. lyrata* lineage probably survived the glaciation in the southeast of North America. The dramatic climatic changes during glaciation and deglaciation cycles promoted not only secondary contact and formation of the allopolyploid hybrid *A. kamchatica*, but also provided the environment that allowed this species to fill a large geographic gap separating the two genetically different *A. lyrata* lineages from Eurasia and North America. With our example focusing on the evolutionary history of the *A. lyrata* species complex, we add substantial information to a broad evolutionary framework for future investigations within this emerging model system in molecular and evolutionary biology.

* Correspondence: mkoch@hip.uni-hd.de
[1]Heidelberg University, Heidelberg Institute of Plant Sciences, Biodiversity and Plant Systematics, Im Neuenheimer Feld 345, D-69120 Heidelberg, Germany

Background

Molecular biological research during the last decade has largely focussed on model organisms such as *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Arabidopsis thaliana*. Now that knowledge in molecular genetics, cell and developmental biology of these organisms has greatly increased, closely related organisms emerge as promising for studying characteristics not possible to elucidate with model and/or single organisms [1,2].

Arabidopsis lyrata L. is a close relative of *A. thaliana*, from which it diverged approximately five million years ago [3,4]. *Arabidopsis lyrata* s.l. represents a small species complex of four species and several putative subspecies with a circumpolar arctic-alpine distribution (Additional file 1, Table S1). 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; [5]] [6]. Most members of the species complex are perennial diploid outbreeders ($2n = 2x = 16$), but also tetraploid cytotypes occur [7,6,8]. There are numerous aspects of the biology of the *A. lyrata* complex that differ from and cannot be addressed in *A. thaliana*, like self-incompatibility and perennial life cycle.

The *A. lyrata* complex has already proven to be a suitable study system for the analysis of character traits such as flowering time [9,10] or pathogen defense [11]. Additionally, molecular mechanisms for the function of sporophytic self-incompatibility have been investigated [12-20], and comparative approaches to analyse sporophytic self-incompatibility in diploids versus polyploids are underway (Jørgensen, unpublished data). Whole genome sequencing of *A. lyrata* was finished last year, and data have been available for a few months (The *A. lyrata* genome sequence assembly v1.0, http://genome.jgi-psf.org/Araly1/Araly1.info.html), enabling direct comparisons with the *A. thaliana* genome. However, in contrast to *A. thaliana*, where the evolutionary history has been analysed in more detail [e.g. [21,22]], evolutionary studies on the *A. lyrata* complex have so far been largely restricted to a small number of populations from Central Europe [23-26] or larger sample sizes with a more general genus-wide perspective [27,28]. A detailed European study has revealed that population structure is dominated by regional genetic bottlenecks, and genetic structure exists within continents [Ansell, personal communication]. This suggests a comprehensive global study is necessary to resolve the evolutionary history of this complex.

In this study we present the first worldwide evolutionary history of the *A. lyrata* complex, covering its whole range of distribution and all taxonomically defined units. We use a widely applied nuclear encoded marker system (ITS, internal transcribed spacer region of nuclear encoded ribosomal DNA) to study gene flow between populations, 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) to investigate migrational movements due to seed dispersal, and the nuclear encoded housekeeping gene *Pgi*C (cytosolic phosphoglucoisomerase), a single copy gene, to discriminate between hybridising taxa. We aim to focus on the following four aspects: (1) Unravelling general phylogeographic patterns of the *A. lyrata* complex by identifying the main genetic lineages, and interpreting genetic variation in space and time [29], in the context of both climatic and geological events throughout Pleistocene glaciation cycles; (2) Evaluating the role of hybridisation and polyploidisation in the origin of *A. kamchatica* (Fisch. ex DC.) K. Shimizu & Kudoh, an amphi-Beringian member of the *A. lyrata* complex; (3) Explaining Pleistocene and postglacial migration routes by analysing genetic diversity statistics: The arctic-alpine *A. lyrata* complex is one of the rare examples among higher plants with a distribution in both Central Europe and North America and, additionally, a circumpolar distribution - other examples are *Cassiope tetragona* [30], and *Saxifraga oppositifolia* [31]; and (4) Studying the role of Beringia as a refugia for populations of arctic *A. lyrata*, since Beringia is assumed to be one of the major refugia for arctic plants during Pleistocene glaciations [30-36].

Methods

Plant material

Altogether 467 accessions of the *A. lyrata* complex were analysed: 295 accessions newly analysed within this study, 39 accessions sequenced by Schmickl et al. [28], and 133 accessions analysed by Koch and Matschinger [27]. 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), LE (The V.L. Komarov Botanical Institute, Russian Academy of Sciences, St. Petersburg), LI (Upper Austrian Provincial Museum, Linz), O (Natural History Museum, University of Oslo, Oslo), W (Natural History Museum, Vienna), and partly collected in the field, documented at HEID (Herbarium University of Heidelberg, Heidelberg). Taxon determination followed the voucher labels, and was verified with

floras and determination keys [e.g. [37]; Flora of North America, Al-Shehbaz, personal communication]. In our study we followed the taxonomy of [37]. Twentyfive accessions of the various subspecies of *A. halleri* (L.) O'Kane & Al-Shehbaz were analysed from throughout the distribution range because of the evidence that this taxon served as a putative parent of *A. kamchatica*, previously also treated as *A. lyrata* ssp. *kamchatica* [[27]; and references therein], an evolutionary scenario which was recently confirmed by [8]. The distribution of the investigated accessions is shown in Figure 1. The accession list is provided with Additional file 2, Table S2.

The following short overview of *A. lyrata* and *A. halleri* taxonomy will introduce the concept of Al-Shehbaz and O'Kane [37]. Elven (ed.) [38] persues a different taxonomic concept, which is summarised in the supplementary material (Additional file 1, Table S1), but will not be discussed here. Apart from *A. thaliana*, the *A. lyrata*

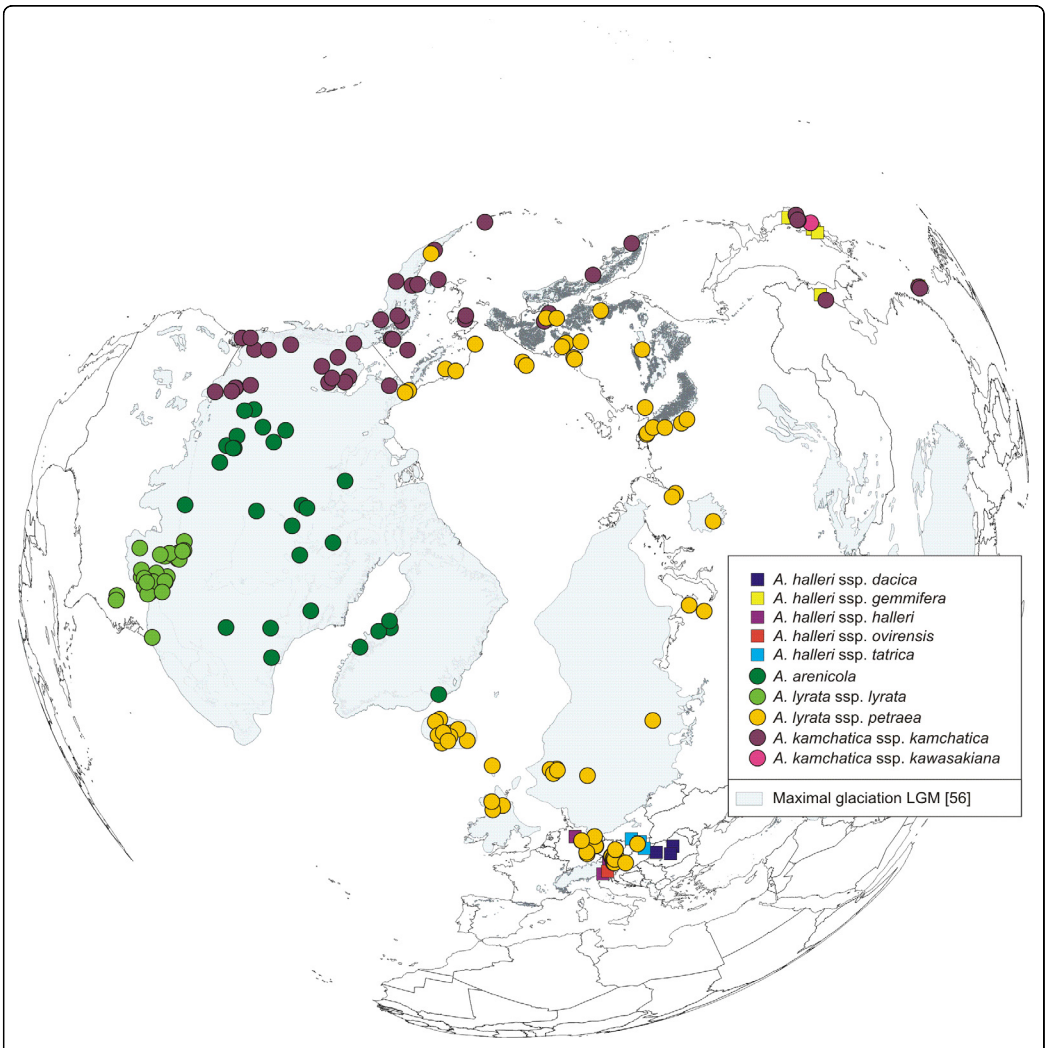


Figure 1 Distribution of *Arabidopsis* accessions investigated. Accessions of the *Arabidopsis lyrata* complex were analysed for nuclear ITS and *Pgi*C regions, and cpDNA *trn*L/F. Accessions of the *Arabidopsis halleri* complex were analysed for *Pgi*C only. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard [56].

complex is one of three major species complexes in the genus *Arabidopsis* [27], the other two being the *A. arenosa* (L.) Lawalrée and *A. halleri* complexes [39,40,6,27,41]. The *A. lyrata* complex is considered by different authors to include different numbers of taxa of various ranking, distribution areas, and ploidy levels (Additional file 1, Table S1). Al-Shehbaz and O'Kane [37] treat the complex as one species, *A. lyrata*, with three subspecies: ssp. *lyrata*, ssp. *petraea* (L.) O'Kane & Al-Shehbaz, and ssp. *kamchatica* (Fisch. ex DC.) O'Kane & Al-Shehbaz. Subspecies *lyrata* is considered to be broadly amphi-Pacific and of two ploidy levels ($2n$ = 16, 32), ssp. *petraea* to be Northern Eurasian and Central European with the same two ploidy levels, and *kamchatica* to be amphi-Pacific and tetraploid ($2n$ = 32). A comprehensive quantitative morphological analysis using multivariate statistics is still not available for the *A. lyrata* complex. Moreover, information on the ploidy level is largely lacking throughout the distribution range, and further cytological investigations might be important for taxon delimitation [7,8]. *Arabidopsis arenicola* (Richardson) Al-Shehbaz, Elven, D.F. Murray & Warwick has only recently been included as part of the *A. lyrata* complex [42]. It has, for a long time, been placed within the genus *Arabis* L. Elven (ed.) [38] considers 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 [Al-Shehbaz, personal communication]. Many of the mentioned taxa include two ploidy levels, suggesting frequent polyploidisation events within the *A. lyrata* complex. An allopolyploid origin of *kamchatica* has already been confirmed based on nuclear DNA sequences, with *A. lyrata* and *A. halleri* ssp. *gemmifera* (Matsum.) O'Kane & Al-Shehbaz as possible parental taxa [[43,27,28,8]; Jørgensen et al., unpublished data]. Otherwise, little is known with regard to the number of polyploid units and their origins.

Five subspecies have been recognised in *A. halleri*: ssp. *halleri*, ssp. *ovirensis* (Wulfen) O'Kane & Al-Shehbaz, ssp. *dacica* (Heuff.) Kolník, comb. nov., ssp. *tatrica* (Pawl.) Kolník, comb. nov., all distributed in Central Europe, and ssp. *gemmifera* in Eastern Asia, supported by both morphometric analysis (Kolnik, unpublished data) and genetic AFLP data (Marhold, unpublished data).

DNA isolation, amplification and sequencing

Total DNA was obtained from dried leaf material and extracted according to the CTAB protocol of Doyle and Doyle [44] 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 RNase A 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), primers and PCR cycling scheme followed the protocol of Dobeš et al. [45], using a PTC200 (MJ Research, Waltham, USA) thermal cycler. The PCR reaction volume of 50 μl contained 1× 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 Biosciences, Chalfont St Giles, England), 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 Dobeš et al. [46]. 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.

Before sequencing PCR products were checked for length and concentrations on 1.5% agarose gels and purified with the NucleoFast Kit (Macherey-Nagel, Düren, Germany). Cycle sequencing was performed using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences, Chalfont St Giles, England) 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.

trnL/F and ITS sequence definition and map reconstruction

Plastidic *trn*L/F sequences were defined as haplotypes and suprahaplotypes following our previous studies [27,41,28]: Haplotypes are characterised by multiple *trn*F pseudogenes in the 3'-region of the *trn*L/F-IGS close to the functional *trn*F gene [47,48,27,49,41,50]. When defining *trn*L/F suprahaplotypes, we excluded the pseudogene-rich region. Pseudogenes evolve with a 10 times higher mutation rate than single nucleotide polymorphisms, which makes them non-applicable for phylogenetic reconstructions on the species level [50]. Additionally, their boundaries are under ongoing discussion. An alternative interpretation is provided by Ansell et al. [51]. In summary, haplotypes belonging to one suprahaplotype share the same base order throughout the whole sequence except for the pseudogene-rich region, where they vary in both length and base content. Suprahaplotypes differ from each other only by single point mutations and/or indels. Newly defined *trn*L/F haplotypes were assigned to GenBank numbers [GQ922894-GQ922903] (Additional file 2, Table S2).

ITS sequences were obtained from direct sequencing of PCR products in order to detect hybrid individuals according to ambiguous base positions. Sequences were defined as ITS types and supratypes following our previous studies [27,41,28]: Most ITS types contained ambiguous sites as a result of the direct sequencing approach. They were combined to ITS supratypes by replacing the ambiguous sites by the bases with higher fluorescence intensity in the electropherogramm. Sequences with equal fluorescence intensity of the two bases at the ambiguous positions were only found between ITS supratypes b and e and labelled b/e ambiguous. Koch and Matschinger [27] already showed that analyses with either ITS types or ITS supratypes produce the same results. However, working with a limited number of ITS supratypes in contrast to a vast number of ITS types contributes to a clearer data display. Multiple ITS types within a single individual indicate the natural variability among ITS loci within an individual as result of either mutations or gene flow between individuals. In these individuals concerted evolution, which denotes the process of DNA sequence homogenisation among members of multigene families by gene conversion and/or unequal crossing over [52-55], is not yet completed. Newly defined ITS types were assigned to GenBank numbers [GQ922904-GQ922910] (Additional file 2, Table S2).

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 (Last Glacial Maximum) were provided by Ehlers and Gibbard [56].

Network analyses and genetic diversity statistics

Network analyses and genetic diversity statistics were exclusively performed using the *trn*L/F suprahaplotypes, as the pseudogene-rich region is not applicable for phylogenetic reconstructions on the species level [50]. The network was constructed using TCS version 1.21 [57] using the statistical parsimony algorithm [58]. Gaps (except polyT stretches) were coded as single additional binary characters. Newly characterised suprahaplotypes were added to the *trn*L/F network of the *A. lyrata* complex published earlier [28]. Genetic diversity statistics were performed with Arlequin version 3.11 [59]. Pairwise genetic differentiation was calculated among the following nine taxonomic and regional groups: *A. lyrata* ssp. *petraea* from (1) unglaciated Central Europe, (2) glaciated northern Europe, (3) northern Russia and western Beringia, (4) eastern Beringia; *A. lyrata* ssp. *lyrata* from (5) unglaciated North America and the glaciated Great Lakes region; (6) *Arabidopsis arenicola* from glaciated North America and Greenland; and *A. kamchatica* from (7) Japan, (8) north-eastern Russia, (9) Alaska and western Canada. F_{ST} values, regarding haplotype frequencies only, and Φ_{ST} values, which take into account the genetic relationship between haplotypes as pairwise character differences, were calculated. The significance of differentiation was examined using a permutation test with 1000 permutations. Additionally, genetic diversity was estimated as nucleotide diversity π.

ITS parsimony analysis

ITS data were analysed based on a 652 bp alignment of ITS supratypes [27]. The total number of variable sites was 44, with 16 of them parsimony informative. From altogether 10 supratypes of both the *A. lyrata* complex and *A. halleri* ssp. *gemmifera*, a strict consensus tree was constructed using maximum parsimony with MEGA version 4.1 [60]. Heuristic searches were performed with 10 random addition sequences and Closest Neighbour Interchange (CNI) branch swapping. Bootstrap values were calculated based on 500 replicates. Length of the most parsimonious trees was 48 mutational steps with a consistency index (CI) = 0.89 and a retention index (RI) = 0.96 (autapomorphies excluded). *Arabidopsis thaliana* was used as outgroup.

Primer design for the nuclear marker *PgiC*

Various higher plants are known to have a duplicated locus of the cytosolic enzyme phosphoglucoisomerase [61-63], and the loci are normally unlinked. Also within the genus *Arabidopsis*, extensive sequencing of the *Pgi*C locus using a cloning strategy revealed a duplication, both in the *A. lyrata* and *A. halleri* complex, and this duplication must have predated the evolutionary split between these two species complexes (Additional file 3, Figure S1; Jørgensen, unpublished data). Both loci were initially simultaneously amplified with the general forward primer 5'-TGCTGTSAGCACTAATCTTGCG-3' and the general reverse primer 5'-TCGAACCCGGGAGAGGTAGACCA-3', following the protocol of Wright et al. [23]. The resulting sequence data showed that a group of alleles at the *Pgi*C1 locus were exclusively found in *A. halleri* and the allopolyploid *A. kamchatica*. It was thus possible to design *A. halleri*-specific primers that worked as a high-throughput and simple PCR-based screening marker to discriminate between genomes of the *A. lyrata* and *A. halleri* complex. Unfortunately, it was not possible to develop a PCR-based reciprocal marker system characterising alleles from the *A. lyrata* gene pool because of a lack of appropriate DNA sequence variation. In general, the alleles from both duplicated *Pgi*C loci were only weakly differentiated among and within species. However, within the *Pgi*C1 locus we found a deletion of 7 bp length in the *A. lyrata* complex compared to the *A. halleri* complex (Additional file 4, Figure S2). Both groups of alleles

were also substantially differentiated by various SNPs. A primer pair with the forward primer located partly within this 7 bp indel (5'-CATTCAACAGATTGTG-3') and the reverse primer 5'-CCAGTAAACATCATGT-3' was developed to amplify a 92 bp fragment within the *Pgi*C1 locus (Additional file 5, Figure S3). The PCR reaction volume of 50 μl contained 1× 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 Biosciences, Chalfont St Giles, England), 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. By screening the absence/presence of the PCR product of this taxon-specific primer pair, we were able to follow the genetic footprint of the *A. halleri* complex in its allopolyploid hybrids throughout its distribution range.

Results

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

The investigated accessions (Figure 1), spanning the whole distribution range of the *A. lyrata* complex, were grouped into three genetic and geographically separated lineages: *trn*L/F suprahaplotype C in Eurasia, A in North America, and B in the amphi-Pacific region (Figure 2). The Eurasian lineage with suprahaplotype C had the largest distribution range: unglaciated Central Europe, formerly 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 (to the Canadian Rocky Mountains in the west), and Greenland. The third and amphi-Pacific lineage spanned from Kamchatka via Beringia into western Canada (with the Canadian Rocky Mountains as eastern border). These three suprahaplotypes were central in the suprahaplotype network (Figure 3), and all were connected to derived and less widely distributed suprahaplotypes: In the Eurasian lineage AC was widespread and found in Central Europe, mainly Austria, AG predominantly in the north (Iceland, Scandinavia, Russia), and AR in western Beringia (incl. Wrangel Island) (Figure 2). 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), Faeroe Islands (G), and Sweden (S). The North American lineage was additionally characterised by the unique suprahaplotypes AQ and BD, and the more widespread BF. Within 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. suprahaplotypes A and B were detected in a few accessions of the Eurasian lineage (Figure 2, Additional file 6, Table S3). This finding is congruent with the observation of central suprahaplotype sharing between the three main species complexes of the genus, *A. lyrata*, *A. halleri*, and *A. arenosa* [27]. This observation has been explained by ancestral cpDNA polymorphism predating the radiation of the genus approximately two million years ago [27].

