Phylogenetic relationships and diversification processes in *Allium* subgenus *Melanocrommyum*

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. Nat.)

Vorgelegt der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –
der Martin-Luther-Universität Halle-Wittenberg

von

Frau M. Sc Maia Gurushidze
geb. am 11.09.1976 in Kvareli, Georgien

Gutachter: 1. Prof. Dr. Martin Röser, MLU Halle-Wittenberg
2. Prof. Dr. Joachim W. Kadereit, Universität Mainz
3. Dr. Frank R. Blattner, IPK Gatersleben

Halle (Saale), 19.05.2009

Nothing in biology makes sense except in the light of evolution

T. Dobzhansky

Everything makes a lot more sense in the light of phylogeny

J. C. Avise

Contents

1. Introduction

1.1 Why phylogeny matters

Phylogenetic analyses of DNA or protein sequences have become an important tool for studying the evolutionary history of organisms from bacteria to humans. Since the rate of sequence evolution varies extensively with gene or DNA sequences one can study the evolutionary relationships at virtually all hierarchical levels (e.g., kingdoms, phyla, families, genera, species, and intraspecific classifications) by using different genes or DNA segments (Nei and Kumar 2000).

The important realization that comparisons of species or gene sequences in a phylogenetic context can provide meaningful insights into different fields of plant biology, including ecology, molecular biology, and physiology, has been illustrated in several reviews (Soltis and Soltis 2000, 2003; Hall et al. 2002; Doyle et al. 2003). Studies of the rate of evolutionary change as well as estimation of the age of diversification pattern of lineages depend directly on knowledge of phylogenetic relationships (Magallon and Sanderson 2001). The phylogenetic analyses of angiosperms have resulted in their reclassification (APG 1998, APGII 2003). This ordinal-level reclassification is perhaps the most dramatic and important change in higher-level angiosperm taxonomy in the past 200 years (Soltis and Soltis 2000). The more a classification reflects evolutionary history (phylogeny) the more useful it will be to point to taxa to look for important features. For example, taxol is a compound known only from the yew family (Taxaceae). If one wishes to look for additional sources of taxol and other taxanes, rather than randomly sample the plant kingdom, one can focus on the nearest relatives as understood from a phylogenetic tree of the seed plants. This in turn leads to investigate the Podocarpaceae because the phylogenetic analyses indicated that this is the sister group to the yew family. Subsequently, taxanes have been reported also in the Podocarpaceae (Stahlhut et al. 1999).

Indeed, phylogenies guide the search for plants of potential commercial importance. Revealing the closest relatives of crops could direct comparative studies and subsequent applications in plant breeding. For instance, *Hordeum bulbosum*, the sister species of cultivated barley (*H. vulgare*) has been widely used in breeding, namely in producing the completely homozygous double-haplotype lines of barley due to the fact that *H. bulbosum* chromosomes are completely eliminated from the *H. vulgare* × *H. bulbosum* hybrid (von Bothmer et. al. 1995). Besides, knowing the clear picture of phylogenetic relationships of the cultivated species can point to taxa that may serve as sources of novel traits for crops.

The value of placing “model organisms” in the appropriate phylogenetic context is another aspect where phylogenies could help answering fundamental evolutionary questions. The most obvious example concerns the model plant *Arabidopsis thaliana* and its relatives, where phylogenetic studies revealed that the closest relatives of *A. thaliana* are the representatives of former genus *Cardaminopsis* (Koch et al. 2001; Clauss and Koch 2006). The accumulated knowledge in the function and evolution of genes that control responses to draught, pathogens, insects and other environmental challenges, as well as

knowledge about the genes involved in the evolution of breeding system can be transferred from the model organism to the closest relatives to study these long-standing evolutionary questions (Mitchell-Olds 2001; Mitchell-Olds et al. 2005)**.**

In addition, a phylogenetic framework has revealed the patterns of evolution of many morphological and chemical characters, including complex pathways such as nitrogen-fixing symbioses, chemical defense mechanisms and mustard oil production (for review see Soltis and Soltis 2000; Daly et al. 2001). For example, mustard oil glucosinolates are known to be produced in at least 15 different plant families. Before the advances in molecular phylogenetics, this biochemical pathway was thought to have arisen and been lost on multiple occasions. Molecular phylogenetic information, however, indicates that all mustard oil producing families with the exception of the genus *Drypetes* (a member of Malpighiales) are part of the same ordinal clade Brassicales, indicating only two evolutionary origins for mustard glucosinolates (Rodman et al. 1993). Moreover, these mustard oils are derived from two different biosynthetic pathways, thus showing that even though the final product may be the same, the ways they were synthesized are not; hence they are not homologous in an evolutionary sense (Daly et al. 2001).

1.2 Plant species-level systematics

As noted above, plant systematics at the generic and higher levels has advanced greatly in the last decades (Chase et al. 1993; Savolainen et al. 2000; Soltis et al. 2000), and the phylogenetic tree of major angiosperm groups is perhaps "better known than any other equivalent group of organisms" (Barraclough and Reeves 2005). These phylogenies have been used to investigate patterns and processes causing diversification in angiosperms (Magallon and Sanderson 2001; Dodd et al. 1999; Barraclough and Savolainen 2001). However, plant species-level systematics has lagged somewhat behind. Now that most family-level angiosperm phylogenetic patterns have been explored, and generic-level studies are underway, it is very challenging to focus on the species level. After all, it is at this level where one can observe the actual processes driving evolution. Species level phylogenies contribute to the better understanding of the mechanisms involved in speciation (be it geographic or ecological), extinction, and evolutionary radiations. As adaptive radiations are characterized by clear ecological and/or morphological diversification coupled with high species production, a phylogenetic framework with complete species representation is one of the best tools (in combination with ecology of the taxa and geological history) in order to distinguish between non-adaptive and adaptive radiations and to understand the mechanisms underlying them. The best known example of utility of phylogenies in studying adaptive radiation is probably the study of Baldwin and Sanderson (1998) in Hawaiian silversword alliance. They demonstrated that the exceptional diversification of the silversword alliance has occurred within the time span encompassed by the emerging modern high islands of the Hawaiian archipelago. Likewise, the shaping of biogeographic distributions, the filling of ecological space, all can be studied most

effectively from species-level perspectives (Zhang et al. 2001, 2004; Linder and Hardy 2005).