Cytosolic phosphoglucose isomerase identifies *Arabidopsis halleri* ssp. *gemmifera* as one parent of the allopolyploid amphi-Pacific *Arabidopsis kamchatica*

The various *Pgi*C alleles detected were defined either as alleles at locus *Pgi*C1 or *Pgi*C2, and, therefore, considered as locus-specific (Additional file 3, Figure S1). However, taxon-specific lineage sorting of the various allele pools after the duplication event was not complete for *Pgi*C2. *A. kamchatica* carries *Pgi*C2 alleles hardly distinguishable from those of *A. septentrionalis* and *A. umbrosa*, which is an additional indicator that an Asian member of the *A. lyrata* complex served as one putative parental taxon. In addition, European *A. lyrata* ssp. *petraea* shares similar alleles with *A. halleri*, also indicating incomplete lineage sorting. The differentiation and lineage sorting of alleles at locus *Pgi*C1 is more taxon-specific. Here, *A. kamchatica* shares alleles most similar to those of *A. halleri* ssp. *gemmifera* from East Asia, and all of these alleles are significantly distinct from those of Eurasian *A. lyrata*. In summary, it is shown that *Pgi*C is not only a suitable marker to screen for hybrid speciation in *A. kamchatica*, but might also be a suitable marker to follow these alleles through space and time. Amplification of *Pgi*C1 alleles was successful in all *A. kamchatica* accessions (Figure 4, Additional file 5, Figure S3), but, as outlined above, failed in Eurasian and North American members of the *A. lyrata* complex (Figure 4, Additional file 5, Figure S3). *Pgi*C1 amplification was positive in all *A. halleri* ssp. *gemmifera* accessions and additionally in European subspecies of *A. halleri* (ssp. *dacica*, ssp. *halleri*, ssp. *tatrica*) (Figure 4, Additional file 5, Figure S3). Hence, presence of *Pgi*C1 alleles without deletion in the forward primer sequence was characteristic for the whole *A. halleri* species complex, except for *A. halleri* ssp. *ovirensis* (data not shown). In this subspecies either a secondary loss of this locus or a complementary mutation in the primer binding site might have occured. However, *A. halleri* ssp. *ovirensis* is a genetically distinct, local endemite at one single place in the southeastern Austrian Alps with an unclear evolutionary history [27,41].

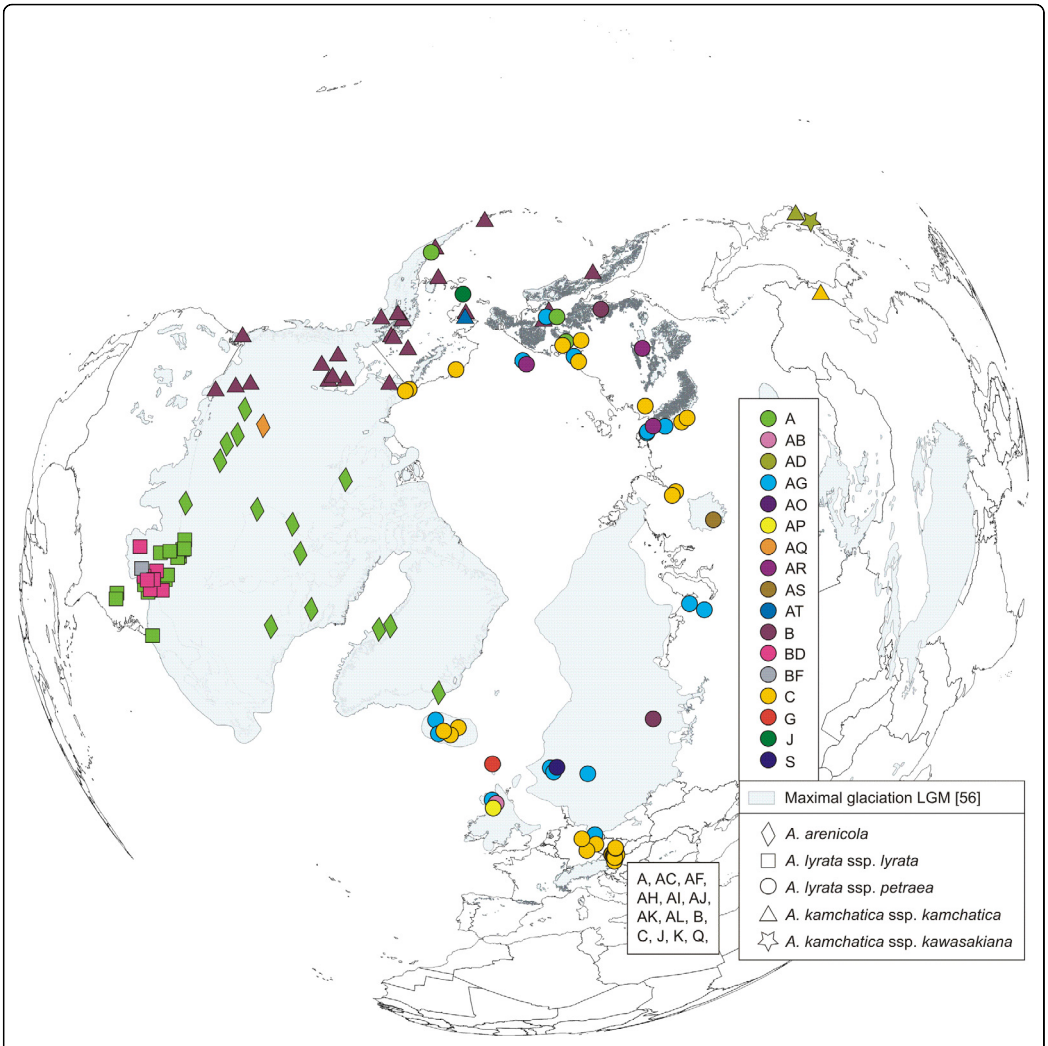


Figure 2 Distribution of cpDNA *trn*L/F suprahaplotypes in the *Arabidopsis lyrata* complex. Data newly presented in this study were combined with previous results from Koch and Matschinger [27] and Schmickl et al. [28]. *Trn*L/F suprahaplotypes were characterised as *trn*L intron and *trn*LF-IGS excluding the pseudogene-rich region in the *trn*LF-IGS. Accessions from Austria are listed separately. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard [56].

Gene diversity statistics show 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 of the *A. lyrata* species complex was divided into nine different groups, according to the evidence for genetic lineages and Pleistocene history (Table 1). The Eurasian lineage (*A. 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 (*A. lyrata* ssp. *lyrata*) was separated into unglaciated North America/glaciated Great Lakes region and glaciated North America/Greenland. Amphi-Pacific *A. kamchatica* was differentiated into groups from Japan,

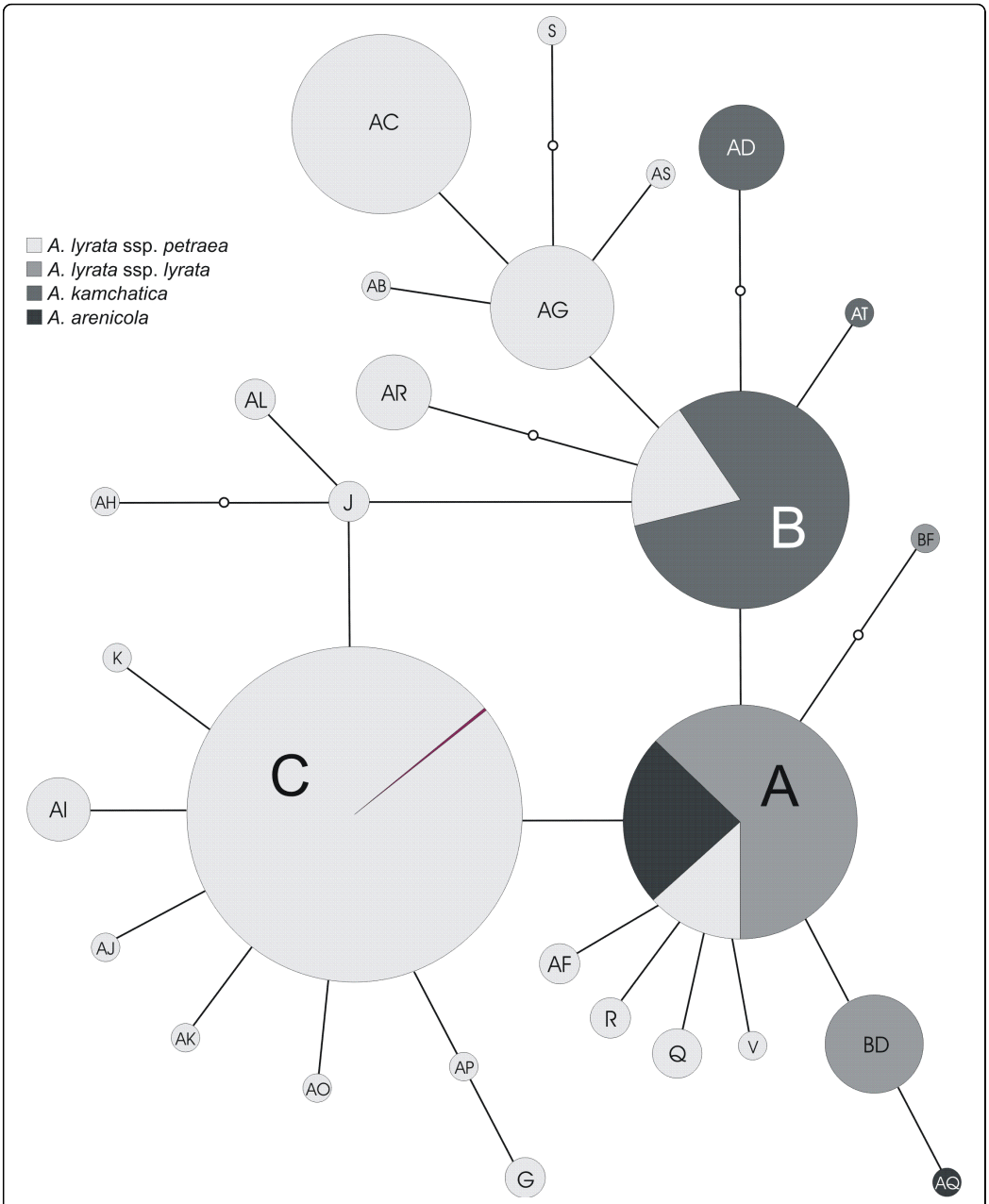


Figure 3 CpDNA *trn*L/F suprahaplotype network of the *Arabidopsis lyrata* complex. Data newly presented in this study were combined with previous results from Koch and Matschinger [27] and Schmickl et al. [28]. *Trn*L/F suprahaplotypes were characterised as *trn*L intron and *trn*LF-IGS excluding the pseudogene-rich region in the *trn*LF-IGS. The sizes of the circles indicate the relative frequency of a suprahaplotype in the whole dataset.

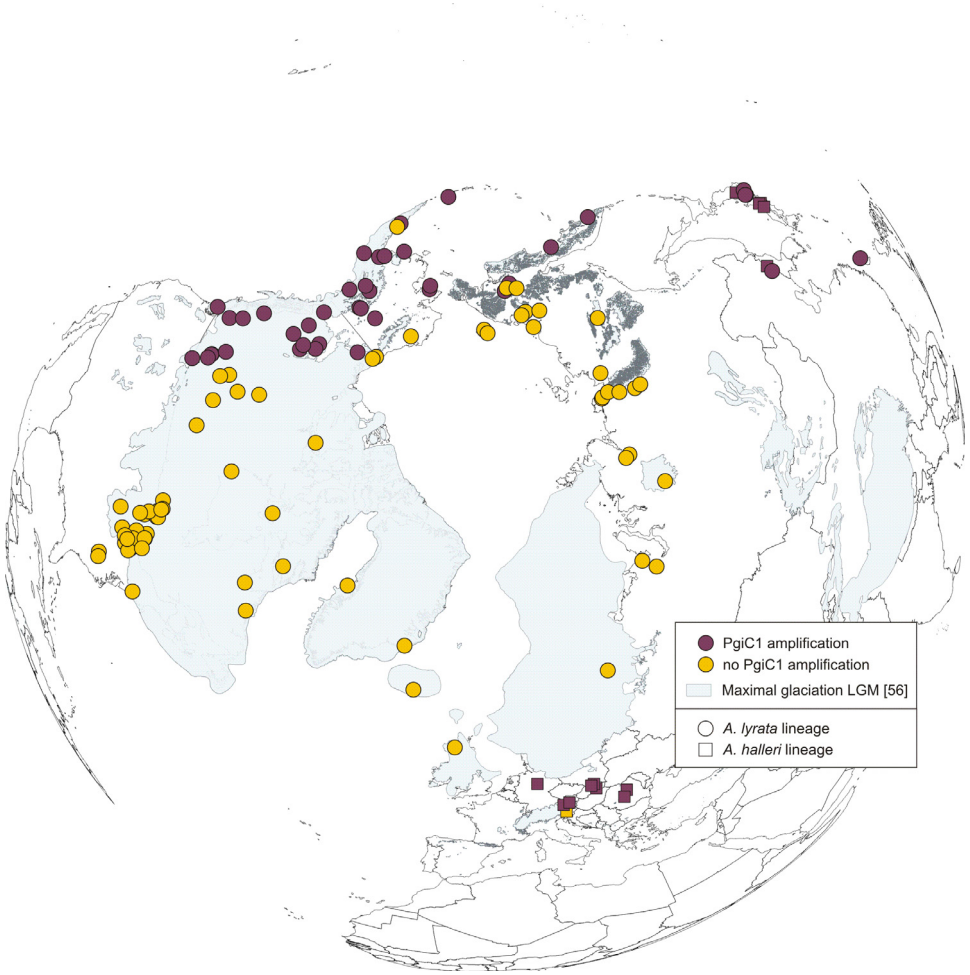


Figure 4 Distribution of accessions with/without *Pgi*C1 amplification in the *Arabidopsis lyrata* complex and *A. halleri*. Amplification was successful in *Arabidopsis halleri* ssp. *gemmifera* from China and Japan, and in Central European subspecies of the *Arabidopsis halleri* complex (ssp. *dacica*, ssp. *halleri*, and ssp. *tatrica*). Amplification was also successful in all *A. kamchatica* accessions. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard [56].

Far Eastern Federal District of Russia, and Alaska/western Canada. The Eurasian lineage of *A. lyrata* ssp. *petraea* showed the highest cpDNA-based (*trn*L/F) nucleotide diversity (π), and it was equally high in unglaciated Central Europe (π = 0.0060), glaciated northern Europe (π = 0.0076), and northern Russia/western Beringia (π = 0.0077), but reduced in eastern Beringia (π = 0.0034) (Table 1). Pairwise Φ_{ST} and F_{ST} calculations showed only slight differentiation among unglaciated Central Europe, glaciated northern Europe, and northern Russia/western Beringia. For eastern Beringia the data were not statistically significant ($P < 0.05$; Table 2). These observations could indicate long-term survival of the *A. lyrata* complex in unglaciated Central Europe and northern Russia/western Beringia, or post-glacial 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 (Figure 2, Additional file 6, Table S3). In Central Europe

Table 1 Taxonomic and regional genetic differentiation based on cpDNA suprahaplotypes.

Geographic region	*n*	$\pi \times 10^{-2}$
A. lyrata ssp. *petraea*: Unglaciated Central Europe	196	0.603 +/- 0.329
A. lyrata ssp. *petraea*: Glaciated N Europe	21	0.764 +/- 0.425
A. lyrata ssp. *petraea*: N Russia, W Beringia	33	0.774 +/- 0.422
A. lyrata ssp. *petraea*: E Beringia	7	0.339 +/- 0.236
A. lyrata ssp. *lyrata*: Unglaciated N America, glaciated Great Lakes region	55	0.080 +/- 0.072
A. arenicola: Glaciated N America, Greenland	17	0.031 +/- 0.042
A. kamchatica: Japan	9	0.000 +/- 0.000
A. kamchatica: Far Eastern Federal District of Russia	9	0.000 +/- 0.000
A. kamchatica: Alaska, W Canada	38	0.007 +/- 0.019

Sample size (n) and nucleotide diversity (π) are provided with the corresponding standard deviation.

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. A high number of unique and rare suprahaplotypes was also found in formerly glaciated northern Europe (AB, AP - Scotland, AO - Iceland, G - Faeroe Islands, S - Norway). Subsequent geographic isolation of these populations during the Holocene warming might have caused restriction of suprahaplotypes to single geographic locations.

In contrast to the Eurasian lineage of *A. lyrata* ssp. *petraea*, the North American lineage showed an approximately tenfold reduction in nucleotide diversity and strong differentiation according to pairwise Φ_{ST} and F_{ST} (Table 2). Nucleotide diversity of accessions from predominantly unglaciated southeastern North America ($\pi = 0.0008$) was higher than from North America and Greenland ($\pi = 0.0003$) (Table 1), which had been under the Laurentide ice sheet during the LGM, possibly indicating genetic bottlenecks with subsequent rapid postglacial immigration.

Extremely reduced nucleotide diversity was reported from amphi-Pacific *A. kamchatica* ($\pi = 0.0000$) (Table 1). When all three genetic groups of *A. kamchatica* were treated separately, only a single *trn*L/F suprahaplotype was found in each group (Japan: AD; China: C; Far Eastern Federal District of Russia/Alaska/western Canada: B) (Additional file 6, Table S3). Additionally, haplotype diversity was low, especially in *A. kamchatica* with suprahaplotype B. Only one haplotype (no. 84) was detected over a vast amphi-Pacific area from Kamchatka to western Canada (Additional file 7, Table S4).

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 (reviewed by Abbott and Brochmann [64] and DeChain [65]), as it remained ice-free during Pleistocene climate oscillations. If we consider only the Eurasian and North American lineage of the *A. lyrata* complex, two major ITS groups met in Beringia (Figure 5): (1) the mainly Eurasian group carrying ITS supratype b (Figure 6), comprising Europe (with additional ITS supratypes a, c, d), northern Russia and western Beringia, and (2) the North American group carrying ITS supratypes e and, extremely rarely, y (Figure 6), including eastern and central North America, Greenland, and eastern Beringia (north of Brooks Range). The main contact zone was 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 most likely due to ancient and/or recent gene flow from populations of the North American lineage into populations of the Eurasian lineage. Throughout the Beringian area ambiguous sites in ITS DNA sequences, caused by multiple ITS copies within a single genome and incomplete concerted evolution [54,28], were mainly found between ITS supratypes b (Eurasian lineage) and e (North American lineage). These results indicate that gene flow between these genetic groups may be counteracting the effects of concerted evolution. Interestingly, the allopolyploid amphi-Pacific lineage (*A. kamchatica*) also showed ITS supratype b like the Eurasian lineage.

Discussion

High genetic diversity of the *Arabidopsis lyrata* complex in Eurasia - postglacial migration from Central European and northern Russianrefugia

For arctic-alpine taxa, centres of species and genetic diversity are, in most cases, considered to concur with Pleistocene refugia [64,66,67]. In the genus *Arabidopsis*, both the number of accepted taxa, *trn*L/F suprahaplotype and ITS supratype diversity are highest in Central Europe, indicating this area as a centre of diversity [37,27]. Especially the Eastern Alps could have served as

Table 2 Pairwise genetic differentiation (F_{ST} and Φ_{ST}) among taxonomic and regional groups.

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.*: Far Eastern Federal District of Russia	*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.: Far Eastern Federal District of Russia	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]	

F_{ST} (above 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).

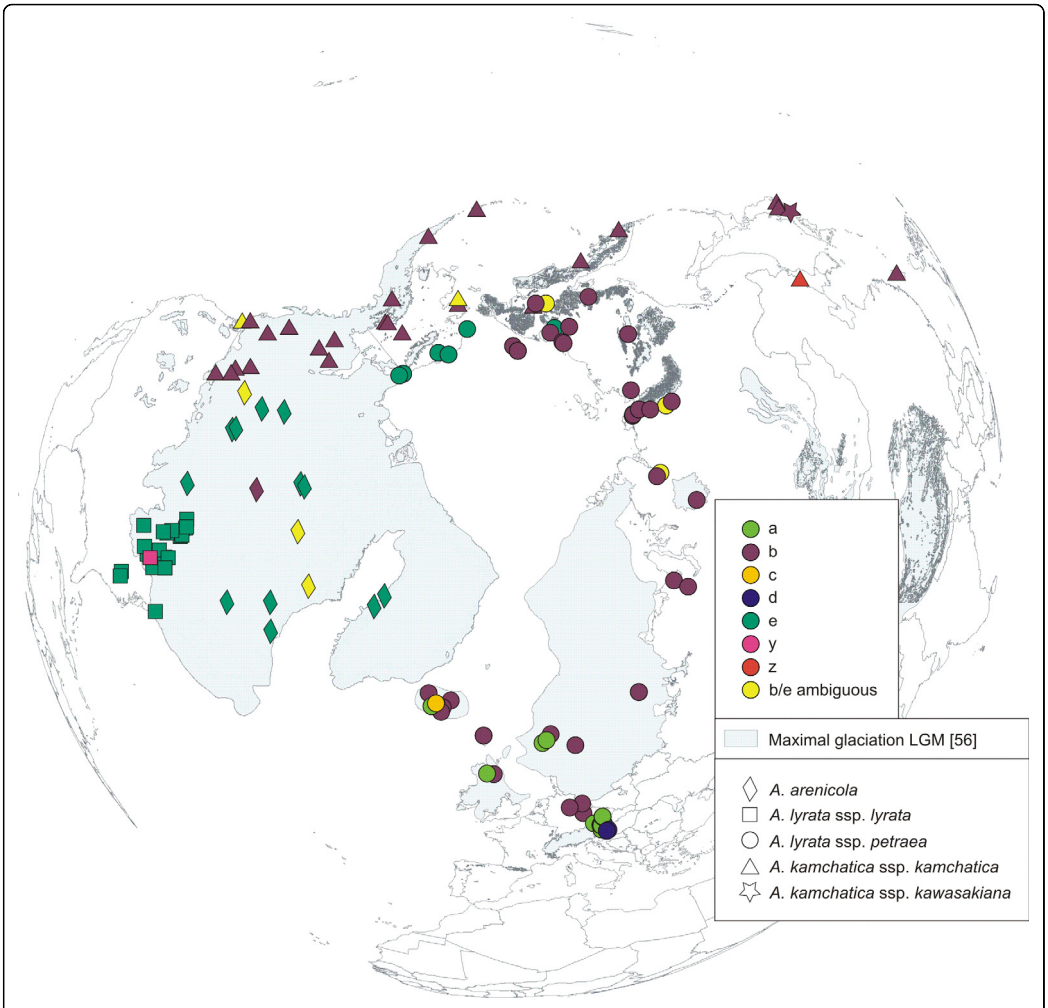


Figure 5 Distribution of ITS supratypes in the *Arabidopsis lyrata* complex. Data newly presented in this study were combined with previous results from Koch and Matschinger [27] and Schmickl et al. [28]. ITS supratypes were characterised as ITS1, 5.8 S rDNA, and ITS2 region, with ambiguous sites replaced by the bases with higher fluorescence intensity in the electropherogramm. Sequences with equal fluorescence intensity of the two bases at the ambiguous positions were only found between ITS supratypes b and e and labelled b/e ambiguous. The ITS supratypes are corresponding to those shown in Figure 6. Maximal glaciation of the LGM is drawn according to Ehlers and Gibbard [56].

a refugium of *A. lyrata* ssp. *petraea*, since both diploid and cytogenetically stabilised tetraploid populations occur there (Schmickl and Koch, unpublished data). For the *A. lyrata* complex, however, Koch and Matschinger [27] concluded that Central Europe is not the only centre of diversity. Our results support this conclusion, showing that the Eurasian lineage of the *A. lyrata* complex is genetically diverse in both Central Europe and arctic Eurasia, including Beringia. To our knowledge such a pattern has not been observed for any other arctic plant with additional distribution in Central Europe. Either genetic diversity is high in the Arctic compared to the Alps, as in *Saxifraga oppositifolia* L., suggesting long-term evolution in Beringia and more recent colonisation of the Alps [33,31], or genetic diversity is highest in the Alps and decreasing towards the Arctic, as observed in *Arabis alpina* L. [68,69], indicating recent and rapid colonisation from Central Europe. Recent

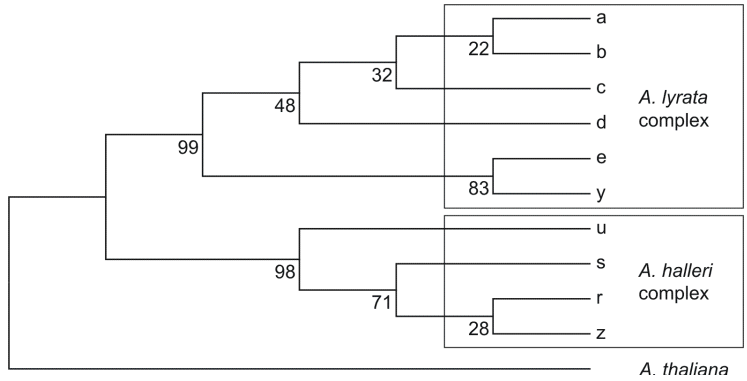


Figure 6 Strict consensus tree from the maximum parsimony analysis of ITS supratypes in the *Arabidopsis lyrata*complex and *Arabidopsis halleri* ssp. *gemmifera*. From altogether 10 ITS supratypes (with ambiguous sites replaced by the bases with higher fluorescence intensity in the electropherogramm) of both the *A. lyrata* complex and *A. halleri* ssp. *gemmifera*, a strict consensus tree (length = 43) was constructed using maximum parsimony with MEGA version 4.1 [60]. Heuristic searches were performed with 10 random addition sequences and Closest Neighbour Interchange (CNI) branch swapping. Bootstrap values were calculated based on 500 replicates. Consistency index (CI) = 0.89, retention index (RI) = 0.96. *Arabidopsis thaliana* was used as outgroup.

migration into the Arctic, associated with a loss of genetic diversity due to repeated bottlenecks, was also found in *Ranunculus glacialis* L. [70] and *Rubus chamaemorus* L. [71]. The overall high genetic diversity in the Eurasian *A. 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 *A. lyrata* was reported only from unglaciated Central Europe, based on microsatellite data [25] and cpDNA sequences [27]. This study showed high genetic diversity also in formerly glaciated northern Europe, caused by numerous 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, as reported from Ansell [personal communication]. Although several authors suggested periglacial survival on nunataks along the Norwegian coastline, mainly based on geomorphological investigations (reviewed by Brochmann et al. [34]), 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 [25]. Riihimäki and Savolainen [9] even found divergent Central and northern European physiological morphotypes, with earlier and more frequent flowering plants in the south, and later and rarer flowering plants in the north.

We observed high genetic diversity within the Eurasian *A. lyrata* lineage, but also some genetic homogenisation. The *trn*L/F suprahaplotypes C and AG were found all over the distribution range, which could be explained by repeated gene flow 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 have bridged distances up to 2000 km [72] or even 4000 km (Westergaard et al., in prep.) and seems to be frequent in the Arctic [34,73].

Ancient split of the Eurasian and North American lineage

The strong genetic differentiation we observed between the Eurasian and North American lineages is probably the result of long-term geographic isolation during Pleistocene glaciations. Muller et al. [26] detected genetic divergence between North American and western European populations in a comparative microsatellite study, and found that the former had a lower diversity. However, they focused on a few populations only and included only one North American population. The strong reduction of genetic diversity we observed in the North American lineage compared to the Eurasian lineage is congruent with nuclear and plastidic marker data of Wright et al. [23], but contradicting Balañα-Alcaide et al. [24], who reported

similar genetic diversity based on two nuclear markers. However, both studies included only two populations from each of North America and Europe, and they may not have been representative. Because of the strongly reduced genetic diversity of the North American lineage, we assume this lineage to be derived from the Eurasian lineage. Future studies of more southern populations from the USA will test this assumption. Migration into North America apparently was associated with a strong genetic bottleneck, which restricted populations to ITS supratype e and *trn*L/F suprahaplotype A. Colonisation of North America is possible from different directions, either from Europe, or from East Asia. Both possibilities have been discussed for various circumboreal species (reviewed by Abbott and Brochmann [64]). In the case of *A. lyrata*, colonisation of North America from Russia seems, however, most likely, as the North American genotype was also rarely detected in western Beringia, but not at all in Iceland or Scandinavia.

An amphi-Beringian *Arabidopsis* hybrid zone - due to allopolyploid success?

The majority of arctic polyploids have a history as post-glacial colonisers [74]. It has frequently been assumed that polyploids have a broader adaptive potential for recolonising formerly glaciated areas [75,76]. As one of the major refugia for arctic plants, Beringia has a higher proportion of diploids compared to areas glaciated during Pleistocene [74]. New adaptations in polyploids may evolve either by genome rearrangements [77] and/or epigenetic changes [78-80] within the first generations after polyploid formation, as reported from rapid gene silencing in the allopolyploid *A. suecica* (Fr.) Norrl. ex O.E. Schulz [78]. Hybridisation is frequently involved in polyploidisation, leading to the formation of allopolyploids with one each of the parental genomes. Allopolyploidisations have been reported for several arctic species complexes, such as the high polyploid *Cerastium alpinum* L. complex [81], high polyploid *Primula* sect. *Aleuritia* [82-84], tetra-/hexaploid *Silene* L. [85,86], tetra- to octoploid *Saxifraga* section *Mesogyne* Sternb. [87,88], and tetra- to dodecaploid *Cardamine digitata* Richardson [89]. Introgression, the integration of genetic material from one species into another through repeated backcrossing, was observed between polyploid *Saxifraga cernua* L. and diploid *Saxifraga sibirica* L. [90]. The first known *Arabidopsis* allopolyploid was *A. suecica* with the maternal parent *A. thaliana* and the paternal parent *A. arenosa*, confirmed by artificial crosses [78]. This species probably developed around 20 000 years ago [91] or between 20 000 and 300 000 years ago [92], with a single origin in Fennoscandia [91,92]. The distribution range of this mainly outcrossing [93] allopolyploid species is rather small. The second natural *Arabidopsis* allopolyploid, *A. kamchatica*, has *A. halleri* ssp. *gemmifera* and a member of the *A. lyrata* complex as parental taxa [[43,27,8]; Jørgensen, unpublished data]. This allopolyploid origin could be confirmed not only for Japanese, but for all *A. kamchatica* accessions. According to chloroplast *trn*L/F data, three different genetic groups were found, which are geographically isolated from each other: (1) accessions of a widespread distribution range from Kamchatka, 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 gene flow could be observed: Either the paternal genome originated from a member of the *A. lyrata* species complex (ITS supratype b) and the maternal genome from *A. halleri* ssp. *gemmifera* (*trn*L/F suprahaplotype AD), as already reported for *A. kamchatica* from Japan [27]. Or *A. halleri* ssp. *gemmifera* represented the paternal genome (ITS supratype z, derived from ITS supratype r exclusively found in *A. halleri* ssp. *gemmifera*), and a member of the *A. lyrata* species complex served as donor of the maternal genome (*trn*L/F suprahaplotype C), as in *A. kamchatica* from China. The North American lineage of the *A. lyrata* complex could be excluded as a parent, as neither ITS supratype e nor *trn*L/F suprahaplotype A were detected in *A. kamchatica*. According to these data we suggest at least three independent origins of *A. kamchatica*, first with maternal *A. halleri* ssp. *gemmifera* in Japan, second with paternal *A. halleri* ssp. *gemmifera* in China, and third with an unknown direction of gene flow in Kamchatka, but in all cases with a member of the Eurasian *A. lyrata* lineage as hybridisation partner.

The most profound change in *A. kamchatica* in contrast to its parental species is the switch from outcrossing (with sporophytic self-incompatibility system) to selfing [ssp. *kamchatica*: [15,8]; ssp. *kawasakiana*: [94]]. Selfing increases the possibility of rapid range expansion, as a population can arise from a single individual independent of pollinators and pollen donors. Such a switch is already well known from *A. thaliana*, dated from around 413 000 [95] to one million years ago [96], and is therefore not necessarily correlated with hybridisation and polyploidisation. Change of mating systems is one of the major driving forces for speciation, initiating reproductive isolation of populations [97,98]. However, it is still unclear if there is a correlation between mating systems and hybridisation and/or polyploidisation. The breakdown of self-incompatibility in an artificial cross between *A. thaliana* and *A. lyrata* [99] could indicate a correlation between hybridisation and a switch in mating system, and it could further indicate possible dominance of the selfer *A. thaliana* over *A. lyrata*.

Otherwise, breakdown of the sporophytic self-incompatibility system has been reported mainly from diploid individuals [16,100].

In contrast to *A. suecica*, allopolyploid *A. kamchatica* has a vast distribution range spanning the whole amphi-Beringian region. Our data suggest that this large distribution is partly due to postglacial colonisation of formerly glaciated areas in eastern Beringia and western Canada. The extremely reduced genetic diversity, particularly of the group with suprahaplotype B, suggests that postglacial immigration may have been rapid, possibly facilitated by de novo adaptations as a result of hybridisation and polyploidisation. Moreover, the success of *A. kamchatica* as a rapid coloniser may have been enhanced by the availability of large, open landscapes, where habitats were frequently disturbed by glacial and/or permafrost activity. However, the change in mating systems may have had a strong impact on the success of *A. kamchatica* as a postglacial coloniser, and also on the establishment of the genetic barrier between *A. kamchatica* and the Eurasian and North American lineages.

Beringia as contact zone between 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* Vahl. [32], *Saxifraga hirculus* L. [36], *Saxifraga oppositifolia* [33,31], and *Vaccinium uliginosum* L. [35,101]. Periglacial survival in Beringia can also be assumed for the *A. lyrata* complex, in particular, for the Eurasian lineage in arctic western and eastern Beringia north of Brooks Range. The nuclear sequence data indicate that gene exchange with populations of the North American lineage (*A. lyrata* ssp. *lyrata*, *A. arenicola*) occurred inter- and postglacially. However, the data do not support glacial survival of the North American lineage in Beringia, as no plastidic sequence types of this lineage were found in Beringia.

Conclusions

By presenting a worldwide evolutionary history of the *Arabidopsis lyrata* species complex, we provide solid knowledge about centres of genetic diversity, different genetic lineages, their contact zones, and hybrid speciation. We could clearly differentiate three genetic lineages, a Eurasian, a North American, and an amphi-Pacific one. The latter is constituted of the allopolyploid A. kamchatica, a hybrid between *A. lyrata and A. halleri*. Further investigations of the population dynamics and the role of selfing within this hybrid species should be conducted to gain a deeper understanding of hybrid establishment in the wild.

Additional file 1: Table S1. The two main taxonomic concepts of the *Arabidopsis lyrata* complex. Summarised from Al-Shehbaz and O'Kane [37], including revision from the Flora of North America [Al-Shehbaz, personal communication], and Elven [38]. *Arabidopsis arenicola* was integrated from Warwick et al. [42]. Elven [38] excluded *Arabidopsis lyrata* from taxonomic treatment, as they assumed it to be a non-arctic, boreal taxon.

Additional file 2: Table S2. Within this list all information about taxonomic unit, name on herbarium sheet, herbarium, herbarium number, locality, latitude/longitude, collector, collection date, accession number, ITS type, ITS GenBank number, ITS supratype, *trn*L intron type, *trn*L GenBank number, *trn*L/F-IGS type, *trn*L/F-IGS GenBank number, *trn*L intron + *trn*L/F-IGS type, *trn*L/F suprahaplotype, and *Pgi*C1 amplification is provided.

Additional file 3: Figure S1. Single most parsimonious tree (length = 553) with bootstrap/jackknife values above 95, based on 37 *Arabidopsis* nuclear DNA *PgiC* sequences. Heuristic searches were performed with 100 random addition sequences and TBR branch swapping, saving three trees per replicate, in TNT [102]. Gaps were treated as fifth state. Consistency index (CI) = 0.69, retention index (RI) = 0.95. Investigated accessions were from the *A. lyrata* complex (ssp. *lyrata*, ssp. *petraea*, *A. septentrionalis*, and *A. umbrosa*) and the *A. halleri* complex (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.

Additional file 4: Figure S2. Alignment of the duplicated phosphoglucoisomerase loci *Pgi*C1 and *Pgi*C2 with the primer binding sites indicated. The forward primer is partly located within the 7 bp deletion between positions 1389 and 1396.

Additional file 5: Figure S3. Selected PCR reactions from the *Pgi*C screening: No *Pgi*C1 amplification in members of the *Arabidopsis lyrata* complex (*A. lyrata* ssp. *lyrata*, *A. arenicola*, *Arabidopsis umbrosa*, and *A. septentrionalis*). Successful *Pgi*C1 amplification in members of *A. halleri* (ssp. *halleri*, ssp. *dacica*, ssp. *tatrica*, and ssp. *gemmifera*), and *A. kamchatica*.

Additional file 6: Table S3. Taxonomic and regional genetic differentiation based on cpDNA suprahaplotypes. Numbers of cpDNA suprahaplotypes occurring in each region are provided.

Additional file 7: Table S4. List of ITS supratypes, ITS types, *trn*L/F suprahaplotypes, and *trn*L/F haplotypes in the *Arabidopsis lyrata* complex with their corresponding frequencies of occurrence throughout the dataset *(italic)*.

Acknowledgements

We thank Ishan Al-Shehbaz (Missouri, USA), Galina Gusarova (Oslo, Norway), Gu Hongya (Beijing, P. R. China), Barbara Mable (Glasgow, Scotland), David L. Remington (Greensboro, USA), Outi Savolainen (Oulu, Finland), and the curators of the Herbariums of the Natural History Museums London and Vienna for providing plant material, Susanne Ball and Stefan Wötzel for laboratory assistance, and Jürgen Ehlers and Phil Gibbard for shapefiles of the LGM. This research was supported by DFG grant KO 2302/5-2 to Marcus A. Koch.

Author details

[1]Heidelberg University, Heidelberg Institute of Plant Sciences, Biodiversity and Plant Systematics, Im Neuenheimer Feld 345, D-69120 Heidelberg, Germany. [2]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066 Blindern, NO-0316 Oslo, Norway.

Authors' contributions

RS carried out the molecular marker studies and statistical analyses, constructed the maps and drafted the manuscript. MHJ constructed the *PgiC* sequence alignment and helped to draft the manuscript. AKB contributed to draft the manuscript. MAK designed and coordinated the

project and drafted the manuscript. All authors read and approved the final manuscript.