1.3 Phylogenetic inference

There are many statistical methods that can be used for reconstructing phylogenetic trees from molecular data. Commonly used methods are classified into three major groups: 1) distance methods, 2) parsimony methods, 3) model-based methods (maximum likelyhood and Bayesian). In the next sections I will give a brief overview of these phylogenetic methods. All of these methods have strengths and weaknesses (Swofford et al. 1996; Lewis 1998; Doyle and Gaut 2000; Huelsenbeck et al. 2002; Nei and Kumar 2000), some of which are summarized in Table 1.

Distance methods

In distance methods evolutionary distances are computed for all pairs of taxa, and phylogenetic trees are constructed by clustering the taxa according to these distance values. Such methods, which are based on overall similarity and include various forms of clustering and ordination, are sometimes called "phenetic" methods. Phenetic methods cluster and classify species based upon the number of identical characters that they share, that means, based upon overall similarity (Sneath and Sokal 1973).

Probably the most often used distance method is the neighbor joining (NJ) method, which is based on the minimum-evolution criterion for phylogenetic trees, i.e. the topology that gives the least total branch length is preferred at each step of the algorithm (Nei and Kumar 2000). Neighbor joining is an iterative algorithm, at each iteration it finds the pair of taxa with the lowest distance value, then creates a node on the tree that joins these two taxa (i.e. join the closest neighbors, as the algorithm name implies). Considering the pair of joined neighbors as a single taxon, the algorithm searches for the next closest neighbor based on distance values. Finally the topology of the tree is defined by successfully joining neighbors and producing new pairs of neighbors (Saitou and Nei 1987). The distance measures used in NJ and other distance methods are typically based on models of nucleotide substitution. The NJ algorithm is fast, appropriate for analyzing a large data set, and is capable of conducting bootstrap tests rapidly. However, it also has important weaknesses. For example, NJ provides only a single tree, precluding comparison with other topologies. Furthermore, because sequence differences are summarized as distance values, it is impossible to identify the specific character changes that support a branch. Taking into account all the above mentioned, most authors argue for a pluralistic approach – that different phylogenetic methods should be used in addition to NJ (e.g. Swofford et al. 1996; Doyle and Gaut 2000; Nei and Kumar 2000).

Maximum parsimony methods

Maximum parsimony (MP) was originally developed for morphological characters and is often classified as a "cladistic" method. The critical feature in cladistics is the identification of derived shared traits, called synapomorphic traits. A synapomorphic character is shared by some taxa but not others because the former inherited it from a common ancestor (Hennig 1953, 1966). The presence of a derived character state in two taxa suggests that they share a common ancestor in which the apomorphy first evolved. Then joining these two taxa as sisters represents the simplest hypothesis that explains the pattern of derived character states, following the principle of parsimony. In general, the theoretical basis of parsimony is Ockham's philosophical idea that the best hypothesis to explain a process is the simplest one and "*ad hoc* hypotheses should be avoided whenever possible" (Swofford et al. 1996). As a consequence, the parsimony methods for inferring phylogenies operate by selecting trees that minimize the total tree length i.e. the number of evolutionary steps (transformations from one character state to another) required to explain a given set of data. For example, the topology of trees which requires the smallest number of nucleotide or amino acid substitutions explaining the entire evolutionary process will be chosen as the best tree(s) (Swofford et al. 1996). As the number of best trees (shortest trees) could be quite high, it is mostly informative to compute a strict consensus tree, a tree which contains only groups or nodes that are represented in all shortest trees.

If there is no homoplasy and the number of nucleotides examined is very large, MP methods are expected to produce the correct (realized) tree. However, in practice nucleotide sequences are often subject to parallel mutations and reversals (homoplasy), and the number of examined nucleotides is rather small. In this case MP methods tend to give incorrect topologies (Nei and Kumar 2000). Furthermore, Felsenstein (1978) has shown that when the rate of nucleotide substitution varies extensively among evolutionary lineages, MP methods may generate incorrect topologies erroneously grouping two or more long branches together as monophyletic. This phenomenon is called long-branch attraction (LBA; Hendy and Penny 1989), and is particularly acute with DNA sequence data, for which each character has only four possible states (as there are four nucleotides), and mutation rates can vary widely. Thus, when DNA substitution rates are high, the probability that two lineages will convergently evolve the same nucleotide at the same site increases. When this happens, parsimony misleadingly interprets this similarity as a synapomorphy, and joins long branches. However, LBA is basically a sampling problem and may be alleviated including taxa that break up the terminating long branches (Felsenstein 2004; for a review see Bergsten 2005).

Maximum likelihood and Bayesian methods

Both maximum likelihood (ML) and Bayesian inference (BI) methods are known as model-based methods because they incorporate ideas about the probability of change (in molecular phylogenetics the different models of sequence evolution). Given the model and the possible tree, the maximum likelihood method calculates the likelihood of observing the

data. The best estimate of the phylogeny is the tree that maximizes this likelihood. In other words, it calculates the probability that a proposed model and hypothesized history (phylogenetic tree) would give rise to the observed data (Swofford et al. 1996). Bayesian inference (BI) methods are closely related to likelihood methods, differing only by maximizing the probability of observing a particular tree, given the model and the data. Unlike the ML method, BI provides and calculates the probabilities of a set of trees so that they can be compared (Huelsenbeck et al. 2001). The major criticism of ML method is that it is computationally very intensive and much slower than other methods. Whereas ML is not feasible for large data sets, BI (as implemented in MrBayes; see Huelsenbeck et al. 2001) incorporates a faster search strategy (using Markov chains) and can be used on data sets of several hundred taxa to find trees, branch lengths, and support. MrBayes uses a simulation technique called Markov chain Monte Carlo (or MCMC) to approximate the posterior probabilities of trees. MCMC methods in BI draw a random sample from the posterior distribution of hypotheses (in this case, trees). It thus becomes possible to make probability statements about true trees (Felsenstein 2004). The basic MCMC algorithm involves two steps: (i) a new tree is proposed by stochastically perturbing the current tree, (ii) this tree is then either accepted or rejected using an acceptance probability described by Metropolis et al. (1953) and Hastings (1970), which actually is comparing likelihood and probability of a new state (in this case tree) to that of the old state. If the new tree is accepted, then it is subject to further perturbation. This process of proposing and accepting/rejecting many states is repeated many thousands or millions of times, and the proportion of the time that any tree is visited is a valid approximation of the posterior probability of that tree (Tierney 1994). Thus, Bayesian analysis samples trees according to their posterior probabilities. Once such a sample is available, it is possible to look for the features that are common among the trees, and construct a consensus tree with the posterior probabilities (pp) of the individual clades. The main criticism of BI is that posterior probabilities can be overestimates (Suzuki et al. 2002).

Bootstrap confidence for phylogenetic trees

The bootstrap is a computer-based technique for assessing the accuracy of almost any statistical estimate (Efron and Tibshirani 1993). It is a resampling method, which was introduced by Felsenstein (1985) to phylogenetics for statistical evaluation of phylogenetic trees. The bootstrapping procedure produces replicate "bootstrap data sets" from randomly sampled characters of the data-matrix. Each of these randomly generated data sets is then analyzed phylogenetically, with a consensus tree constructed to summarize the results of all replicates.

Table 1. Comparison of methods of phylogeny reconstruction (modified from Soltis and Soltis 2003).

Method	General	Advantages	Disadvantages
Parsimony	Simplest explanation is the best (Ockham's razor)	By minimizing no. of steps, it also minimizes the no. of additional hypothesis (parallel or reversal nucleotide substitutions)	Different results may be obtained based on the entry order of sequences (therefore, perform multiple searches)
	Select the tree or trees that minimize the amount of change (no. of steps)	Searches identify numerous equally parsimonious (shortest) trees; treats multiple hits as an inevitable source of false similarity (homoplasy)	Relatively slow (compared with NJ) with large data sets
		Basic method can be modified by weighting schemes to compensate for multiple hits	Highly unequal rates of base substitution may cause difficulties (e.g. long branch attraction)
		Can identify individual characters that are informative or problematic	
		Can infer ancestral states	
NJ	Involves estimation of pair-wise distances between nucleotide sequences	Fast	Different results may be obtained based on the entry order of sequences
	Pair-wise distances compensate for multiple hits by transforming observed percent differences into an estimate of the no. of nucleotide substitutions using one of several models of molecular evolution	Provides branch lengths	Only a single tree produced; cannot evaluate other trees
	Minimum evolution is a common distance criterion for selecting an optional tree, i.e. sum of all branch lengths is the smallest	Uses molecular evolution model	Branch lengths presented as distances rather than as discrete characters (steps)
	NJ algorithm provides a good approximation of the minimum evolution tree	Readily implemented in PAUP*, MEGA and several other programs	Cannot identify characters that are either informative or problematic
			Cannot infer ancestral character states

Maximum Likelihood	Involves estimating the likelihood of observing a set of aligned sequences given a model of nucleotide substitution and a tree	A statistical test (the likelihood ratio test) can be used to evaluate properties of trees	Computationally very intensive (much slower than other methods). However, recently introduced fast searching algorithms as implemented in a software RAxML are relatively fast
		Nucleotide substitution models are used directly in the estimation process	Practical with smaller data sets. However, as mentioned above, the newly implemented searching strategies (program RAxML) allow analyzing relatively large data matrices.
		Flexible, different models can incorporate parameters of base frequencies, substitution rates, and variation in substitution rates.	
		Easily implemented in PAUP* and PHYLIP	
		Uses all of the data (invariable sites and unique mutations are still informative, unlike parsimony analysis)	
Bayesian	Uses a likelihood function and an efficient search strategy	Based on the likelihood function, from which it inherits many of its favorable statistical properties	Very high memory demands
	Based on a quality called the posterior probability of a tree	Uses models as in ML	
	Researcher may specify belief in a prior hypothesis prior to analysis	Can be used to analyze relatively large data sets	
		Readily implemented in MRBAYES	
		Provides support values	Posterior probabilities (measure of internal support) can be overestimates

This widely used technique provides assessments of "confidence" for each clade of an observed tree, based on the proportion of bootstrap trees showing that same clade. The proportion of trees/replicates in which a grouping is recovered is presented as a measure of bootstrap support (bs) for that group. However, interpretations of bootstrap values vary (for review, see Soltis and Soltis, 2003b), some consider values of 70% or more as indicators of strong support, whereas others reserve "strong support" for values of 90% or 95% and

above. Zharkikh and Li (1992; Li and Zharkikh 1994) examined the statistical properties of bootstrap analyses and showed that the bootstrap supports were underestimates. Hillis and Bull (1993) carried out a large simulated study that reached the same conclusion. They argued that the value of 70% might indicate a significantly supported group. Many further studies showed that from a practical point of view, bootstrap values are more conservative and might fail to support a true node, whereas posterior probability values are more liberal and may fail to reject a false node (Archibald et al. 2003). This observation is consistent with other empirical results (e.g., Huelsenbeck et al. 2002) and reflects the observation that these values are measurements of different statistical properties of phylogenetic trees (e.g., Alfaro et al. 2003), and therefore it is always desirable to have both values (bootstrap values and posterior probabilities) for testing the significance of a clade resulting from phylogenetic analyses.