Received: 27 November 2009 Accepted: 8 April 2010
Published: 8 April 2010

References

1. Mitchell-Olds T: ***Arabidopsis thaliana*** **and its wild relatives: a model system for ecology and evolution.** *Trends Ecol Evol* 2001, **16**:693-700.
2. Benfey PN, Mitchell-Olds T: **From genotype to phenotype: systems biology meets natural variation.** *Science* 2008, **320**:495-497.
3. Koch MA, Haubold B, Mitchell-Olds T: **Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae).** *Mol Biol Evol* 2000, **17**:1483-1498.
4. Koch MA, Haubold B, Mitchell-Olds T: **Molecular systematics of the Brassicaceae: evidence from coding plastidic *MATK* and nuclear *CHS* sequences.** *Am J Bot* 2001, **88**:534-544.
5. Turner TL, Von Wettberg EJ, Nuzhdin SV: **Genomic analysis of differentiation between soil types reveals candidate genes for local adaptation in *Arabidopsis lyrata*.** *PLoS One* 2008, **3**:e3183.
6. Clauss MJ, Koch MA: **Poorly known relatives of *Arabidopsis thaliana*.** *Trends Plant Sci* 2006, **11**:449-459.
7. Zhukova PG, Petrovsky VV, Plieva TV: **The chromosome numbers and taxonomy of some plant species from Siberia and Far East.** *Botanicheskiy Zhurnal* 1973, **58**:1331-1342.
8. Shimizu-Inatsugi R, Lihová J, Iwanaga H, Kudoh H, Marhold K, Savolainen O, Watanabe K, Yakubov VV, Shimizu KK: **The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*.** *Mol Ecol* 2009, **18**:4024-4048.
9. Riihimäki M, Savolainen O: **Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae).** *Am J Bot* 2004, **91**:1036-1045.
10. Riihimäki M, Podolsky R, Kuittinen H, Koelewijn H, Savolainen O: **Studying genetics of adaptive variation in model organisms: flowering time variation in *Arabidopsis lyrata*.** *Genetica* 2005, **123**:63-74.
11. Clauss MJ, Dietel S, Schubert G, Mitchell-Olds T: **Glucosinolate and trichome defenses in a natural *Arabidopsis lyrata* population.** *J Chem Ecol* 2006, **32**:2351-2373.
12. Kusaba M, Dwyer K, Hendershot J, Vrebalov J, Nasrallah JB, Nasrallah ME: **Self-incompatibility in the genus *Arabidopsis*: characterization of the S locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*.** *Plant Cell* 2001, **13**:627-643.
13. Schierup MH, Mable BK, Awadalla P, Charlesworth D: **Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*.** *Genetics* 2001, **158**:387-399.
14. Charlesworth D, Mable BK, Schierup MH, Bartolomé C, Awadalla P: **Diversity and linkage of genes in the self-incompatibility gene family in *Arabidopsis lyrata*.** *Genetics* 2003, **164**:1519-1535.
15. Mable BK, Beland J, di Berardo C: **Inheritance and dominance of self-incompatibility alleles in polyploid *Arabidopsis lyrata*.** *Heredity* 2004, **93**:476-486.
16. Mable BK, Robertson AV, Dart S, diBerardo C, Witham L: **Breakdown of self-incompatibility in the perennial *Arabidopsis lyrata* (Brassicaceae) and its genetic consequences.** *Evolution* 2005, **59**:1437-1448.
17. Charlesworth D, Kamau E, Hagenblad J, Tang C: **Trans-specificity at loci near the self-incompatibility loci in *Arabidopsis*.** *Genetics* 2006, **172**:2699-2704.
18. Hagenblad J, Bechsgaard J, Charlesworth D: **Linkage disequilibrium between incompatibility locus region genes in the plant *Arabidopsis lyrata*.** *Genetics* 2006, **173**:1057-1073.
19. Schierup MH, Bechsgaard JS, Nielsen LH, Christiansen FB: **Selection at work in self-incompatible *Arabidopsis lyrata*: mating patterns in a natural population.** *Genetics* 2006, **172**:477-484.
20. Schierup MH, Bechsgaard JS, Christiansen FB: **Selection at work in self-incompatible *Arabidopsis lyrata*. II. Spatial distribution of S haplotypes in Iceland.** *Genetics* 2008, **180**:1051-1059.
21. Sharbel TF, Haubold B, Mitchell-Olds T: **Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe.** *Mol Ecol* 2000, **9**:2109-2118.
22. Beck JB, Schmuths H, Schaal BA: **Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics.** *Mol Ecol* 2008, **17**:902-915.
23. Wright SI, Lauga B, Charlesworth D: **Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*.** *Mol Ecol* 2003, **12**:1247-1263.
24. Balañá-Alcaide D, Ramos-Onsins SE, Boone Q, Aguadé M: **Highly structured nucleotide variation within and among *Arabidopsis lyrata* populations at the FAH1 and DFR gene regions.** *Mol Ecol* 2006, **15**:2059-2068.
25. Clauss MJ, Mitchell-Olds T: **Population genetic structure of *Arabidopsis lyrata* in Europe.** *Mol Ecol* 2006, **15**:2753-2766.
26. Muller MH, Leppälä J, Savolainen O: **Genome-wide effects of postglacial colonization in *Arabidopsis lyrata*.** *Heredity* 2008, **100**:47-58.
27. Koch MA, Matschinger M: **Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*.** *Proc Natl Acad Sci USA* 2007, **104**:6272-6277.
28. Schmickl R, Jørgensen MH, Brysting AK, Koch MA: **Phylogeographic implications for the North American boreal-arctic *Arabidopsis lyrata* complex.** *Plant Ecol Div* 2008, **1**:245-254.
29. Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC: **Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics.** *Annu Rev Ecol Syst* 1987, **18**:489-522.
30. Eidesen PB, Carlsen T, Molau U, Brochmann C: **Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic.** *J Biogeogr* 2007, **34**:1559-1574.
31. Abott RJ, Comes HP: **Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (Purple saxifrage).** *New Phytol* 2004, **161**:211-224.
32. Tremblay NO, Schoen DJ: **Molecular phylogeography of *Dryas integrifolia*: glacial refugia and postglacial recolonization.** *Mol Ecol* 1999, **8**:1187-1198.
33. Abbott RJ, Smith LC, Milne RI, Crawford RM, Wolff K, Balfour J: **Molecular analysis of plant migration and refugia in the Arctic.** *Science* 2000, **289**:1343-1346.
34. Brochmann C, Gabrielsen TM, Nordal I, Landvik JY, Elven R: **Glacial survival or tabula rasa? The history of North Atlantic biota revisited.** *Taxon* 2003, **52**:417-450.
35. Alsos IG, Engelskjøn T, Gielly L, Taberlet P, Brochmann C: **Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species.** *Mol Ecol* 2005, **14**:2739-2753.
36. Oliver C, Hollingsworth PM, Gornall RJ: **Chloroplast DNA phylogeography of the arctic-montane species *Saxifraga hirculus* (Saxifragaceae).** *Heredity* 2006, **96**:222-231.
37. Al-Shehbaz IA, O'Kane SL Jr: **Taxonomy and phylogeny of *Arabidopsis* (Brassicaceae).** *The Arabidopsis Book* Rockville, MD: American Society of Plant Biologists 2002.
38. Elven R, (Ed): **Checklist of the Panarctic Flora (PAF) Vascular Plants.** *Version* 2007 [http://www.binran.ru/infsys/paflist/index.htm].
39. O'Kane SL Jr, Al-Shehbaz IA: **A synopsis of *Arabidopsis* (Brassicaceae).** *Novon* 1997, **7**:323-327.
40. Al-Shehbaz IA, O'Kane SL Jr, Price RA: **Generic placement of species excluded from *Arabidopsis* (Brassicaceae).** *Novon* 1999, **9**:296-307.
41. Koch MA, Wernisch M, Schmickl R: ***Arabidopsis thaliana*** **'s wild relatives: an updated overview on systematics, taxonomy and evolution.** *Taxon* 2008, **57**:933-943.
42. Warwick SI, Al-Shehbaz IA, Sauder CA: **Phylogenetic position of *Arabis arenicola* and generic limits of *Aphragmus* and *Eutrema* (Brassicaceae) based on sequences of nuclear ribosomal DNA.** *Can J Bot* 2006, **84**:269-281.
43. Shimizu KK, Fujii S, Marhold K, Watanabe K, Kudoh H: ***Arabidopsis kamchatica*** **(Fisch. ex DC.) K. Shimizu & Kudoh and *A. kamchatica* subsp. *kawasakiana* (Makino) K. Shimizu & Kudoh, new combinations.** *Acta Phytotax Geobot* 2005, **56**:163-172.
44. Doyle JJ, Doyle JL: **A rapid DNA isolation procedure for small quantities of fresh leaf tissue.** *Phytochem Bull* 1987, **19**:11-15.
45. Dobeš CH, Mitchell-Olds T, Koch MA: **Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A.* × *divaricarpa*, and *A. holboellii* (Brassicaceae).** *Mol Ecol* 2004, **13**:349-370.
46. Dobeš C, Mitchell-Olds T, Koch MA: **Intraspecific diversification in North American *Boechera stricta* (= *Arabis drummondii*), *Boechera* × *divaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and**

chloroplast molecular markers - an integrative approach. *Am J Bot* 2004, **91**:2087-2101.
47. Koch MA, Dobeš C, Matschinger M, Bleeker W, Vogel J, Kiefer M, Mitchell-Olds T: **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* 2005, **22**:1032-1043.
48. Dobeš C, Kiefer C, Kiefer M, Koch MA: **Plastidic *trnF* UUC pseudogenes in North American genus *Boechera* (Brassicaceae): mechanistic aspects of evolution.** *Plant Biol* 2007, **9**:502-515.
49. Koch MA, Dobeš C, Kiefer C, Schmickl R, Klimeš L, Lysak MA: **Supernetwork identifies multiple events of plastid *trn* F(GAA) pseudogene evolution in the Brassicaceae.** *Mol Biol Evol* 2007, **24**:63-73.
50. Schmickl R, Kiefer C, Dobeš C, Koch MA: **Evolution of *trn* F(GAA) pseudogenes in cruciferous plants.** *Plant Syst Evol* 2009, **282**:229-240.
51. Ansell SW, Schneider H, Pedersen N, Grundmann M, Russell SJ, Vogel JC: **Recombination diversifies chloroplast *trn*F pseudogenes in *Arabidopsis lyrata*.** *J Evol Biol* 2007, **20**:2400-2411.
52. Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E: **Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes.** *Proc Natl Acad Sci USA* 1980, **77**:7323-7327.
53. Liao D: **Concerted evolution: molecular mechanism and biological implications.** *Am J Hum Genet* 1999, **64**:24-30.
54. Koch MA, Dobeš C, Mitchell-Olds T: **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* 2003, **20**:338-350.
55. Carson AR, Scherer SW: **Identifying concerted evolution and gene conversion in mammalian gene pairs lasting over 100 million years.** *BMC Evol Biol* 2009, **9**:156.
56. Ehlers J, Gibbard PL, (Eds): *Quaternary glaciations - extent and chronology* Amsterdam: Elsevier 2004.
57. Clement M, Posada D, Crandall KA: **TCS: a computer program to estimate gene genealogies.** *Mol Ecol* 2000, **9**:1657-1659.
58. Templeton AR, Crandall KA, Sing CF: **A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation.** *Genetics* 1992, **132**:619-633.
59. Excoffier L, Laval G, Schneider S: **Arlequin (version 3.0): An integrated software package for population genetics data analysis.** *Evol Bioinform Online* 2005, **1**:47-50.
60. Kumar S, Nei M, Dudley J, Tamura K: **MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences.** *Brief Bioinform* 2008, **9**:299-306.
61. Gottlieb LD: **Evidence for duplication and divergence of the structural gene for phosphoglucoisomerase in diploid species of *Clarkia*.** *Genetics* 1977, **86**:289-308.
62. Ghatnekar L: **A polymorphic duplicated locus for cytosolic PGI segregating in sheep's fescue (*Festuca ovina* L.).** *Heredity* 1999, **83**:451-459.
63. Ford VS, Gottlieb LD: **Single mutations silence PgiC2 genes in two very recent allotetraploid species of *Clarkia*.** *Evolution* 2002, **56**:699-707.
64. Abbott RT, Brochmann C: **History and evolution of the arctic flora: in the footsteps of Eric Hulten.** *Mol Ecol* 2003, **12**:299-313.
65. DeChain EG: **A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants.** *Plant Ecol Div* 2008, **1**:197-207.
66. Tribsch A, Schönswetter P: **Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps.** *Taxon* 2003, **52**:477-497.
67. Hewitt GM: **The structure of biodiversity - insights from molecular phylogeography.** *Front Zool* 2004, **1**:4.
68. Koch MA, Kiefer C, Ehrich D, Vogel J, Brochmann C, Mummenhoff K: **Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae).** *Mol Ecol* 2006, **15**:825-839.
69. Ehrich D, Gaudeul M, Assefa A, Koch MA, Mummenhoff K, Nemomissa S, Brochmann C: **Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains.** *Mol Ecol* 2007, **16**:2542-2559.
70. Schönswetter P, Paun O, Tribsch A, Niklfeld H: **Out of the Alps: colonization of northern Europe by east alpine populations of the glacier buttercup *Ranunculus glacialis* L. (Ranunculaceae).** *Mol Ecol* 2003, **12**:3373-3381.
71. Ehrich D, Alsos IG, Brochmann C: **Where did the northern peatland species survive the dry glacials: cloudberry (*Rubus chamaemorus*) as an example.** *J Biogeogr* 2008, **35**:801-814.
72. Savile DBO: **Arctic adaptations in plants.** *Canada Department of Agriculture Monograph* 1972, **6**:1-81.
73. Alsos IG, Eidesen PB, Ehrich D, Skrede I, Westergaard K, Jacobsen GH, Landvik JY, Taberlet P, Brochmann C: **Frequent long-distance plant colonization in the changing Arctic.** *Science* 2007, **316**:1606-1609.
74. Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen AC, Elven R: **Polyploidy in arctic plants.** *Biol J Linn Soc* 2004, **82**:521-536.
75. Manton I: **The problem of *Biscutella laevigata* L. - II. The evidence from meiosis.** *Ann Bot* 1937, **1**:439-462.
76. Stebbins GL: **Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach.** *Bot Helv* 1984, **94**:1-13.
77. Soltis PS, Soltis DE: **The role of genetic and genomic attributes in the success of polyploids.** *Proc Natl Acad Sci USA* 2000, **97**:7051-7057.
78. Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B: **Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids.** *Plant Cell* 2000, **12**:1551-1568.
79. Wendel JF: **Genome evolution in polyploids.** *Plant Mol Biol* 2000, **42**:225-249.
80. Liu B, Wendel JF: **Epigenetic phenomena and the evolution of plant allopolyploids.** *Mol Phylogenet Evol* 2003, **29**:365-379.
81. Brysting AK, Oxelman B, Huber KT, Moulton V, Brochmann C: **Untangling complex histories of genome mergings in high polyploids.** *Syst Biol* 2007, **56**:467-476.
82. Guggisberg A, Mansion G, Kelso S, Conti E: **Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*.** *New Phytol* 2006, **171**:617-632.
83. Guggisberg A, Baroux C, Grossniklaus U, Conti E: **Genomic origin and organization of the allopolyploid *Primula egaliksensis* investigated by in situ hybridization.** *Ann Bot* 2008, **101**:919-927.
84. Guggisberg A, Mansion G, Conti E: **Disentangling reticulate evolution in an arctic-alpine polyploid complex.** *Syst Biol* 2009, **58**:55-73.
85. Popp M, Erixon P, Eggens F, Oxelman B: **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* 2005, **30**:302-313.
86. Popp M, Oxelman B: **Origin and evolution of North American polyploid *Silene* (Caryophyllaceae).** *Am J Bot* 2007, **94**:330-349.
87. Brochmann C, Xiang QY, Brunsfeld SJ, Soltis DE, Soltis PS: **Molecular evidence for polyploid origins in *Saxifraga* (Saxifragaceae): the narrow arctic endemic *S. svalbardensis* and its widespread allies.** *Am J Bot* 1998, **85**:135-143.
88. Jørgensen MH, Elven R, Tribsch A, Gabrielsen TM, Stedje B, Brochmann C: **Taxonomy and evolutionary relationships in the *Saxifraga rivularis* complex.** *Syst Bot* 2006, **31**:702-729.
89. Jørgensen MH, Carlsen T, Skrede I, Elven R: **Microsatellites resolve the taxonomy of the polyploid *Cardamine digitata* aggregate (Brassicaceae).** *Taxon* 2008, **57**:882-892.
90. Kapralov MV, Gabrielsen TM, Sarapultsev IE, Brochmann C: **Genetic enrichment of the arctic clonal plant *Saxifraga cernua* at its southern periphery via the alpine sexual *Saxifraga sibirica*.** *Mol Ecol* 2006, **15**:3401-3411.
91. Säll T, Jakobsson M, Lind-Halldén C, Halldén C: **Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*.** *J Evol Biol* 2003, **16**:1019-1029.
92. Jakobsson M, Hagenblad J, Tavaré S, Säll T, Halldén C, Lind-Halldén C, Nordborg M: **A unique recent origin of the allotetraploid species *Arabidopsis suecica*: Evidence from nuclear DNA markers.** *Mol Biol Evol* 2006, **23**:1217-1231.
93. Säll T, Lind-Halldén C, Jakobsson M, Halldén C: **Mode of reproduction in *Arabidopsis suecica*.** *Hereditas* 2004, **141**:313-317.
94. Sugisaka J, Kudoh H: **Breeding system of the annual Cruciferae, *Arabidopsis kamchatica* subsp. *kawasakiana*.** *J Plant Res* 2008, **121**:65-68.
95. Bechsgaard JS, Castric V, Charlesworth D, Vekemans X, Schierup MH: **The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr.** *Mol Biol Evol* 2006, **23**:1741-1750.

96. Tang C, Toomajian C, Sherman-Broyles S, Plagnol V, Guo YL, Hu TT, Clark RM, Nasrallah JB, Weigel D, Nordborg M: **The evolution of selfing in *Arabidopsis thaliana*.** *Science* 2007, **317**:1070-1072.
97. Schoen DJ, Johnston MO, L'Heureux AM, Marsolais JV: **Evolutionary history of the mating system in *Amsinckia* (Boraginaceae).** *Evolution* 1997, **51**:1090-1099.
98. Rieseberg LH, Willis JH: **Plant speciation.** *Science* 2007, **317**:910-914.
99. Nasrallah JB, Liu P, Sherman-Broyles S, Schmidt R, Nasrallah ME: **Epigenetic mechanisms for breakdown of self-incompatibility in interspecific hybrids.** *Genetics* 2007, **175**:1965-1973.
100. Mable BK: **Genetic causes and consequences of the breakdown of self-incompatibility: case studies in the Brassicaceae.** *Genet Res* 2008, **90**:47-60.
101. Eidesen PB, Alsos IG, Popp M, Stensrud Ø, Suda J, Brochmann C: **Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species.** *Mol Ecol* 2007, **16**:3902-3925.
102. Goloboff PA, Farris JS, Nixon KC: **TNT (Tree analysis using New Technology).** Published by the authors, Tucumán 2000.

doi:10.1186/1471-2148-10-98
Cite this article as: Schmickl *et al.*: The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. *BMC Evolutionary Biology* 2010 **10**:98.

V

VI

Interspecific and interploidal gene flow in Central European *Arabidopsis* (Brassicaceae)

Marte H. Jørgensen[1, 4], Dorothee Ehrich[2], Roswitha Schmickl[3], Marcus A. Koch[3], and Anne K. Brysting[1]

[1]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, P.O. Box 1066 Blindern, NO-0316 Oslo, Norway

[2]Institute for Arctic and Marine Biology, University of Tromsø, NO-9037 Tromsø, Norway

[3]Centre for Organismal Studies (COS) Heidelberg, Department of Biodiversity and Plant Systematics, University of Heidelberg, Im Neuenheimer Feld 345, D-69120 Heidelberg, Germany

[4]Author for correspondence: martej@ulrik.uio.no, +47 92046378

Abstract

Background: Effects of polyploidisation on gene flow between natural populations are little known. Central European diploid and tetraploid populations of *Arabidopsis arenosa* and *A. lyrata* are here used to study interspecific and interploidal gene flow, and we attempt to distinguish between ancient and recent events.

Results: Ploidy levels were confirmed by flow cytometry. Network analyses clearly separated diploids according to species. Tetraploids and diploids were highly intermingled within species, and some tetraploids intermingled with the other species, as well. Isolation with migration analyses suggested interspecific introgression from tetraploid *A. arenosa* to tetraploid *A. lyrata* and vice versa, and some interploidal gene flow within taxa, more prevailing in *A. lyrata* than *A. arenosa*.

Conclusions: Interspecific genetic isolation at diploid level combined with introgression at tetraploid level indicates that polyploidy may buffer against negative consequences of interspecific hybridisation. The fact that we found species-specific differences in interploidal gene flow between sympatric *A. lyrata* and allopatric *A. arenosa* cytotypes suggests that further use of IM analysis might be one way to go to establish criteria for distinguishing between ancient and recent gene flow.

Background

Polyploidy, i.e. whole genome duplication, has long been considered a major evolutionary force in the Plant Kingdom [see e.g., 1, 2-5], and even though large advances in our understanding of polyploidy have been made during the last couple of decades, there are still many questions unanswered [reviewed by 6]. We still don't have a general agreement on classification of polyploids, for instance. Some authors work with strict taxonomic definitions; autopolyploids are the result of polyploidisation events involving only a single species, and allopolyploids are the result of interspecific hybridisation [e.g., 3]. Others base their definitions on inheritance patterns and the presence or absence of multivalents [e.g., 7]. However, most would agree that auto- and allopolyploids are the extremes of a continuous range. There are also still controversies about how polyploids should be treated taxonomically. Soltis et al. [8] suggest that autopolyploids deserve species rank taxonomically, with ploidy level as part of the name. Others do not even give allopolyploids species status due to lack of morphological distinctness (e.g., lack of diagnostic qualitative and descrete characters), and include them as subspecies of one of the parents [e.g., 9]. Yet others separate morphological and biological species where the first may contain several of the latter [e.g., 10].

Traditionally, polyploidisation events have been considered to result in total reproductive isolation of the new polyploid from the parent (s), and thus regarded as instant speciation [e.g., 11]. More recent research has shown that recurrent formation of polyploids and triploid bridges contribute to interploidal gene flow [3, 12, 13]. To what extent, however, is still not known [6]. Multiple independent polyploidisation events have been shown to be common for both allopolyploids [e.g., 14, 15] and autopolyploids [e.g., 16, 17, 18]. Population studies and modelling of sympatric *Chamerion angustifolium* (L.) Holub revealed that autotetraploids are not necessarily instantly isolated from their diploid progenitors, but

that the isolation can become more prevalent through time [19, 20]. Slotte et al. [21] showed that there is unidirectional gene flow from diploid *Capsella rubella* Reuter to its allotetraploid descendant *C. bursa-pastoris* (L.) Medicus. Furthermore, if polyploidisation events result in immediate isolation from the progenitors, the result should be a major bottleneck. However, several studies have shown higher genetic diversity in polyploids compared to their progenitors [22-26], although this is not always the case [14, 27-30]. The increased diversity may be the result of either recurrent formation of the polyploids [e.g., 1], or past or ongoing interploidal gene flow through, for instance, triploid bridges [e.g., 19]. These different models of polyploidisation can be seen as a gradient. Single event polyploidisation with subsequent reproductive isolation represents one end of this gradient whereas polyploidisation with ongoing gene flow or recurrent polyploidisation represent the other. Instances where polyploidisation is followed by historical gene flow which later stopped, or where polyploidisation is followed by reproductive isolation and subsequent gene flow in the form of secondary contact [31], could be considered as intermediate forms. Criteria and methods to distinguish between these different categories, though, have not yet been proposed [6].

Arabidopsis (DC.) Heynh. is a small genus consisting mostly of diploids, but includes both allopolyploids [14, 32-35] and taxonomic autopolyploids [36-41]. As the genus includes the geneticists' pet plant *A. thaliana* L., plenty of molecular tools are available also for its relatives [e.g., the recent release of the A. lyrata genome, 42], making the genus ideal for studying polyploid evolution. In Central Europe two species have been recorded with two ploidy levels each: *A. arenosa* (L.) Lawalrée (hereafter *arenosa*) represents a complex species aggregate [38] and diploid populations occur mainly in the Carpathians and possibly in a few regions futher south in Hungary and Croatia, whereas tetraploid *arenosa* is found in most of Central Europe. In contrast *A. lyrata* (L.) O'Kane & Al-Shehbaz (hereafter *lyrata*) is mostly

diploid throughout its European distribution range, but several tetraploid populations are found in the Eastern Austrian Forealps and neighboring regions [39, 43, 44]. Schmickl and Koch (submitted) detected significant levels of introgression from *arenosa* into *lyrata* in the Eastern Austrian Forealps using microsatellite markers and morphometrics. These results, mainly based on genetic admixture and differentiation, suggest past and ongoing gene flow between the two taxa. To distinguish between ancient and recent gene flow, especially in polyploid complexes, has long been controversial, and no clear criteria have yet been commonly acknowledged [6]. Here we use low-copy nuclear and plastid DNA sequences from the *Arabidopsis* model system and different analytical methods to study interspecific and interploidal gene flow in *lyrata* and *arenosa*, specifically asking, 1) whether interploidal gene flow takes place in one or both directions, 2) how polyploidization affects interspecific introgression and 3) whether it is possible to distinguish between ongoing gene flow and retention of ancestral polymorphism.

Methods

Material

Three to five Central European populations of each ploidy level (2*x* and 4*x*) from both *arenosa* and *lyrata* ssp. *petraea* (L.) O'Kane & Al-Shehbaz were included in this study (Table 1; Fig. 1) with a total of 16 populations. *Lyrata* 2*x*, *lyrata* 4*x* and *arenosa* 4*x* were mostly sampled in close proximity in the Eastern Austrian Forealps. *Arenosa* 2*x* was sampled in Slovakian Carpathians. The material was collected as living plants or seeds from defined single mother plants and grown in the Botanical Garden, University of Heidelberg and the Phytotron, University of Oslo. Fresh leaves of 1-43 plants from most populations (274

individuals in total; Table 1) were later collected for flow cytometry analyses. Leaves from two specimens per population were dried using silica gel to preserve DNA before extraction.

Flow cytometry

Relative nuclear DNA content of 274 specimens from 14 populations (Table 1) were estimated by flow cytometry analyses performed by G. Geenen (Plant Cytometry Services, Schijndel), using DAPI staining, the *Arabidopsis* buffer described in Doležel and Suda [45] and *Ilex crenata* Thunb. 'Fastigiata' as internal standard, otherwise following the protocol described in Jørgensen *et al*. [30]. The populations a4_GER and a4_AUT3 were not included in the analyses as we did not have living material from these at the time of the analysis. However, microsatellite data for these populations indicate that they are tetraploid (Schmickl and Koch submitted; Schmickl and Koch unpublished). T-tests were done in SPSS 16.0 (SPSS Inc., Chicago) to test for differences in means of nuclear DNA content between the taxa.

DNA extraction, cloning, and sequencing

Whole genomic DNA was extracted from leaf tissue using the DNeasy Plant Mini protocol (Qiagen, Hilden). Polymerase chain reaction (PCR) of the low-copy nuclear regions chalcone synthase (CHS) and short chain alcohol dehydrogenase (*sc*ADH) was carried out in 25µl volumes with 1X DyNAzyme EXT buffer (Finnzymes, Espoo), 0.2 mM of each dNTP, 0.6 µM of each primer (Additional files Table S1), 0.2 U DyNAzyme EXT DNA polymerase (Finnzymes), and 2 µl 10 times diluted DNA template. Thermocycling conditions consisted of 3 min at 94˚C, and 35 cycles of 30 s at 93˚C, 30 s at 55˚C, 2.5-3 min at 70˚C, and a final extension for 5 min at 70˚C. PCR products were cloned using the TOPO-TA kit for sequencing with the pCR4-TOPO vector (Invitrogen, Carlsbad). Colonies were checked for

inserts by running a PCR with M13 or T7 primers. At least six insert-containing clones from each PCR reaction were sequenced in both directions. The plastid region *trn*L-F was amplified using PuReTaq Ready-To-Go PCR beads (GE Healthcare, Waukesha) with 0.6 μM of each primer (Additional files Table S1) and 2 μL 10 times diluted DNA template. For each DNA region, both strands were sequenced using BigDye v3.1 7 cycle sequencing kit (Applied Biosystems, Foster City) and M13F/M13R or T7 primers. Products of the cycle-sequencing reactions were separated on an ABI 3700 Genetic Analyzer (Applied Biosystems). The resulting sequences were assembled and edited using Vector NTI advance 10 (Invitrogen), and consensus sequences representing each allele and alignments were made using BioEdit version 7.0.5 [46]. For the plastid *trn*L-F region, only the first 700 base pairs were included in the analyses, to avoid unambigous alignment due to the presense of pseudogenes [47, 48].

PCR-mediated recombinants (chimeras) constitute a well-known problem in PCR-based cloning protocols [e.g. 49, 50-54], and to distinguish between PCR-mediated and real recombinants is not possible vía PCR-based methods. However, the risk of obtaining recombinants as PCR artefacts increases with the concentration of template [50], and the expected frequency of these should be lower than for real recombinants [49, 51]. In this study we omitted clones that were recombinants of other cloned sequences from the same individual, and that were present at low frequencies, as PCR artefacts. Discrepant bases supported by only a single clone were assumed to be due to polymerase reading error and were corrected based on consensus sequences of other clones from the same individual.

Data analyses

Intragenic recombination events may be relatively common [55], and should be taken into consideration when chosing methods for phylogenetic analysis [56]. Minimum number of

recombination events [RM; 57] per region was calculated using DnaSP [58]. As we found substantial recombination for most regions (Additional files Table S1), phylogenetic relationships were analysed for each region using neighbour networks [59], with Jukes-Cantor distances in the program SplitsTree4 [60]. Gaps were included following the simple coding strategy introduced by Simmons and Ochoterena [61] as implemented in the software SeqState [62]. For all marker systems the datasets were analysed by: 1) splitting the individuals into subsets according to taxa, 2) including all individuals, and 3) including only diploids.

When two or three different alleles are found in a tetraploid individual, it is not possible to determine the true genotype although dosage may give an indication. Computational methods based among others on the EM algorithm have been developed to infer genotypes assuming random mating and populations at equilibrium [e.g. 63]. As we have sampled two individuals from several distinct populations, we cannot assume equilibrium, and chose not to use any statistical method. To roughly assess the impact of assuming different numbers of allele copies, we reconstituted genotypes at random using the following approach: For each tetraploid individual with two or three alleles, a random number between 1 and 3 was generated. For individuals with two alleles, 1 corresponded to three copies of the first allele and one copy of the second (the order was arbitrary), 2 corresponded to two copies of each allele and 3 corresponded to one copy of the first allele and three copies of the second. For individuals with three alleles, 1 corresponded to duplicating the first allele, 2 to duplicating the second allele and 3 to duplicating the third allele. Three different datasets (D1, D2, D3) were generated using this approach. Assuming that the three allele proportions 1:3, 2:2, 3:1 are equally probable for tetraploids with two distinct alleles at a locus leads to having a deficit of 2:2 individuals compared to equilibrium expectations. Therefore we created a fourth dataset (D22) where all individuals with two

184 distinct alleles were considered to have a 2:2 genotype. The four datasets were used both in
185 diversity calculations and isolation with migration (IM) analyses.
186 Diversity indices were calculated for each molecular region and each species/ploidy
187 level, separately. For the tetraploids we calculated the indices for the duplicated datasets D1,
188 D2, D3, and D22, and averaged the estimates. DnaSP [58] was used to calculate gene
189 diversity (Hd), nucleotide diversity averaged over loci (π), and average number of nucleotide
190 differences (k).
191 The parameters of the isolation with migration (IM) model were estimated as
192 implemented in the program IMa2 [64, 65] to assess the importance and direction of gene
193 flow between the ploidy levels and species. Samples of *arenosa* and *lyrata* were first
194 analysed separately, to determine whether there is gene flow between ploidy levels and in
195 which direction it occurs. Second, in order to assess gene flow between species, both taxa
196 with both ploidy levels were analysed together in an analysis with four populations. We
197 assumed one ancestral population for each species, and one ancestral population for the
198 whole complex. In such an analysis with four populations a large number of parameters have
199 to be estimated, requiring a large amount of data to obtain reliable estimates. As the three loci
200 available here were somewhat limited in that respect, we also analysed only the two diploid
201 taxa to assess evidence for interspecific gene flow. All analyses involving tetraploids were
202 carried out for the four different datasets of tetraploid genotypes. One of the assumptions of
203 the IM model is that there is no important gene exchange with populations not included in the
204 analysis. As there was evidence for significant gene flow between ploidy levels in both
205 species, we did not analyse the two tetraploids together. Another assumption of IMa2 is that
206 there is no recombination. Tests for recombination [57, 66] showed, however, that there was
207 considerable recombination in the nuclear sequences used here. The program IMgc [67] was
208 used to find the largest subsets of the data matrix without signs of recombination

209 (nonrecombining blocks) by removing either sequences or variable sites. The program can
210 prioritise the number of sequences kept or the number of variable sites. We first used the
211 default value 1 for the prioritising parameter. As some of the sequences got very short
212 (datasets def; Additional files Table S2), we produced additional subsets using a value of 0.5,
213 retaining more variable sites and fewer sequences (datasets seq). This option was used for
214 *arenosa* and *lyrata*, but it could not be used for the total and diploid datasets, because it
215 reduced the number of sequences of diploid *arenosa* to three or less (Table S2). Reducing the
216 data to non-recombining blocks reduces the amount of data and leads to a loss of diversity,
217 which may lead to a downward bias in estimates of effective population sizes obtained from
218 IM. Divergence time and gene flow estimates have, however, been shown to be largely
219 unaffected [68].
220 The parameter estimates provided by IMa2 are scaled by the mutation rate. In order to
221 convert them to demographic estimates, a mutation rate needs to be assumed. We followed
222 the procedure of Slotte et al. [21]: we assumed a substitution rate of $6.5 \cdot 10^{-9}$ [69, 70] as a
223 lower boundary for the mutation rate and the synonymous substitution rate of $1.5 \cdot 10^{-8}$ per site
224 per generation [71] as an upper boundary, and used the mean of these two estimates to
225 calculate per-fragment mutation rates (Additional files Table S3).
226 IMa2 uses a Bayesian approach and Markov Chain Monte Carlo (MCMC)
227 simulations to estimate parameters. Priors for effective population sizes ($q = 4N_e*\mu$, where μ
228 is the migration rate per fragment, not per bp), time since divergence ($t = time*\mu$) and
229 migration rates ($m=$migration rate$/\mu$) were initially chosen as recommended in the IM manual
230 [72] and adjusted according to the results of preliminary analyses. For the final runs the
231 following values were used (all values are scaled by $\mu=1.21 \cdot 10^{-5}$): q=17, t=10 and m=5 for
232 *arenosa*, q=20, t=7 and m=5 for *lyrata*, q=15, t=10 and m=5 for the total dataset, and q=12,
233 t=10 and m=2 for the diploid dataset. The number of chains and the heating scheme was also

tested in several preliminary runs. For the final runs we used 20 chains and heating parameters of ha=0.96 and hb=0.9 for analyses with two populations and 80 chains, and ha=0.999 and hb=0.3 for analyses with four populations. The length of the burnin was 1 000 000 steps and estimates were based on between 10 and 27 million steps. Mixing was assessed by trend plots for estimates over the runs and by effective sample size (ESS) values. Convergence was assessed by repeating runs several times with different random seeds. Some IMa2 runs were performed on BioHPC, Computational Biology Service Unit, Cornell University.

Results

DNA content

The ploidy levels for the 14 populations examined are given in Table 1. Only two populations showed signs of more than one ploidy level. The tetraploid *arenosa* population a4_AUT1 from Kernhof in Austria included one diploid individual, and the diploid *lyrata* population l2_CZE from NW Ivanice in the Czech Republic contained one triploid. The T-test showed that the two taxa had significantly different DNA content within ploidy levels for both diploids and tetraploids, *lyrata* having a slightly larger nuclear DNA content than *arenosa* in both cases (0.23 vs. 0.20, and 0.44 vs. 0.43, $P<0.001$).

Sequence variation and diversity

The obtained sequences were deposited to GenBank with accession numbers GQ386471-GQ386654; 75 sequences of CHS, 59 sequences of *sc*ADH, and 32 sequences of *trn*L-F (Additional files Table S1). Substantial recombination has taken place for both low-copy

257 nuclear regions. Minimum number of recombination events was 16 for CHS and 24 for
258 *sc*ADH (Additional files Table S1). For the plastid *trn*L-F region, only a single
259 recombination event was detected.
260 When analysed alone, the diploids were separated into two groups corresponding to
261 named taxa in the neighbour networks based on the nuclear markers (Fig. 2a,c), and partly
262 also the plastid region (Fig. 2e). CHS split the *lyrata* diploids into two distinct groups with
263 absolutely no geographical structure; both clusters included specimens from Germany, the
264 Czech Republic, and Austria (Fig. 2a). There was no apparent geographical structure among
265 *arenosa* specimens either. The analysis of *sc*ADH gave no additional information (Fig. 2d).
266 The two taxa didn't share *trn*L-F haplotypes, but all three *lyrata* haplotypes clustered closer to
267 *arenosa* than to each other (Fig. 2e).
268 Adding the tetraploids to the neighbour networks complicated the picture (Fig. 2b, d,
269 f). The majority of the tetraploids clustered according to taxa; tetraploid *lyrata* clustered with
270 diploid *lyrata*, and tetraploid *arenosa* with diploid *arenosa*. There were, however, exceptions
271 for all the marker systems. The CHS network grouped five tetraploid *lyrata* sequences with
272 tetraploid *arenosa* (Fig. 2b). These represent three specimens (with a mixture of *lyrata*- and
273 *arenosa*-like alleles; Additional files Table S4) of which two are from the same population
274 (l4_AUT1), collected in Wachau, and the last one from Schrambach, also in Lower Austria
275 (population l4_AUT4). In the *sc*ADH network, one of the specimens from the Wachau
276 population (l4_AUT1_11) shared an allele with a tetraploid *arenosa* collected just a few
277 kilometres away (a4_AUT2_18), whereas the specimen from the Schrambach population
278 (14_AUT4_2) clustered with a tetraploid *arenosa* from Wachau (a4_AUT4_15; Fig. 2d;
279 Additional files Table S4).
280 In the CHS network ten tetraploid *arenosa* sequences clustered with the *lyrata* groups
281 (Fig. 2b). These represent seven specimens (with a mixture of *lyrata*- and *arenosa*-like

alleles; Additional files Table S4), most of them from Lower Austria (populations a4_AUT1, a4_AUT2, and a4_AUT4), but a single one from Germany (a4_GER). Only one of these tetraploid *arenosa* specimens (a4_AUT1_2) contained a *lyrata*-like *sc*ADH allele and clustered with *lyrata* in the network (Fig. 2d; Additional files Table S4).

The plastid *trn*L-F network separated specimens according to taxa with one exception: the same tetraploid *lyrata* specimen from Wachau (l4_AUT1_11), which clustered with *arenosa* also in the CHS and *sc*ADH networks, shared a haplotype with diploid and tetraploid *arenosa* (a2_SVK2, a4_AUT2, and a4_GER, Fig. 2f).

To summarise, these networks basically told the same story with major splits between *lyrata* and *arenosa*, and with ploidy levels to a high degree intermingled within each taxon. Deviations from this pattern were found more or less in the same populations for the different markers; tetraploid *arenosa*: a4_AUT1, a4_AUT2, a4_AUT4 and tetraploid *lyrata*: l4_AUT1 and l4_AUT4, all populations from Lower Austria where the two taxa are sympatric (Table 1; Additional files Table S4).

Analysing the taxa separately for all marker systems showed that the specimens did not cluster according to ploidy level (Fig. 3). For both species, the CHS and *sc*ADH networks separated groups of a few tetraploids from the remaining specimens, corresponding to the deviations mentioned above (Fig. 3a, b). Otherwise specimens of different ploidy levels are completely intermingled.