Networks – genealogical methods

The phylogenetic methods listed above are constructing bifurcating trees, because they assume that DNA sequence splits into two descendant sequences either at the time of speciation or gene duplication. Therefore, phylogenetic trees are usually bifurcating. If one is analyzing relationships among sequences of different species, where relationships are hierarchical, the tree-building methods could be best choice, particularly if the sequence variation is optimal for the level of relationships studied (family, generic, etc.). In contrast, relationships among sequences sampled from individuals of the same species are not hierarchical, because they are the result of sexual reproduction and recombination (tokogenetic relationships). Therefore, several major assumptions of the phylogenetic methods are violated at the population level. These violations are (i) existence of ancestral sequence types (haplotypes) together with their descendant sequences in the populations, (ii) the fact related to the persistence of ancestral haplotypes is that a single ancestral sequence will often give rise to several descendant alleles, yielding a haplotype tree with true multifurcations (instead of bifurcating tree), and (iii) recombination between genes and hybridization between the lineages, which results in reticulate relationships. The above mentioned phenomena lead to poor resolution of the resulting trees or inadequate portrayal of intraspecific (genealogical) relationships. It became obvious that phylogenetic methods that allow for persistent ancestral nodes, multifurcations and reticulations are needed to take these population phenomena into account. As a solution, network methods are developed for estimating within species relationships (Posada and Crandall 2001). In addition, networks provide a way of representing information present in a data set (e.g. presence of loops could point to recombination, or indicate reverse or parallel mutations), which are not visualized by bifurcating trees. In this thesis I use a statistical parsimony network approach (Templeton et al. 1992), which connects the haplotypes in a single network according to the numbers of mutational steps among them until all haplotypes are included or until parsimony connection limit is reached. Parsimony limit is the maximum number of differences among haplotypes as a result of single substitutions (i.e. those that are not the

result of multiple substitutions at a single site) with a 95% statistical confidence (Templeton et al. 1992), and thus depends mainly on the length of haplotype sequences to be analyzed.

1.4 Molecular clock

The idea of a molecular clock was developed by Zuckerkandl and Pauling (1962), and later mathematically formalized by Kimura (1968; Kimura and Ohta 1971). A global molecular clock implies that amino acid and DNA sequences evolve at a constant evolutionary rate over the whole phylogeny. The molecular clock indeed received much attention, as a full understanding of the patterns and processes of biological diversification requires dating of evolutionary events.

Molecular clock-based dating of divergence events (i.e. nodes in a phylogeny) involves the following steps: (i) obtaining the genetic distance between two sequences or taxa in the analysis, one of which must have a known age, usually determined from fossils; (ii) calculating the substitution rate by dividing the genetic distance by the known age; (iii) using that rate to convert genetic distances between taxa of interest into estimates of their absolute ages. However, as mentioned above, the utility of a molecular clock depends on the rate constancy of molecular evolution at a genetic locus across time and across lineages. Unfortunately, the strict clock turned out to be an oversimplified model, and it became clear that rate variation rather than rate constancy is a common phenomenon (Britten 1986; Li 1997; Arbogast et al. 2002). Particularly studies in plants have uncovered striking heterogeneity in the rate of nucleotide substitution among lineages (Klak et al. 2004). To overcome the problem of rate heterogeneity, methods that allow for evolutionary rate to change over time, i.e. relax molecular clock assumption, were developed. Here I briefly review the methods I used for dating of diversification events.

Nonparametric rate smoothing (NPRS)

The basic assumption in NPRS (Sanderson 1997) is autocorrelation (Gillespie 1991), which implies that the rate of a daughter lineage is similar, but maybe not the same, as the rate of the mother lineage, i.e. relies on minimization of ancestor-descendant local rate changes. The method therefore minimizes rate differences between the mother and daughter lineages; the optimization is made by penalizing large rate variations, according to a squared function of rate changes between adjacent branches. The NPRS algorithm calculates different rates over the whole tree at the same time. Multiple fossil age constraints can be used to calibrate the tree.

Penalized likelihood (PL)

Penalized likelihood (PL) is a semi-parametric method (Sanderson 2002), which combines a model-based likelihood with a roughness penalty regulated by a smoothing parameter. Substitutions along a branch are assumed to have a Poisson distribution. PL uses a saturated model for rate change that allows rates to vary freely over the tree. This model

actually means that there are more parameters than observations, and therefore a penalty for sudden rate changes is needed. The smoothing parameter is objectively chosen by a cross validation process, which sequentially removes data to find the smoothing that best fits the data. If the smoothing parameter is large, the function is dominated by the roughness penalty, and this leads to a clock-like model. If it is low, the smoothing will be effectively unconstrained, and the method will then behave similarly to NPRS.

1.5 Study group – *Allium* subgenus *Melanocrommyum*

Taxonomic history and systematics

The onion genus *Allium* L. is probably the largest genus of the petaloid monocotyledons with about 800 species (Friesen et al. 2006). *Allium* is a member of family Alliaceae subfamily Allioideae Herb. (Fay and Chase 1996). According to Fay and Chase (1996) and Friesen et al. (2000) subfamily Allioideae consists only of the genus *Allium*. The genus is naturally distributed in the northern hemisphere, with the main center of diversity in southwest and central Asia and a second smaller one in North America. *Allium* includes some economically important species like common onion, garlic, chives and leek, and also some species with medicinal properties and others with horticultural merit (Fritsch and Friesen 2002). Most *Allium* species produce cysteine sulphoxides, causing the characteristic odor and taste. The content and chemistry of these compounds vary tremendously among different representatives of the genus, and show correlation with three major phylogenetic lineages within the genus (Fritsch and Keusgen 2006).

No comprehensive monograph of the genus has been compiled since Regel (1875) and the taxonomy is complicated, with a proliferation of synonyms, as well as the infrageneric groups. The history of infrageneric classification of the genus *Allium* traces back to Linneaus (Linné 1753) who accepted 30 species in three alliances. Later studies (Don 1832; Vvedensky 1935; Traub 1968; Kamelin 1973; Hanelt et al. 1992) recognized an increased number of infrageneric groups and species. Recent molecular phylogenetic studies radically changed our understanding of the evolution of the genus. Phylogenetic analyses of *Allium* based on the nuclear rDNA internal transcribed spacer region (ITS) proved several subgenera and sections to be non-monophyletic, allowed to circumscribe intrageneric groups and resulted in a revision and new classification of the genus (Friesen et al. 2006). According to this latest classification of the genus *Allium*, about 800 species are assigned to 15 subgenera based on the sequences of nuclear rDNA ITS**,** karyological and morphological characters. Among these subgenera subgenus *Melanocrommyum* is one of the largest and taxonomically most complicated, comprising approximately 140 species according to the last taxonomic treatment (Khassanov and Fritsch 1994). Due to their large and often dense inflorescences and peculiar flower colors, some members of this subgenus are cultivated as ornamentals like *A. cristophii, A. karataviense, A. macleanii, A. stipitatum,* and *A. hollandicum*, including hybrids derived from them (Friesen et al. 1997).

Well-developed leaf sheaths restricted to subterranean parts, the extremely short developmental period, and several anatomical characters, such as two opposite rows of

vascular bundles in leaf blades and true palisade parenchyma, are considered synapomorphies for the subgenus (Fritsch 1992a; Hanelt et al. 1992).

The first species of this taxonomic group, *A. magicum* (a name now rejected in favour of *A. nigrum*, nom. cons), has already been recognized by Linnaeus (1753) in the first edition of "Species Plantarum" under the informal group "foliis caulinis planis". Only a century later, this species was separated as section *Melanocrommyum* by Webb and Berthelot (1848). Another species, *A. caspium* (Pallas) M. Bieb., was separated under another section (section *Kaloprason* C. Koch) by Koch (1849). However, the last monographer of the genus, Regel (1875), did not accept these sections and included the above mentioned species and related ones under section *Molium* G. Don ex Koch.

Many botanists followed this sectional affiliation for the next decades. Only in the middle of the 20th century it became evident that section *Molium* in this wide sense consisted of two only distantly related subgroups differing by chromosomal basic number (x = 7 versus x = 8) correlated with far-reaching differences in the anatomical structure of vegetative and generative organs. As a result, Wendelbo (1969) raised section *Melanocrommyum* to subgeneric rank and separated some morphologically well circumscribed groups of species at sectional level. Since then a characteristic feature of the taxonomic history of the subgenus is an increasing number of sections and subsections. Besides, the description of new species and subspecies of subgenus *Melanocrommyum* continued (Khassanov and Fritsch 1994; Kamelin and Seisums 1996; Fritsch et al. 1998, 2002), and the different opinions where best to affiliate new taxa remained an unsolved problem.

Taxonomically important characters for sectional delimitation within the subgenus are density, shape and developmental mode of the inflorescences, shape and posture of the tepals, length of the scape, and shape and texture of the leaf blades. All existing classifications were based on these morphological characters. However, a formal cladistic analysis of morphological characters was never conducted. Unfortunately, many of these features are quantitative rather than qualitative, difficult to discern, or subjective and open to interpretation, and therefore are likely responsible for the extremely different taxonomic treatments of subgenus *Melanocrommyum*.

The classifications proposed by Kamelin (1973), Tscholokaschwili (1975), Khassanov (1992), Fritsch (1992a, 1993), Seisums (1994), and Khassanov and Fritsch (1994) differ considerably accepting different numbers of sections and subsections, and some even maintained the traditional use of section *Molium* s. lat. for subgenus *Melanocrommyum* (Tzagolova 1977, 1983). This clearly reflects uncertainties about relationships within the subgenus. The most recent classification affiliated about 140 species and subspecies to 15 sections and 20 subsections (Khassanov and Fritsch 1994).

A. cardiostemon

A. aflatunense

A. nigrum

A. stipitatum

A. karataviense

A. isakulii subsp. *nuratense*

Geographical distribution and ecology

Subgenus *Melanocrommyum* has a Eurasian distribution (some species also reported from northern part of Africa) extending from the western Mediterranean to the Near and Middle East, reaching north-western China and Pakistan in the east, and southern Siberia in the north (Fig. 1). The center of diversity is in the mountainous regions of Central and Southwest Asia (Hanelt et al. 1992; Khassanov and Fritsch 1994). Although the majority of *Allium* species have adapted to different habitats (different types of forests, subalpine pastures, and moist subalpine and alpine grasslands), the species of subgenus *Melanocrommyum* are exclusively found in arid habitats. Dry steppes, semi-deserts, or even sandy and gravellous deserts, stony slopes of arid mountains, more rarely shrublands and open dry forests are the preferred habitats of *Melanocrommyum* species (Hanelt et al. 1992). The severe climatic conditions enforced an ephemeroid growth rhythm (less pronounced in other subgenera of *Allium*) which is characterized by long summer-winter dormancy and an extremely shortened life cycle from spring to early summer (Hanelt et al. 1992). Based on synapomorphies and specializations as well as due to distinct ecological and geographical restrictions, the subgenus was hypothesized to be a phylogenetically derived group, which rapidly diversified during the late Tertiary when by the closing of the Tethys Sea enormous regions could be newly colonized (Hanelt 1992).