Both nuclear regions showed high levels of gene diversity (Hd = 0.93-0.99; Fig. 4, Additional files Table S5). There were no large differences among the duplicated tetraploid datasets (D1-22), and no clear differences between taxa or ploidy level. For *trn*L-F, the diversity was somewhat lower (Hd = 0.60-0.71), particularly for *arenosa*. Different patterns of nucleotide diversity (π; Fig. 4, Additional files Table S5) were found among the markers. For CHS the level of diversity was the same for 2*x arenosa*, and 2*x* and 4*x lyrata*, whereas 4*x*

307 *arenosa* was more diverse. For *sc*ADH the pattern was different: *arenosa* was more diverse
308 than *lyrata*, and while diploid *arenosa* was more diverse than tetraploid *arenosa*, it was the
309 other way around in *lyrata*. For *trn*L-F there was no difference in diversity between ploidy
310 levels within taxa, but *lyrata* was slightly more diverse than *arenosa*. The average number of
311 nucleotide differences (*k*) showed a similar pattern as nucleotide diversity for the nuclear
312 regions, except that tetraploid *lyrata* was more diverse than diploid *lyrata*. For *trn*L-F, *k* was
313 generally very low, but higher for tetraploid than diploid *arenosa*, and the other way for
314 *lyrata*.
315

316 **Isolation with migration results**

317 For analyses of pairs of populations all runs reached ESS values > 1000 and mixing seemed
318 good based on trend plots. Repeated runs indicated good convergence. The analyses with four
319 populations did, however, not perform equally well and ESS values remained < 50 for several
320 parameters. We will therefore in the first place base our conclusions on the pairwise runs, and
321 only mention the results of the four population run as indicative.
322 The main aim of this study was to assess evidence for gene flow between ploidy
323 levels and species. For *arenosa*, the IM analysis revealed strong support for gene flow from
324 diploids to tetraploids, but not in the other direction (Fig. 5, Additional files Fig. S1, S2). The
325 95% highest posterior density intervals (HPD) for the migration rate from diploids to
326 tetraploids excluded 0 for seven of eight datasets where it was estimated reliably (four
327 variants of tetraploid genotypes x two options of largest non-recombinant blocks; Additional
328 files Table S2), whereas the estimate of m was at the lowest value for migration from
329 tetraploids to diploids in all cases. Estimates of the number of migrants from diploids to
330 tetraploids were between 2.1 and 4.1, but HPD intervals were large and overlapped
331 considerably among both migration directions (Additional files Table S2). The posterior

distribution for time since divergence did not go down towards 0 at the upper limit of the prior interval, independent of prior choice, and divergence time could thus not be properly estimated. Effective population size estimates were 1.5 to 2 times higher for tetraploids than for diploids, but as these are estimates of the effective number of genes, the estimated number of tetraploid individuals was in fact somewhat lower than for diploids.

For *lyrata*, there was also clear support for gene flow between the ploidy levels. HPD intervals for gene flow from diploids to tetraploids excluded 0 for five of eight datasets (Additional files Table S2) and all estimates of m were larger than 0. For gene flow from tetraploids to diploids, HPD distributions did not reach low levels at high values for gene flow from tetraploids to diploids, making them unreliable. Still the HPD intervals excluded 0 for three out of four A datasets, and estimates of m were at the lowest point of the distribution only for one A dataset and two B datasets, indicating gene flow from tetraploids to diploids. The discrepancies between the different datasets resulted from the fact that different parts of the sequences were kept by IMgc (Additional files Tables S2). The effective number of migrants per generation was estimated as 1.5 to 1.8 from diploids to tetraploids and as 1.2 to 1.4 from tetraploids to diploids. As for *arenosa*, the posterior distributions for time since divergence were not unimodal and did not provide any reliable estimates. Estimates of effective population sizes varied also somewhat between datasets and differed in general not between ploidy levels.

The results of the analysis including all four taxonomic entities (Additional files Table S2) were largely consistent with the results of the pairwise runs, although these estimates have to be considered unreliable due to poor performance of the MCMC (and are therefore not shown). In addition, these runs suggested gene flow between species, from diploid and tetraploid *arenosa* into tetraploid *lyrata*, and for one dataset also from diploid *lyrata* into tetraploid *arenosa*.

Analysing the two diploid species together (Additional files Table S2) showed that there was no evidence what so ever for gene flow between them.

Discussion

The overall picture indicated by our results is the following. Both the neighbour nets and the IM analysis show that there is no gene flow between the two diploid taxa. There is, however, evidence for gene flow from diploids to tetraploids in *arenosa* and possibly for interploidal gene flow in both directions in *lyrata*. This is consistent with the intermingling of sequences of both ploidy levels revealed by the networks. The possibility of gene flow from tetraploids to diploids in *lyrata* is further supported by the triploid individual found in population l2_CZE. The networks including both species and ploidy levels clearly indicate mixing of lineages between species. As there seems to be no gene flow between diploids, we assume that this mixing resulted from gene flow into tetraploids. Consistent with this assumption, the IMa2 analysis of all four taxonomic entities suggest gene flow from diploid and tetraploid *arenosa* into tetraploid *lyrata*, and possibly, with much lower frequency, also from diploid *lyrata* into tetraploid *arenosa*.

Interspecific gene flow

It has been suggested that genotypes of tetraploids are buffered against the shock of absorbing foreign genomes, and that extensive introgression often takes place at the tetraploid level between species that are isolated from each other at the diploid level [73, 74]. Our analyses of *Arabidopsis* in Central Europe show that *arenosa* and *lyrata* are good biological species at the diploid level. The network analyses show no sharing of alleles, and the main splits are between the two taxa, which is in agreement with a comprehensive large-scale

analysis of the genus [75]. Furthermore, isolation with migration (IM) analyses of diploids show no gene flow from 2*x arenosa* to *lyrata* or the other way around.

We do, however, find signs of introgression in the tetraploids on both sides; several tetraploid *lyrata* sequences cluster with *arenosa* in the network analyses (especially for the CHS region, but also for the other two regions), and several tetraploid *arenosa* sequences cluster with *lyrata*.

To our knowledge there are not many studies that have dealt with gene flow between sister species that contain two (or more) ploidy levels. Luttikhuizen *et al.* [23] found higher genetic diversity in autotetraploid *Rorippa amphibia* (L.) Besser compared to conspecific diploids using microsatellites, and suggested that introgression as well as multiple origins of the tetraploids might have contributed to the tetraploid diversity. Stift *et al.* [76] used crossing experiments to show that there are limited reproductive barriers between *R. amphibia* and the sympatric tetraploid *R. sylvestris* (L.) Besser, and concluded that gene flow between the two tetraploids is the probable reason for the high diversity found in tetraploid *R. amphibia* by Luttikhuizen *et al.* [23]. Kloda *et al.* [77] studied gene flow among diploid *Ononis* (*O. spinosa* L. and *O. intermedia* C.A.Mey. ex. A.K.Becker) and tetraploids (*O. repens* L. and *O. maritima* Dumort.) using microsatellites, and found that there were restrictions to gene flow between, but not within the ploidy levels. In the genus *Paeonia* L. homoploid hybrid species have been derived from allotetraploids, but not from the diploid progenitors, suggesting that chromosomal structural differences induced by polyploidy might create new opportunities for interspecific gene flow [78]. In line with this previous evidence, our study suggests that polyploids might tolerate introgression better than their diploid progenitors, as suggested by de Wet & Harlan [73] and Harlan & de Wet [74].

Interploidal gene flow

In general, high genetic diversity in polyploids, as our network analyses indicated for both *arenosa* and *lyrata*, can be explained by multiple independent origins of the polyploids, continuing introgression from the diploids, introgression from other polyploid species (see above), or as a result of subsequent evolution following one or more old polyploidisation events [see e.g., 3, 6]. Our network analyses for each species show no apparent clustering according to ploidy levels; i.e. we do not see a separation of diploids and tetraploids. For *arenosa* our IM analyses suggest that the diploids and the tetraploids are two distinct groups, with some migration from the diploid to the tetraploid level subsequent to the separation of the two gene pools. As the diploids from the Carpathians and the tetraploids from the Eastern Austrian Forealps and Germany are no longer sympatric, the migration we observe is probably the result of recurrent origin/introgression in the past. For *lyrata* our IM analysis suggests that the gene flow is bidirectional.

In our flow cytometry data, we find signs of mixed-ploidy populations: a triploid plant was detected in the mainly diploid l2_CZE population of *lyrata* from the Czech Republic, and a diploid plant was detected in the mainly tetraploid a4_AUT1 population of *arenosa* from the Eastern Austrian Foreland. Also, the Central European *Arabidopsis* populations often have restricted distributions, and populations of different ploidy levels are sometimes only a few kilometres apart. Thus, introgression is possible, especially along disturbed sites. However, the generally low migration rates suggest that introgression remains a rare event. If this is the case, the small genus *Arabidopsis* contains polyploids with very different evolutionary fates when it comes to introgression. The allotetraploid *A. suecica* (Fr.) Norrl. ex O.E.Schulz has very low genetic diversity and has probably originated only once [34, 79]. The allotetraploid *A. kamchatica* (Fisch. ex DC.) K.Shimizu & Kudoh has probably originated multiple times, and additionally experienced some later introgression from the

parental diploids [14, 35]. The tetraploid *arenosa* shows some signs of introgression, and the tetraploid *lyrata* shows signs of extensive ongoing introgression (cf. Schmickl and Koch. submitted). We still do not have proper criteria or methods to distinguish between recurrent formation and introgression from progenitors as sources of diversity in polyploids [6]. In this study we tried a new analysis method [65] in an attempt to make this distinction. Although estimates were not always consistent and credible intervals were large, the results provided indications that the two species are placed in different positions along the gradient of possible polyploid evolution models. Furthermore, what the results indicate suggests that the genus *Arabidopsis* might be a good model system to use for developing the criteria called for by Soltis *et al.* [6].

Gene flow from diploids to tetraploid derivatives has long been acknowledged as relatively common [see e.g., 2, 3, 6 and references therein, 12]. The question of gene flow in the opposite direction is more controversial. Stebbins [80] states that interploidal gene flow is usually unidirectional from diploids to tetraploids for two reasons: 1) offspring of triploid hybrids are usually tetraploid or close to it in chromosome number, and 2) diploids and tetraploids are often so highly incompatible that triploid offspring cannot be formed at all. However, studies involving natural triploids in euploid hybrid swarms show that triploids may produce 1x, 2x, and 3x gametes, and may therefore contribute to gene flow in both directions [19, 20, 81]. Indeed, gene flow from tetraploids to diploids has been observed in some taxa, e.g. *Dactylorhiza maculata* (L.) Soó [82] and *Betula* L. [83]. In our study, we did not observe gene flow from tetraploid to diploid *arenosa* (m = 0), but as the different ploidy levels are allopatric, we cannot distinguish between genetic/genomic and geographic barriers. However, for *lyrata*, with diploids and polyploids in close proximity, gene flow seems to be bidirectional.

Conclusions

In this study, we looked at the effect of polyploidisation on interspecific introgression, and interploidal gene flow using Central European *Arabidopsis* as a model system. There was no evidence for interspecific gene flow between 2*x arenosa* and 2*x lyrata* which can be considered as good biological entities, but some support for gene flow into 4*x lyrata* and possibly 4*x arenosa*. Thus, whole genome duplication might decrease vulnerability to interspecific hybridisation and buffer negative effects of introgression. Interploidal gene flow was detected from 2*x* to 4*x* in both species, and from 4*x* to 2*x* in *lyrata*. For *arenosa*, the two ploidy levels are allopatric, and the lack of gene flow could be the result of geographic as well as genetic barriers. In *lyrata*, however, where geographic barriers are limited, gene flow is bidirectional.

Authors' contributions

MHJ carried out the molecular work and statistical analyses, and drafted the manuscript. DE carried out the isolation with migration analyses and helped to draft the manuscript. RS and MAK provided the samples and contributed to draft the manuscript. AKB was project leader, contributed in the molecular work, and drafted the manuscript.

Acknowledgement

The authors would like to thank Cecilie Mathiesen for assistance in the lab, and Barbara Mable and the botanical group in Oslo for reading and commenting on the manuscript.

Literature cited

1. Soltis DE, Soltis PS: **Molecular data and the dynamic nature of polyploidy**. *Critical Reviews in Plant Sciences* 1993, **12**(3):243-273.
2. Soltis DE, Soltis PS: **Polyploidy: recurrent formation and genome evolution**. *Trends in Ecology & Evolution* 1999, **14**(9):348-352.
3. Soltis DE, Soltis PS, Tate JA: **Advances in the study of polyploidy since *Plant speciation***. *New Phytologist* 2003, **161**:173-191.
4. Fawcett JA, Maere S, Van de Peer Y: **Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(5737-5742).
5. Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH: **The frequency of polyploid speciation in vascular plants**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(33):13875-13879.
6. Soltis DE, Buggs RJA, Doyle JJ, Soltis PS: **What we still don't know about polyploidy**. *Taxon* 2010, **59**(5):1387-1403.
7. Stift M, Berenos C, Kuperus P, Tienderen PHv: **Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to *Rorippa* (yellow cress) microsatellite data**. *Genetics* 2008, **179**:2113-2123.
8. Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS: **Autopolyploidy in angiosperms: have we grossly underestimated the number of species?** *Taxon* 2007, **56**(1):13-30.
9. O'Kane SL, Al-Shehbaz IA: **A synopsis of *Arabidopsis* (Brassicaceae)**. *Novon* 1997, **7**:323-327.
10. Doyle JJ, Doyle JL, Rauscher JT, Brown AHD: **Evolution of the perennial soybean polyploid complex (*Glycine* subgenus *Glycine*): a study of contrasts**. *Biological Journal of the Linnean Society* 2004, **82**:583-597.
11. Hendry AP: **Evolutionary biology: speciation**. *Nature* 2009, **458**:162-164.
12. Chapman MA, Abbott RJ: **Introgression of fitness genes across a ploidy barrier**. *New Phytologist* 2010, **186**:63-71.
13. Kim S-T, Sultan SE, Donoghue MJ: **Allopolyploid speciation in *Persicaria* (Polygonaceae): Insights from a low-copy nuclear region**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(34):12370-12375.
14. Shimizu-Inatsugi R, Lihová J, Iwanaga H, Kudoh H, Marhold K, Savolainen O, Watanabe K, Yakubov VV, Shimizu KK: **The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri***. *Molecular Ecology* 2009, **18**(19):4024-4048.
15. Symonds VV, Soltis PS, Soltis DE: **Dynamics of polyploid formation in *Tragopogon* (Asteraceae): recurrent formation, gene flow, and population structure**. *Evolution* 2010, **64**(7):1984-2003.
16. Soltis DE, Soltis PS: **Genetic consequences of autopolyploidy in *Tolmiea* (Saxifragaceae)**. *Evolution* 1989, **43**:586-594.

17. Segraves KA, Thompson JN, Soltis PS, Soltis DE: **Multiple origins of polyploidy and the geographic structure of *Heuchera grossulariifolia***. *Molecular Ecology* 1999, **8**:253-262.
18. Parisod C, Besnard G: **Glacial *in situ* survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigita* (Brassicaceae)**. *Molecular Ecology* 2007, **16**:2755-2767.
19. Husband BC: **The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations**. *Biological Journal of the Linnean Society* 2004, **82**:537-546.
20. Husband BC, Sabara HA: **Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae)**. *New Phytologist* 2003, **161**:703-713.
21. Slotte T, Huang H, Lascoux M, Ceplitis A: **Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae)**. *Molecular Biology and Evolution* 2008, **25**(7):1472-1481.
22. Abbott RJ, Ireland HE, Rogers HJ: **Population decline despite high genetic diversity in the new allopolyploid species *Senecio cambrensis* (Asteraceae)**. *Molecular Ecology* 2007, **16**:1023-1033.
23. Luttikhuizen PC, Stift M, Kuperus P, Tienderen PHv: **Genetic diversity in diploid vs. tetraploid *Rorippa amphibia* (Brassicaceae)**. *Molecular Ecology* 2007, **16**:3544-3553.
24. Facey PD, Lee PLM, Smith MNE, Hipkin CR: **Conservation of genetic diversity in British populations of the diploid endemic *Coincya monensis* ssp. *monensis* (Isle of Man Cabbage): the risk of hybridisation with the tetraploid alien, *Coincya monensis* ssp. *cheiranthos***. *Conservation Genetics* 2007, **8**:1029-1042.
25. Koch MA: **Genetic differentiation and speciation in prealpine *Cochlearia* (Brassicaceae): Allohexaploid *Cochlearia bavarica* (Brassicaceae) compared to its diploid acestor *Cochlearia pyrenaica* in Germany and Austria**. *Plant Systematics and Evolution* 2002, **232**(1-2):35-49.
26. Koch MA, Bernhardt K-G: **Comparative biogeography of the cytotypes of annual *Microthlaspi perfoliatum* (Brassicaceae) in Europe using isozymes and cpDNA data: refugia, diversity centers, and postglacial colonization**. *American Journal of Botany* 2004, **91**(1):115-124.
27. Hedrén M, Nordström S, Ståhlberg D: **Polyploid evolution and plastid DNA variation in the *Dactylorhiza incarnata*/*maculata* complex (Orchidaceae) in Scandinavia**. *Molecular Ecology* 2008, **17**:5075-5091.
28. Paun O, Greilhuber J, Temsch EM, Hörandl E: **Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae)**. *Molecular Ecology* 2006, **15**:897-910.
29. Paun O, Stuessy TF, Hörandl E: **The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex**. *New Phytologist* 2006, **171**:223-236.
30. Jørgensen MH, Elven R, Tribsch A, Gabrielsen TM, Stedje B, Brochmann C: **Taxonomy and evolutionary relationships in the *Saxifraga rivularis* complex**. *Systematic Botany* 2006, **31**(4):702-729.
31. Stebbins GL: **Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach**. *Botanica Helvetica* 1984, **94**(1):1-13.
32. Shimizu KK, Fujii S, Marhold K, Watanabe K, Kudoh H: ***Arabidopsis kamchatica* (Fisch. ex DC.) K.Shimizu & Kudoh and *A. kamchatica* subsp. *kawasakiana* (Makino) K.Shimizu & Kudoh, new combinations**. *Acta Phytotaxonomica et Geobotanica* 2005, **56**(2):163-172.

33. O'Kane SL, Schaal BA, Al-Shehbaz IA: **The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences**. *Systematic Botany* 1996, **21**(4):559-566.
34. Jakobsson M, Hagenblad J, Tavaré S, Säll T, Halldén C, Lind-Halldén C, Nordborg M: **A unique recent origin of the allotetraploid species *Arabidopsis suecica*: evidence from nuclear DNA markers**. *Molecular Biology and Evolution* 2006, **23**(6):1217-1231.
35. Schmickl R, Jørgensen MH, Brysting AK, Koch MA: **The evolutionary history of the *Arabidopsis lyrata* complex: A hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier**. *BMC Evolutionary Biology* 2010, **10**:98.
36. Berkutenko AN, Tzytlenok SI, Pulkina SV: **Chromosome numbers and dispersal of the Brassicaceae family in the Magadan district**. *Botanicheskiy Zhurnal* 1984, **69**(1):75-80.
37. Elven R, Murray DF, Razzhivin VY, Yurtsev BA: **Checklist of the Panarctic Flora (PAF) vascular plants**. Oslo: National Centre for Biosystematics, University of Oslo; 2008.
38. Koch MA, Wernisch M, Schmickl R: ***Arabidopsis thaliana*'s wild relatives: an updated overview on systematics, taxonomy and evolution**. *Taxon* 2008, **57**(3):933-943.
39. Polatschek A: **Cytotaxonomische Beiträge zur Flora der Ostalpenländer, I**. *Österreichische botanische Zeitschrift* 1966, **113**:1-46.
40. Zhukova PG, Petrovsky VV, Plieva TV: **The chromosome numbers and taxonomy of some plant species from Siberia and Far East**. *Botanicheskiy Zhurnal* 1973, **58**:1331-1342.
41. Zhukova PG, Petrovsky VV: **Chromosome numbers and taxonomy of some plant species from the northern Asia regions**. *Botanicheskiy Zhurnal* 1987, **72**(12):1617-1624.
42. Hu TT, Pattyn P, Bakker EG, Cao J, Cheng J-F, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H *et al*: **The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change**. *Nature Genet* 2011, **43**(5):476-481.
43. Dobeš C, Vitek E: **Documented chromosome number checklist of Austrian vascular plants**. Wien: Verlag des Naturhistorischen Museums Wien; 2000.
44. Schmickl R: **Reticulate evolution in glacial refuge areas - the genus *Arabidopsis* in the eastern Austrian Danube Valley (Wachau)**. *PhD thesis.* Heidelberg: Ruperto-Carola University of Heidelberg; 2009.
45. Doležel J, Suda J (eds.): **Flow Cytometry with Plant Cells**. Weinheim: Wiley VCH Verlag GmbH & Co. KGaA; 2007.
46. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. *Nucleic Acids Symposium Series* 1999, **41**:95-98.
47. Schmickl R, Kiefer C, Dobeš C, Koch MA: **Evolution of *trn*F(GAA) pseudogenes in cruciferous plants**. *Plant Systematics and Evolution* 2008, **282**(3-4):229-240.
48. Koch MA, Dobeš C, Matschinger M, Bleeker W, Vogel J, Kiefer M, Mitchell-Olds T: **Evolution of the *trnF*(GAA) gene in *Arabidopsis* relatives and the Brassicaceae family: monophyletic origin and subsequent diversification of a plastidic pseudogene**. *Molecular Biology and Evolution* 2005, **22**(4):1032-1043.
49. Bradley RD, Hillis DM: **Recombinant DNA sequences generated by PCR amplification**. *Molecular Biology and Evolution* 1997, **14**(5):592-593.