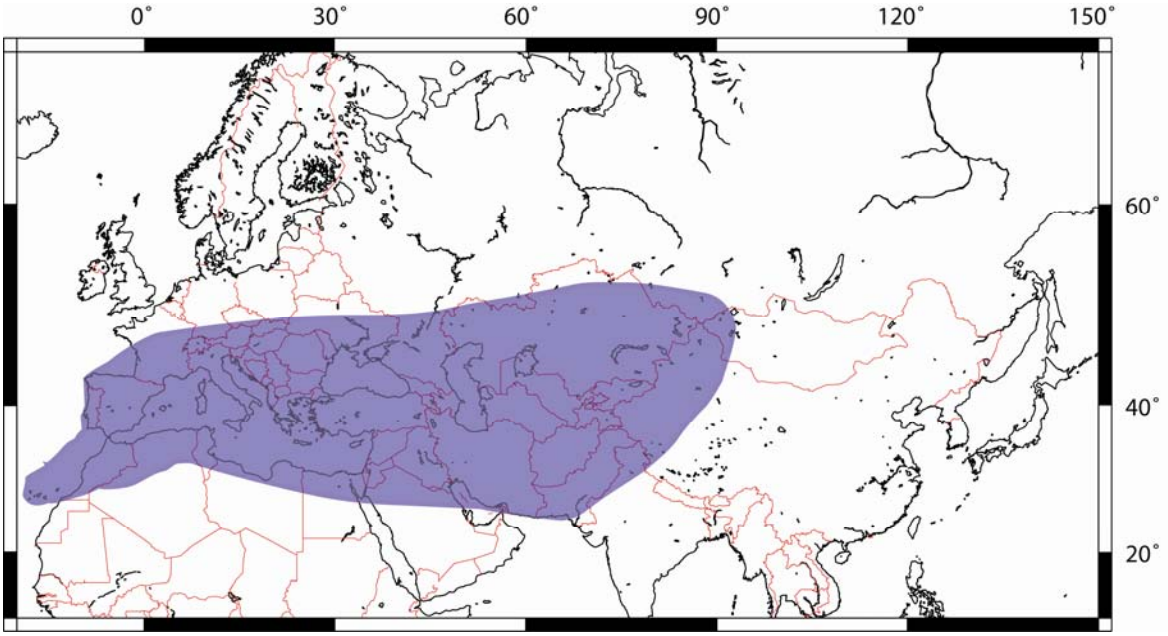


Fig. 1. Distribution map of *Allium* subgenus *Melanocrommyum*.

Phylogenetic relationships within the subgenus

Previous works. A phylogenetic analysis of some members of the subgenus based on PCR-RFLP of noncoding regions of chloroplast DNA was carried out by Mes et al. (1999). They found severe incongruence between morphology and chloroplast data, though the resolution in their phylogenetic trees was low and the study did not cover all morphological and geographical groups of the subgenus. They also reported the existence of polymorphic internal transcribed spacer regions (ITS) of the nuclear ribosomal DNA within individuals based on RFLP analyses of amplified ITS fragments and hypothesized high level of interspecific gene flow within the subgenus. Another published phylogeny of the subgenus based on ITS sequences was restricted to only 13 species, each represented by a single individual (Dubouzet and Shinoda 1998). The results of the above mentioned studies are not comparable due to incomplete sampling (and therefore analyzing relationships among different species) and different markers and methods used. However, both studies report conflicts between taxonomical treatment and molecular data. Other molecular investigations focused on the infrageneric differentiation and classification of the entire genus *Allium* (Linne von Berg et al. 1996; Mes et al. 1997; Dubouzet and Shinoda 1999; Fritsch and Friesen 2002; Friesen et al. 2006), or the phylogenetic position of North American (mainly Californian) taxa and their adaptation to serpentine soils (Nguyen et al. 2008). In the absence of comprehensive molecular analyses, the phylogenetic relationships within subgenus *Melanocrommyum* remained essentially unclear.

Phylogeny based on nuclear and chloroplast data. In this thesis I report analyses of DNA sequences from both the chloroplast and nuclear genome. The nuclear genome is represented by sequences of the internal transcribed spacers (ITS) of nuclear rDNA. The ITS region has a great potential to resolve plant phylogenies at specific, generic and even family levels (Baldwin et al. 1995; Wissemann 2003; Hörandl et al. 2005). Favorable properties that allow widespread utilization of ITS are nearly universal primers useful for amplifying the region from most plant and fungal phyla (White et al. 1990; Blattner 1999). However, the sole usage of the ITS region for phylogenetic inference might also face some problems due to existence of paralogs (Álvarez and Wendel 2003; Blattner 2004), pseudogenes (Harpke and Peterson 2007), or possible uniparental inheritance due to unidirectional homogenization of rDNA repeats after hybridization (Álvarez and Wendel 2003). Therefore, cloning of ITS amplicons and additional molecular information from chloroplast data could become essential to detect paralogous copies within individuals, and reticulations caused by hybridization. Data from the chloroplast genome come from the leucine transfer RNA (*trn*L) gene and the intergenic spacer between *trn*L and *trn*F (*trn*L-*trn*F) (Taberlet 1991; Gilley and Taberlet 1994). Both, nuclear and plastid genomes were covered for two reasons: (1) to compare phylogenetic hypotheses from different genomes and (2) to help resolve more distal portions of the phylogeny. The expectation in starting this project was that overall the *trn*L- *trn*F region would have more slowly evolving sites. This information would be useful for resolving the deeper nodes in the phylogeny, whereas

ITS would show higher overall substitution rates that would aid in resolving the distal portions of the phylogeny (Soltis and Soltis 1998; Beardsley and Olmstead 2002). Phylogenetic analyses of nuclear and chloroplast sequences confirmed the monophyly of the subgenus, while most sections were either para- or polyphyletic (Chapters 2 and 3). The splits of the large sections are supported by anatomical characters (differences in the anatomy of flower nectaries). ITS data (i) demand a new treatment at sectional level, (ii) do not support the hypotheses of frequent gene flow among species, (iii) indicate that multiple rapid radiations occurred within different monophyletic groups of the subgenus, and (iv) detected separately evolving lineages (cryptic species) within three morphologically clearly defined species (Chapter 2). In chapter 4, I attempt to infer the reasons of the present biogeographic pattern and disjunct distribution of the genus *Allium* and a time frame of the diversifications of arid-adapted lineages (subgenus *Melanocrommyum*) using a molecular clock approach.

Karyology and ploidy level

The subgenus consists mostly of diploid (x = 8) perennial species. Karyological analyses revealed uniform karyotypes without any clear species-specific or section-specific characteristics (Fritsch and Astanova 1998). Despite the uniform karyotypes, genome size in *Melanocrommyum* is quite variable and ranges from 26 to 50 pg for 2C DNA content (Ohri et al. 1998; Gurushidze et al. in prep.). Several studies found the correlation of genome size with life history, breeding system and ecology within several taxonomic groups (Albach and Greilhuber 2004; Weiss-Schneeweiss 2006). As the members of the subgenus do not show differences in life form, karyology, ploidy level, or in habitat ecology, it becomes particularly challenging to address the evolution of genome size variation in a phylogenetic framework. In my thesis (chapter 5) I address the genome size variation and evolution in *Melanocrommyum* in a phylogenetic context, to test if there is phylogenetic constraint on genome size in this group.

2. Phylogenetic analysis of nuclear rDNA ITS sequences of the subgenus *Melanocrommyum* infers cryptic species and demands a new sectional classification

The content of this chapter is already published in *Molecular Phylogenetics and Evolution*, 2008. 49: 997 – 1007; doi: 10.1016/j.ympev.2008.09.003

2.1 Introduction

The internal transcribed spacer region of nuclear ribosomal DNA is one of the most often employed markers for phylogenetic analyses at the generic and subgeneric level (Baldwin et al. 1995; Álvarez and Wendel 2003; Wissemann 2003; Hörandl et al. 2005) and has shown to be informative also in *Allium* (Friesen et al. 2006; Gurushidze et al. 2007). In combination with chloroplast data or via cloning of ITS paralogs, the marker can also give insights into reticulate events caused by hybridization (Wendel and Doyle 1998). The presence of paralogous ITS copies caused by gene duplication or hybridization has been reported quite often (Álvarez and Wendel 2003; Blattner 2004), which together with concerted evolution of the ribosomal DNA clusters could complicate phylogenetic inference (Blattner 2004). Therefore, it is particularly important to distinguish homologous and non-homologous ITS sequences within individuals in *Melanocrommyum*, as a high level of hybridization was hypothesized by Mes et al. (1999).

In an attempt to clarify the phylogenetic relationships within the subgenus, I analyzed ITS sequences (directly sequenced and cloned amplicons) of multiple accessions of more than 100 species covering all existing taxonomic groups and the whole geographic distribution of subgenus *Melanocrommyum*. As shape and position of the septal nectaries and their excretory canals differ remarkably among several taxonomic groups of *Allium* (Fritsch 1992b), nectary anatomy was studied in subgenus *Melanocrommyum* to see if groups defined by ITS data are supported by morphological structures.

2.2 Materials and Methods

Plant material

First of all, as a subgenus and a section within the subgenus have the same name (*Melanocrommyum)*, I emphasize that under *Melanocrommyum* species (or taxa) subgenus *Melanocrommyum* is referred to, while I always mention "section *Melanocrommyum*" when referring to the section.

A total of 430 accessions were included in the study representing 106 species, covering all sections and almost all subsections of the subgenus, together with eight outgroup species representing subgenera *Reticulatobulbosa, Butomissa, Caloscordum, Anguinum, Vvedenskya,* and *Porphyroprason*. Outgroups were selected based on the phylogenetic study of the genus *Allium* (Friesen et al. 2006). In case when individuals from different populations were available, those from different geographical areas were selected to control for intraspecific sequence variation. Leaf material for DNA extraction was obtained from

the Taxonomic *Allium* Reference Collection of the IPK Gatersleben (Germany). Most accessions in the collection stem from natural populations collected during field trips and afterwards propagated vegetatively. Herbarium specimens were used when living plants were not available. When accessions occupied unexpected positions in initial phylogenetic analyses, additional accessions of the same species were included. Voucher specimens were deposited in the herbarium of the IPK Gatersleben (GAT). Several accessions recently collected in Iran and Tajikistan turned out to represent new species. Descriptions of these taxa are not formally published (in prep.), therefore they were numbered as: *A. sp. 1 … 8* followed by the proposed epithets in quotation marks. Accession numbers, sectional classification, and the origin of plant material are listed in the appendix Table 1.

Anatomy of septal nectaries

The ovaries of 96 species were collected and fixed at the beginning of full anthesis from the IPK living collection. The shape and position of the nectaries and their canals were studied using serial cross sections of ovaries stained with safranin and Astra® blau (Merck, Darmstadt, Germany). The nectaries were easily detectable by the dense granular cytoplasm of secretory and surrounding cells. The excretory canals possess epidermal cells with a cuticle surrounding the cavity. The major types of nectaries were defined by the shape, position, and relative sizes of nectaries and their excretory canals. Schematic sketches of the entire cross sections of the ovaries were used to represent the defined nectary types (Fig. 2).

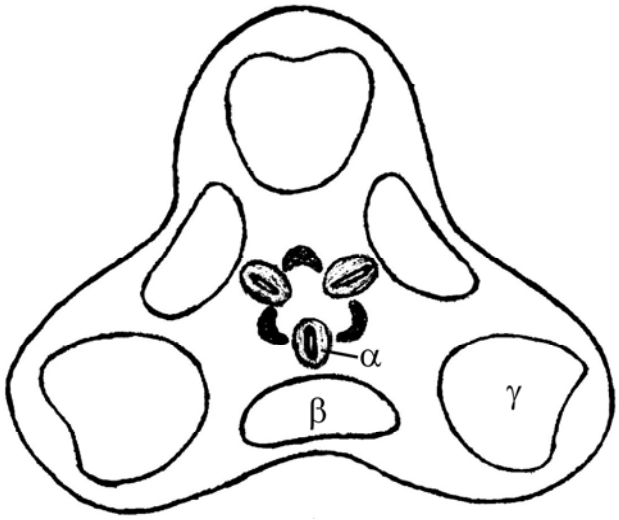


Fig. 2. Schematic representation of the cross-section of an ovary; α – septal nectary, β – excretory canal, γ – locule. In all remaining nectary types in subgenus *Melanocrommyum* excretory canals are not shown (Fig. 3), because they start in the lower part of the nectaries and are not visible on the cross sections of the ovaries.