50. Judo MSB, Wedel AB, Wilson C: **Stimulation and suppression of PCR-mediated recombination**. *Nucleic Acids Research* 1998, **26**(7):1819-1825.
51. Popp M, Oxelman B: **Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences**. *Molecular Phylogenetics and Evolution* 2001, **20**(3):474-481.
52. Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF: **PCR-mediated recombination in amplification products derived from polyploid cotton**. *Theoretical and Applied Genetics* 2002, **104**:482-489.
53. Brysting AK, Mathiesen C, Marcussen T: **Challenges in polyploid phylogenetic reconstruction: a case story from the arctic-alpine *Cerastium alpinum* complex**. *Taxon* 2011, **60**(2):333-347.
54. Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM, Zhou J: **Evaluation of PCR-generated chimeras, mutations and heteroduplexes with 16S rRNA gene-based cloning**. *Applied and Environmental Microbiology* 2001, **67**(2):880-887.
55. Kelly LJ, Leitch AR, Clarkson JJ, Hunter RB, Knapp S, Chase MW: **Intragenic recombination events and evidence for hybridspeciaton in *Nicotiana* (Solanaceae)**. *Molecular Biology and Evolution* 2010, **27**(4):781-799.
56. Schierup MH, Hein J: **Consequences of recombination on traditional phylogenetic analysis**. *Genetics* 2000, **156**:879-891.
57. Hudson RR, Kaplan NL: **Statistical properties of the number of recombination events in the history of a sample of DNA sequences**. *Genetics* 1985, **111**:147-164.
58. Librado P, Rozas J: **DnaSP v5: A software for comprehensive analysis of DNA polymorphism data**. *Bioinformatics* 2009, **25**:1451-1452.
59. Bryant D, Moulton V: **Neighbor-net: an agglomerative method for the construction of phylogenetic networks**. *Molecular Biology and Evolution* 2004, **21**(2):255-265.
60. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies**. *Molecular Biology and Evolution* 2006, **23**(2):254-267.
61. Simmons MP, Ochoterena H: **Gaps as characters in sequence-based phylogenetic analyses**. *Systematic Biology* 2000, **49**(2):369-381.
62. Müller K: **SeqState: Primer design and sequence statistics for phylogenetic DNA datasets**. *Applied Bioinformatics* 2005, **4**(1):65-69.
63. De Silva HN, Hall AJ, Rikkerink E, McNeilage MA, Fraser LG: **Estimation of allele frequencies in polyploids under certain patterns of inheritance**. *Heredity* 2005, **95**:327-334.
64. Hey J, Nielsen R: **Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics**. *proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(8):2785-2790.
65. Hey J: **Isolation with migration models for more than two populations**. *Molecular Biology and Evolution* 2010, **27**(4):905-920.
66. Hudson RR: **Estimating the recombination parameter of a finite population model without selection**. *Genetical Research* 1987, **50**:245-250.
67. Woerner AE, Cox MP, Hammer MF: **Recombination-filtered genomic datasets by information maximization**. *Bioinformatics* 2007, **23**:1851-1853.
68. Strasburg JL, Rieseberg LH: **How robust are "isolation with migration" analyses to violations of the IM model? A simulation study**. *Molecular Biology and Evolution* 2010, **27**(2):297-310.

69. Wolfe KH, Li W-H, Sharp PM: **Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs**. *Proceedings of the National Academy of Sciences of the United States of America* 1987, **84**:9054-9058.
70. Wright SI, Lauga B, Charlesworth D: **Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata***. *Molecular Ecology* 2003, **12**:1247-1263.
71. Koch MA, Haubold B, Mitchell-Olds T: **Molecular systematics of the Brassicaceae: evidence from coding plastidic *MatK* and nuclear *Chs* sequences**. *American Journal of Botany* 2001, **88**(2):534-544.
72. **Documentation for IMa2** [http://genfaculty.rutgers.edu/hey/software]
73. de Wet JMJ, Harlan JR: **Apomixis, polyploidy, and speciation in *Dichanthium***. *Evolution* 1970, **24**(270-277).
74. Harlan JR, de Wet JMJ: **On Ö. Winge and a prayer: the origins of polyploidy**. *The Botanical Review* 1975, **41**(4):361-390.
75. Koch MA, Matschinger M: **Evolution and genetic differentiation among relatives of *Arabidopsis thaliana***. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(15):6272-6277.
76. Stift M, Bregman R, Oostermeijer JGB, van Tienderen PH: **Other tetraploid species and conspecific diploids as sources of genetic variation for an autotetraploid**. *American Journal of Botany* 2010, **97**(11):1858-1866.
77. Kloda JM, Dean PDG, Maddren C, MacDonald DW, Mayes S: **Using principal component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example from British Restharrows (*Ononis spinosa* and *Ononis repens*)**. *Heredity* 2008, **100**:253-260.
78. Ferguson D, Sang T: **Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*)**. *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98**(7):3915-3919.
79. Lind-Halldén C, Halldén C, Säll T: **Genetic variation in *Arabidopsis suecica* and its parental species *A. arenosa* and *A. thaliana***. *Hereditas* 2002, **136**:45-50.
80. Stebbins GL: **Chromosomal evolution in higher plants**. London: Edward Arnold (Publishers) Ltd.; 1971.
81. Henry IM, Dilkes BP, Young K, Watson B, Wu H, Comai L: **Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid respons**. *Genetics* 2005, **170**:1979-1988.
82. Ståhlberg D: **Habitat differentiation, hybridization and gene flow patterns in mixed populations of diploid and autotetraploid *Dactylorhiza maculata* s.l. (Orchidaceae)**. *Evolutionary Ecology* 2009, **23**:295-328.
83. Thórsson ÆT, Salmela E, Anamthawat-Jónsson K: **Morphological, cytogenetic, and molecular evidence for introgressive hybridisation in birch**. *The Journal of Heredity* 2001, **92**(5):404-408.

Additional files

Additional file 1: Table S1. Dataset summary with primers, GenBank numbers, alignment lengths and minimum number of recombination events per region.

Additional file 2: Table S2. Isolation with migration (IM) analyses: description of datasets.

Additional file 3: Table S3. Isolation with migration analyses: mutation rates per fragment per generation.

Additional file 4: Table S4. Signs of introgression among tetraploid *Arabidopsis arenosa* and *A. lyrata*.

Additional file 5: Table S5. Diversity indeces.

Additional file 6: Figure S1. Isolation with migration analyses: effective population sizes and estimates of time for the def datasets.

Additional file 7: Figure S2. Isolation with migration analyses of the seq datasets.

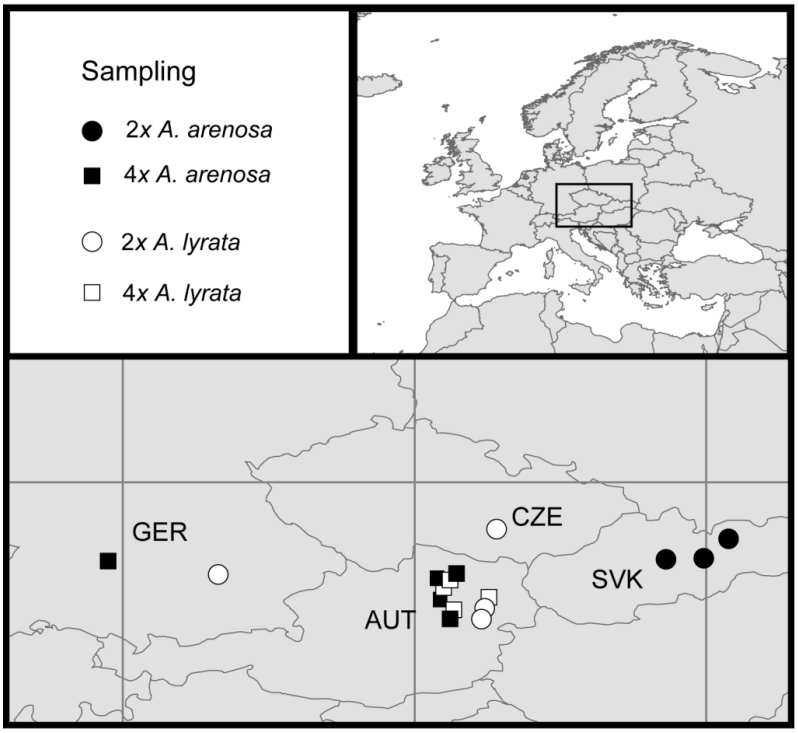
733 **Figure 1** Central European *Arabidopsis* populations included in this study. Taxon is given by
734 color: *A. arenosa* - black and *A. lyrata* - white. Ploidy levels are given as circle - diploid and
735 square - tetraploid. Country names are abbreviated: AUT - Austria, CZE - the Czech
736 Republic, GER - Germany, and SVK - Slovakia.
737
738 **Figure 2** Neighbour network analyses of diploid only (a,c,e) and both diploid and tetraploid
739 (b,d,f) Central European *Arabidopsis arenosa* and *A. lyrata* based on (a,b) the low-copy
740 nuclear CHS region; (c,d) the low-copy nuclear *sc*ADH region; and (e,f) the plastid region
741 *trn*L-F. Only specimens deviating from the majority within each taxon are named. Taxon is
742 given by color: *A. arenosa* -black and *A. lyrata* - grey. Ploidy levels are given as circle -
743 diploid and square - tetraploid.
744
745 **Figure 3** Neighbour network analyses of diploid and tetraploid Central European *Arabidopsis*
746 *arenosa* (a,c,e) and *A. lyrata* (b,d,f) species. (a,b) are based on the low-copy nuclear CHS
747 region; (c,d) are based on the low-copy nuclear *sc*ADH region; and (e,f) are based on the
748 plastid region *trn*L-F. Only specimens identified as deviating from the majority within each
749 taxon in Fig. 2 are named. Taxon is given by color: *A. arenosa* - black and *A. lyrata* - grey.
750 Ploidy levels are given as circle - diploid and square - tetraploid.
751
752 **Figure 4** Diversity analyses of Central European *Arabidopsis arenosa* 2*x* (black) and 4*x*
753 (dark grey), and *A. lyrata* 2*x* (grey) and 4*x* (light grey). (a) Gene diversity, $\hat{H}$; (b) Nucleotide
754 diversity, π; and (c) Average number of nucleotide differences, k.
755
756 **Figure 5** Isolation with migration (IM) analyses of the datasets D1-D22 (see text for details):
757 migration rates. (a) Migration from diploid to tetraploid (grey) and from tetraploid to diploid

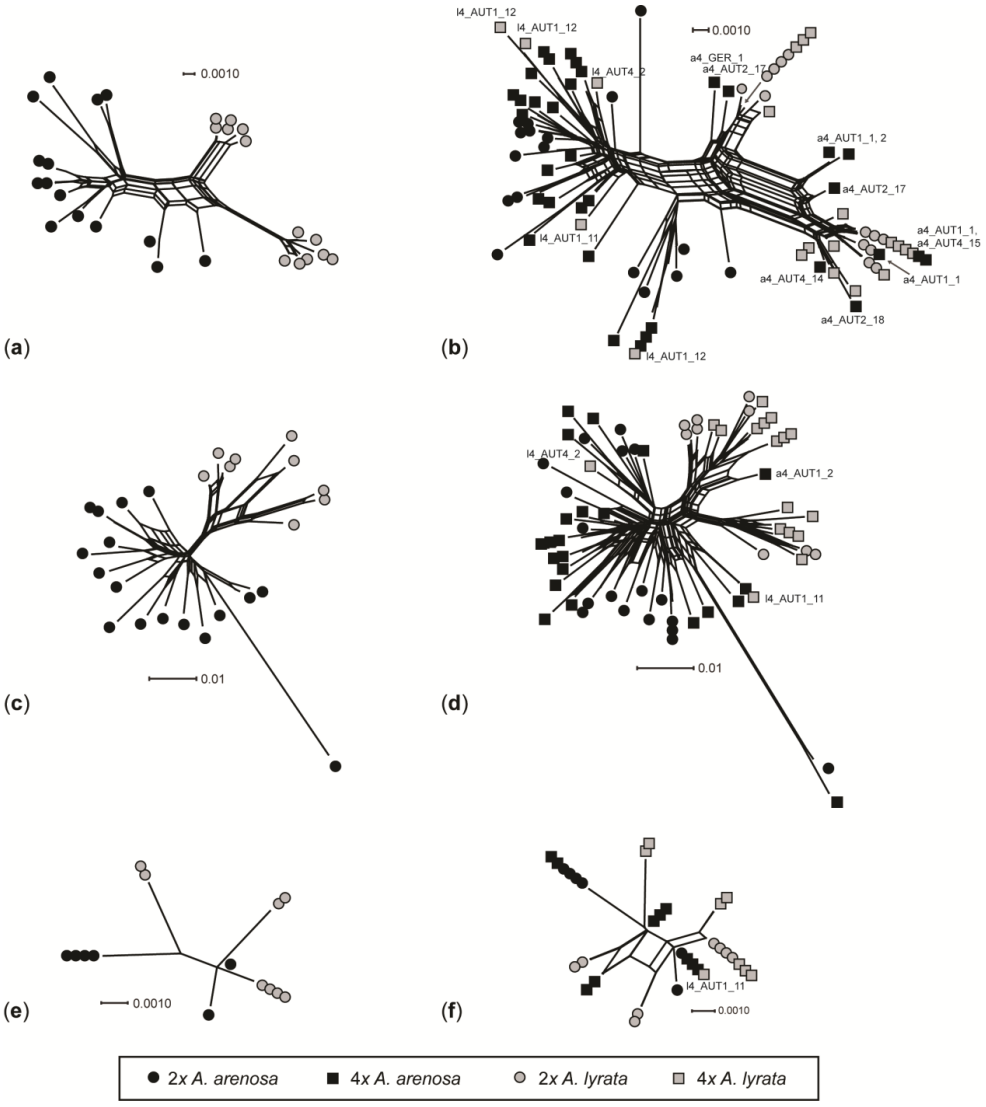
(black) *A. arenosa*. (b) Migration from diploid to tetraploid (grey) and from tetraploid to diploid (black) *A. lyrata*.

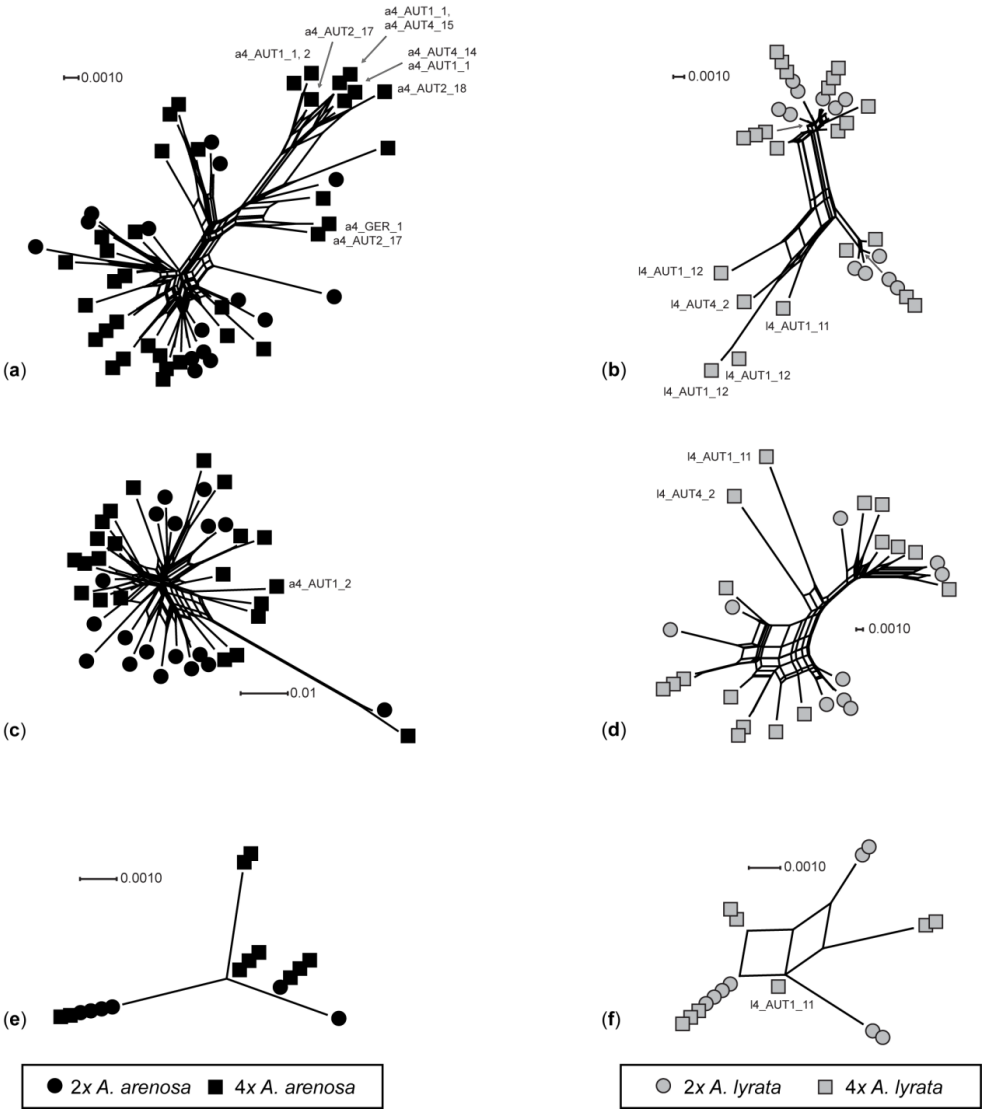
761 **Table 1**. Sampling of *Arabidopsis arenosa* and *A. lyrata* included in this study.

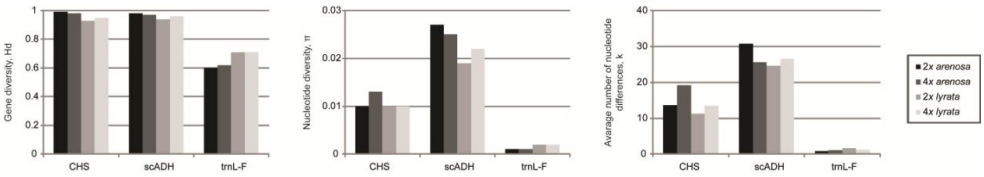
Taxon	Ploidy level	Population[1]	Locality[2]	# FC[3]
arenosa	2*x*	a2_SVK1	SVK: Vysoké Tatry; Prešovský kraj; Belianske Tatry; Zadné Meďodoly Valley; Kopské Sedlo (131)	5
		a2_SVK2	SVK: Nízke Tatry Mts.; Pusté Pole (915140)	30
		a2_SVK3	SVK: Veľká Fatra Mts.; Harmanec; Malý Šturec Sedlo (915141)	33
	4*x*	a4_GER	GER: Southern Germany; Swabian Alps; Wental; Felsenmeer (123)	-
		a4_AUT1[4]	AUT: Lower Austria; Eastern Alps; SSW St. Aegyd am Neuwalde; Kernhof; rocky batter next to street opposite depot (81 or 915142)	26
		a4_AUT2	AUT: Lower Austria; Waldviertel; Wachau; NNE Weißenkirchen; Achleiten (3)	12
		a4_AUT3	AUT: Lower Austria; Kamp Valley; S Stiefern; parking site with view on railway bridge (89)	-
		a4_AUT4	AUT: Lower Austria; Waldviertel; Wachau; forest road from Scheibenbach towards Pfaffental (20)	4
lyrata	2*x*	l2_GER	GER: Bavaria; Veldenstein Forest; street from Velden to Pfaffenhofen (112)	17
		l2_CZE[5]	CZE: SW Brno; NW Ivanice; between Nova Ves and Oslavany; slope above Oslava River (96)	9
		l2_AUT1	AUT: Lower Austria; street from Pernitz to Pottenstein (88 or 915143)	43
		l2_AUT2	AUT: Lower Austria; S Vienna; Bad Vöslau; rocks near Vöslauer Hütte (74 or 915145)	28
	4*x*	l4_AUT1	AUT: Lower Austria; Waldviertel; Wachau; E Dürnstein; small hill N Franzosendenkmal (13)	24
		l4_AUT2	AUT: Lower Austria; S Vienna; Mödling; Castle ruin Mödling (66 or 915144)	21
		l4_AUT3	AUT: Lower Austria; Dunkelstein Forest; Wachau; N Bacharnsdorf (50)	21
		l4_AUT4	AUT: Lower Austria; Schrambach between Freiland and Lilienfeld (116)	1

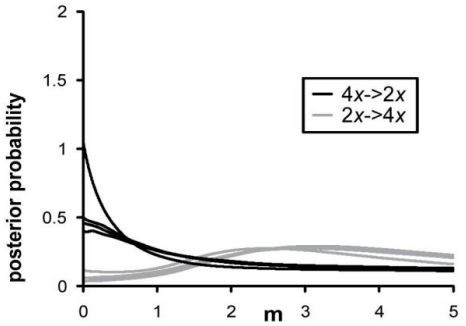
762 [1]Two individuals from each populations were sequenced.
763 [2]Country names are abbreviated as follows: AUT - Austria, CZE - Czech Republic, GER -
764 Germany, and SVK - Slovakia. Brackets following localities give original collection number
765 (Schmickl and Koch submitted).
766 [3]# FC gives the number of individuals analysed with flow cytometry. The populations
767 a4_GER and a4_AUT3 were not included in the flow cytometry analysis, but multi-allelic
768 microsatellite loci suggest they are tetraploid (Schmickl and Koch, unpublished).
769 [4]a4_AUT1 contained a single diploid individual, the others were tetraploid.
770 [5]l2_CZE contained a single triploid individual, the others were diploid.
771


Sampling
2x A. arenosa
4x A. arenosa
2x A. lyrata
4x A. lyrata
GER
CZE
SVK
AUT

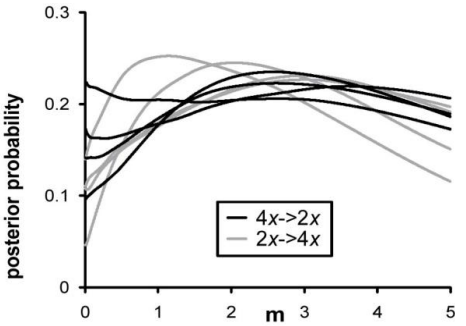

0.0010
(a)
l4_AUT1_12
l4_AUT1_12
0.0010
a4_GER_1
a4_AUT2_17
l4_AUT4_2
a4_AUT1_1, 2
a4_AUT2_17
a4_AUT1_1,
a4_AUT4_15
l4_AUT1_11
a4_AUT4_14
a4_AUT1_1
a4_AUT2_18
l4_AUT1_12
(b)
0.01
(c)
l4_AUT4_2
a4_AUT1_2
l4_AUT1_11
0.01
(d)
0.0010
(e)
l4_AUT1_11
0.0010
(f)
2x A. arenosa
4x A. arenosa
2x A. lyrata
4x A. lyrata


a4_AUT1_1,
a4_AUT4_15
a4_AUT2_17
a4_AUT1_1, 2
a4_AUT4_14
a4_AUT1_1
a4_AUT2_18
0.0010
a4_GER_1
a4_AUT2_17
(a)
0.0010
l4_AUT1_12
l4_AUT4_2
l4_AUT1_11
l4_AUT1_12
l4_AUT1_12
(b)
a4_AUT1_2
0.01
(c)
l4_AUT1_11
l4_AUT4_2
0.0010
(d)
0.0010
(e)
0.0010
l4_AUT1_11
(f)
2x A. arenosa
4x A. arenosa
2x A. lyrata
4x A. lyrata


Gene diversity, Hd
1
0.8
0.6
0.4
0.2
0
CHS
scADH
trnL-F
Nucleotide diversity, π
0.03
0.02
0.01
0
CHS
scADH
trnL-F
Average number of nucleotide differences, k
40
30
20
10
0
CHS
scADH
trnL-F
2x arenosa
4x arenosa
2x lyrata
4x lyrata


posterior probability
2
1.5
1
0.5
0
0
1
2
m
3
4
5
4x->2x
2x->4x

(a)


posterior probability
0.3
0.2
0.1
0
0
1
2
m
3
4
5
4x->2x
2x->4x

(b)

Additional file 1: Table S1. Primers and annealing temperatures used for PCR prior to cloning and sequencing. CHS primers are taken from Koch et al. [1], *sc*ADH primers from Wright et al. [2], and *trn*L-F primers from Taberlet et al. [3]. Length gives the length of the final alignments. # seq gives number of sequences included in the final alignments. R_M gives minimal number of recombination events for each region [4].

Region	GenBank no.	Primer names	Primer sequences (5'-3')	Length /bp	# seq	R_M
CHS	GQ386503-GQ386588	CHS-FOR1 CHS-REV5	CTTCATCTGCCCGTCCATCTAACC GGAACGCTGTGCAAGAC	1452	86	16
*sc*ADH	GQ386589-GQ386654	*sc*ADH-F *sc*ADH-R	GGCATTCCTCCAGCGAC CTTCCGTCGTCGTCTCTTC	1626	67	24
*trn*L-F	GQ386471-GQ386502	c f	CGAAATCGGTAGACGCTACG ATTTGAACTGGTGACACGAG	700	32	1

1. Koch MA, Haubold B, Mitchell-Olds T: **Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae)**. *Molecular Biology and Evolution* 2000, **17**(10):1483-1498.
2. Wright SI, Lauga B, Charlesworth D: **Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata***. *Molecular Ecology* 2003, **12**:1247-1263.
3. Taberlet P, Gielly L, Pautou G, Bouvet J: **Universal primers for amplification of three non-coding regions of chloroplast DNA**. *Plant Molecular Biology* 1991, **17**:1105-1109.
4. Hudson RR, Kaplan NL: **Statistical properties of the number of recombination events in the history of a sample of DNA sequences**. *Genetics* 1985, **111**:147-164.

Additional file 2: Table S2. Isolation with migration (IM) analyses: reduction of datasets to nonrecombining blocks using IMgc. Fasta file indicates which dataset is included: all - including both taxa and ploidy levels, are - including *A. arenosa* only, lyr - including *A. lyrata* only, 2x - including diploids only. See text for explanation of D1-D22. Recombination gives Hudson's recombination parameter R [1] and minimal number of recombination events R_M [2]. We created two different kinds of IM input files through IMgc by choosing different prioritising parameters (pp): def prioritises number of sequences (pp = 1), whereas seq prioritises length of sequences (pp = 0.5). Start and stop gives the positions of the recombining blocks in the original alignments. The final number of sequences in each block is split into diploid and tetraploid *A. arenosa* (a2 and a4), and diploid and tetraploid *A. lyrata* (l2 and l4).

IMgc input					IMa2 input								
Fasta file	Region length/bp	# sequences	Recombination R	R_M	IM input file	Start	Stop	Block length/bp	# sequences a2	a4	l2	l4	Comment
all-CHS-D1.fas	1452	100	21.3	16	all-CHS-D1-def	263	677	414	10	31	14	29	
					all-CHS-D1-seq	203	896	693	3	15	14	24	
all-CHS-D2.fas			22.8	16	all-CHS-D2-def	263	701	438	10	35	14	29	
					all-CHS-D2-seq	203	896	693	6	17	14	24	
all-CHS-D3.fas			22.5	16	all-CHS-D3-def	263	701	438	9	33	14	28	
					all-CHS-D3-seq	215	953	738	2	18	14	24	too little a2
all-CHS-D22.fas			22.9	16	all-CHS-D22-def	263	776	513	8	32	14	27	
					all-CHS-D22-seq	263	990	727	1	13	14	26	too little a2
are-CHS-D1.fas	1452		39	13	are-CHS-D1-def	215	701	486	10	35			
					are-CHS-D1-seq	215	761	546	8	33			
are-CHS-D2.fas					are-CHS-D2-def	215	701	486	10	36			
					are-CHS-D2-seq	215	776	561	8	34			

are-CHS-D3.fas					are-CHS-D3-def	215	701	486	9	34		
					are-CHS-D3-seq	215	776	561	7	31		
are-CHS-D22.fas					are-CHS-D22-def	215	701	486	10	35		
					are-CHS-D22-seq	215	776	561	8	34		
lyr-CHS-D1.fas	1452		2	10	lyr-CHS-D1-def	423	1452	1029			16	25
					lyr-CHS-D1-seq	215	1452	1237			14	25
lyr-CHS-D2.fas					lyr-CHS-D2-def	423	1452	1029			16	25
					lyr-CHS-D2-seq	215	1452	1237			14	25
lyr-CHS-D3.fas					lyr-CHS-D3-def	423	1452	1029			16	25
					lyr-CHS-D3-seq	215	1452	1237			14	25
lyr-CHS-D22.fas					lyr-CHS-D22-def	215	1452	1237			14	26
					lyr-CHS-D22-seq	162	1452	1290			14	25
all-scADH-D1.fas	1626		70.05	24	all-scADH-D1-def	23	230	207	8	25	16	28
					all-scADH-D1-seq	527	806	279	5	21	14	30
all-scADH-D2.fas					all-scADH-D2-def	565	755	190	5	25	16	29
					all-scADH-D2-seq	527	806	279	5	22	14	29
all-scADH-D3.fas					all-scADH-D3-def	454	659	205	7	28	14	26
					all-scADH-D3-seq	527	806	279	5	22	12	29
all-scADH-D22.fas					all-scADH-D22-def	565	755	190	5	25	14	30
					all-scADH-D22-seq	565	818	253	4	21	14	30
are-scADH-D1.fas	1626	50	36.6	16	are-scADH-D1-def	1044	1467	423	8	32		
					are-scADH-D1-seq	1044	1626	582	8	28		
are-scADH-D2.fas			34.3	16	are-scADH-D2-def	1129	1626	497	10	38		
					are-scADH-D2-seq	849	1463	614	8	25		
are-scADH-D3.fas					are-scADH-D3-def	1129	1626	497	10	39		
					are-scADH-D3-seq	1010	1606	596	8	22		
are-scADH-D22.fas					are-scADH-D22-def	1129	1626	497	10	38		
					are-scADH-D22-seq	1010	1463	453	8	26		
lyr-scADH-D1.fas	1626	48	37.7	16	lyr-scADH-D1-def	353	704	351			14	28
					lyr-scADH-D1-seq	565	1010	445			11	24

lyr-scADH-D2.fas				lyr-scADH-D2-def	353	755	402		16	28	
				lyr-scADH-D2seq	565	1028	463		11	22	
lyr-scADH-D3.fas				lyr-scADH-D3-def	565	1010	445		11	28	
				lyr-scADH-D3-seq	556	1109	553		9	27	
lyr-scADH-D22.fas		33.5	16	lyr-scADH-D22-def	565	1010	445		11	29	
				lyr-scADH-D22-seq	556	1109	553		9	27	
2x-CHG.fas				2x-CHS-D1-def	445	953	508	7	16		
				2x-CHS-D1-seq	395	953	558	5	14		
2x-scADH.fas	1626			2x-scADH-D1-def	1130	1455	325	9	14		
				2x-scADH-D1-seq	181	1156	975	0	11		too little a2

1. Hudson RR: **Estimating the recombination parameter of a finite population model without selection**. *Genetical Research* 1987, **50**:245-250.
2. Hudson RR, Kaplan NL: **Statistical properties of the number of recombination events in the history of a sample of DNA sequences**. *Genetics* 1985, **111**:147-164.

Additional file 3: Table S3. Isolation with migration analyses: mutation rates per fragment per generation. The rates are based on a per site mutation rate of $1.075 \cdot 10^{-8}$ scaled by the length of each fragment following Slotte et al. [1].

Region	Mutation rate
CHS	$2.13 \cdot 10^{-5}$
*sc*ADH	$1.55 \cdot 10^{-5}$
*trn*L-F	$5.30 \cdot 10^{-6}$
Mean	$1.21 \cdot 10^{-5}$

1. Slotte T, Huang H, Lascoux M, Ceplitis A: **Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae)**. *Molecular Biology and Evolution* 2008, **25**(7):1472-1481.

Additional file 4: Table S4. Signs of introgression among tetraploid *Arabidopsis arenosa* and *A. lyrata*. are-like and lyr-like give number of alleles clustering with *A. arenosa* and *A. lyrata*, respectively, for each of the two low-copy nuclear regions (CHS and *sc*ADH).

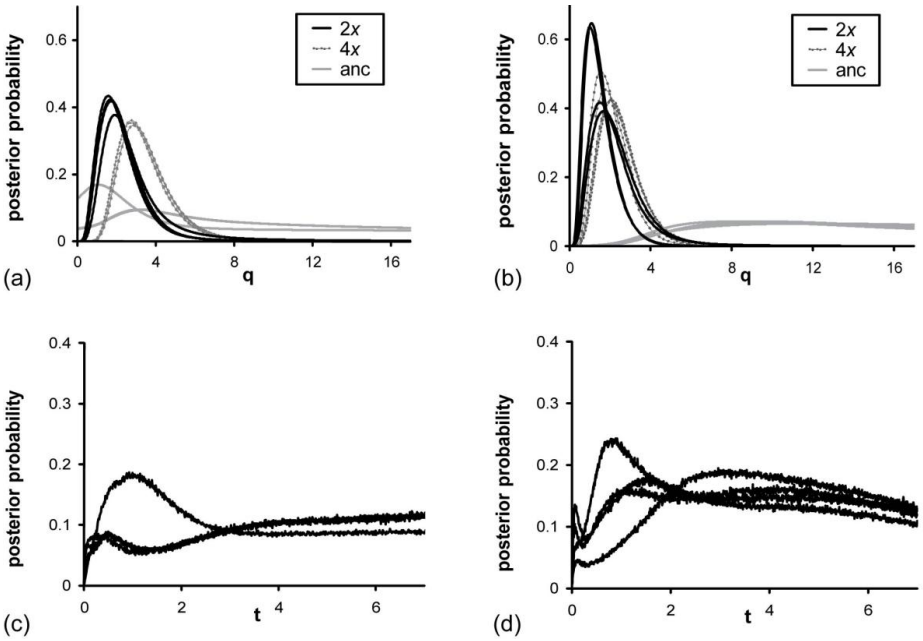
Taxon	Specimen	CHS are-like	CHS lyr-like	*sc*ADH are-like	*sc*ADH lyr-like
A. arenosa	a4_AUT1_1	3	3	3	-
	a4_AUT1_2	2	1	2	1
	a4_AUT2_17	2	2	3	-
	a4_AUT2_18	3	1	2	-
	a4_AUT4_14	3	1	3	-
	a4_AUT4_15	3	1	3	-
	a4_GER_1	1	1	2	-
A. lyrata	l4_AUT1_11	1	2	1	1
	l4_AUT1_12	3	1	-	2
	l4_AUT4_2	1	3	1	1

Additional file 5: Table S5. Diversity indices calculated for the different regions and the different data subsets: *are* 2*x* and *lyr* 2*x* are diploid *arenosa* and *lyrata*, respectively, *are* 4*x* D1-D22 and *lyr* 4*x* D1-D22 are tetraploid subsets created through different sequence duplication schemes, and *are* 4*x* and *lyr* 4*x* are the means of these subsets (CHS and *sc*ADH), or the unduplicated tetraploid subsets (*trn*L-F). # seq gives the number of sequences, Lenght gives sequence lengths excluding gap sites, S gives number of segregating sites, # hapl gives number of haplotypes, Hd ± sd gives gene diversity ± standard deviation, π gives nucleotide diversity, k gives avarage number of nucleotide differences.

		are 2x	are 4x D1	are 4x D2	are 4x D3	are 4x D22	are 4x	lyr 2x	lyr 4x D1	lyr 4x D2	lyr 4x D3	lyr 4x D22	lyr 4x
CHS	# seq	12	40	40	40	40	40	16	32	32	32	32	32
	Length	1417	1422	1422	1422	1422	1422	1421	1422	1422	1422	1422	1422
	S	43	76	76	76	76	76	25	56	56	56	56	56
	# hapl	11	28	28	28	28	28	10	17	17	17	17	17
	Hd ± sd	0.99 ± 0.04	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.98	0.93 ± 0.04	0.94 ± 0.03	0.94 ± 0.03	0.96 ± 0.02	0.96 ± 0.02	0.95
	π	0.01	0.013	0.013	0.014	0.013	0.013	0.01	0.01	0.01	0.01	0.01	0.01
	k	13.61	19.42	19.08	19.23	18.92	19.16	11.3	13.3	13.75	13.66	13.1	13.45
*sc*ADH	# seq	10	40	40	40	40	40	16	32	32	32	32	32
	Length	1133	1035	1035	1035	1035	1035	1305	1188	1188	1188	1188	1188
	S	104	144	144	144	144	144	66	94	94	94	94	94
	# hapl	9	24	24	24	24	24	9	17	17	17	17	17
	Hd ± sd	0.98 ± 0.05	0.97 ± 0.01	0.97 ± 0.01	0.97 ± 0.01	0.98 ± 0.01	0.97	0.94 ± 0.03	0.96 ± 0.02	0.95 ± 0.02	0.95 ± 0.02	0.96 ± 0.01	0.96
	π	0.027	0.024	0.024	0.025	0.026	0.025	0.019	0.023	0.022	0.022	0.022	0.022
	k	30.733	24.52	24.86	26.2	27.06	25.66	24.57	27.51	26.01	26.16	26.47	26.54
*trn*L-F	# seq	6					10	8					8
	Length	688					688	689					688
	S	2					3	4					3
	# hapl	3					3	3					3

Hd ± sd	0.6		0.62	0.71	0.71
	± 0.21		± 0.14	± 0.12	± 0.12
π	0.001		0.001	0.002	0.002
k	0.867		1.07	1.71	1.29

Additional file 6: Figure S1. Isolation with migration analyses of the def datasets (Table S2). (a,b) effective population sizes ($4N\mu$) for *A. arenosa* and *A. lyrata*, respectively. (b,d) time since population split for the same datasets.



Additional file 7: Figure S2. Isolation with migration analyses of the seq datasets (Table S2). (a, b) migration rates, (c, d) effective population sizes ($4N\mu$), and (e, f) time since population split for *A. arenosa* (a, c, e) and *A. lyrata* (b, d, f).