Molecular methods

In most cases the collected material was dried in silica-gel or frozen in liquid nitrogen. Total genomic DNA-isolation was carried out with the DNeasy Plant Kit (Qiagen) according to the instructions of the manufacturer. DNA concentrations were estimated on 1% agarose gels stained with ethidium bromide. The internal transcribed spacer region (ITS1, 5.8 DNA and ITS2) was amplified with primers ITS-A and ITS-B (Blattner 1999) in 50 µl reaction volume containing 1 U *Taq* DNA polymerase (Qiagen), the supplied reaction buffer, 0.2 µM of each dNTP, 50 pmol of each primer and approximately 20 ng of total DNA. In case of herbarium material ITS-1 and ITS-2 were amplified separately using initial amplification primers (ITS-A and ITS-B) in combination with internal primers binding within the 5.8S rDNA (Blattner, 1999). Due to differences in the primer binding site for ITS-2 in *Melanocrommyum*, a *Melanocrommyum* specific forward primer ITS-D_{mel} (5'-ACTCTTGGCAATGGATATCTT-3') was designed. Thus, for separate amplifications of ITS-1 the primers ITS-A and ITS-C were used, while ITS-2 was amplified with primers ITS-D_{mel} and ITS-B. PCR was performed in a Gene Amp 9700 PCR System (PE Biosystems) with the following cycling program: 95°C – 3 min, 38 cycles of [95°C – 30 sec, 53°C – 45 sec, 68°C – 1 min] and a post-treatment at 70°C for 8 min. In order to weaken DNA secondary structures, Q-solution (Qiagen) was added to the reactions with a final concentration of 20%. Amplicons of approximately 700 base pairs (bp) in length were purified either using Nucleofast 96 PCR plates (Macherey-Nagel) or cut out from an agarose gel, purified with the Nucleospin Gel extraction kit (Macherey-Nagel) and resuspended in 35 µl TE buffer. After checking DNA concentration on 1.8%-agarose gels, 20-40 ng PCR product was directly sequenced on the MegaBACE 1000 automatic DNA sequencer using the respective dye-terminator sequencing technology (Amersham Biosciences). Cycle sequencing was performed using the original amplification primers, with the exception of a nested forward primer ITS-SF_{mel} (5'-CAAGGTTTTCG-TAGGTGAACCTGCG-3'), which was modified from ITS-SF (Blattner et al., 2001) according to the sequences of *Melanocrommyum* species. To test for multiple ITS copies within individuals (intra-individual ITS polymorphism) PCR was performed in three parallel reactions, which were afterwards pooled, purified, and the amplicons were cloned in the pGEM-T Easy vector (Promega) and 8-16 clones per individual were sequenced with the Templi-Phi DNA Sequencing Template Amplification Kit (Amersham Biosciences). Forward and reverse sequences from each template were manually edited and combined in single consensus sequences. The sequences were deposited in the EMBL nucleotide database under accession numbers FM177239 – FM177465.

As in several cases ITS sequences from single species (i.e. different individuals of the same species) fell into different clades in phylogenetic analysis and hybridization was assumed to be frequent in the subgenus (Mes et al. 1999), I tested in this case for ITS capture (homogenization of the rDNA clusters towards one parental sequence after hybridization), by random amplified polymorphic DNA (RAPD) fingerprinting (Welsh and McClelland 1990; Williams et al. 1990). The rationale here was that gene conversion might

occur in the tandemly repeated rDNA but should not involve most other parts of the genome. As genomic fingerprinting methods like RAPD screen arbitrarily wide parts of the genome they should result in different banding patterns if we deal with two cryptic species, while a certain number of common fragments can be expected if hybridization and gene conversion at the rDNA loci was the reason for different intraspecific ITS sequences. The RAPD procedure was carried out for a subset of taxa (30 accessions of 10 species) using 10mer primers A04, B10, C07, C9, D03, C14 and G14, which were obtained from Operon Technologies (Alameda, CA). As reproducibility of bands is a crucial aspect of RAPD analysis, an optimized PCR protocol for *Allium* taxa (Maaß and Klaas 1995, Friesen and Blattner 2000) was used in this study. Among important factors that affect the robustness of RAPD amplification are nuclear DNA content and polyploidy (Bachmann 1997). Considering the huge DNA amount of the taxa under study, the template DNA first was treated with restriction endonuclease EcoR1 to increase the reproducibility of bands. The objective of this step was to reduce non-specific amplification by digesting the genomic DNA (Riede et al. 1994). PCR reaction was carried out in a total volume of 25 µl, containing 1U *Taq* DNA polymerase in the supplied PCR buffer (Qiagen), 0.2 mM of each dNTP, 1.25 mM $MgCl_2$, 5 pmol of primer and 20 ng of digested template DNA. PCR was performed with the following program: 94°C – 3min, initial 10 cycles of [94°C – 30 sec, 43°C – 30 sec, 72°C – 2 min] and next 35 cycles of [94°C – 30 sec, 37°C – 30 sec, 72°C – 2 min], with a post-treatment at 72°C for 10 min. One-third of the reaction mixtures was separated on 1.5% agarose gels in 0.5 TBE and the bands were visualized by ethidium bromide staining (Sambrook et al. 1989).

Data analyses

Aligning of the ITS sequences was done manually by introducing gaps of 1-14 bp lengths at 32 alignment positions. In case of multiple identical sequences resulting from cloned PCR products, only one sequence was included in the data matrix. For many species I sequenced individuals from up to 12 populations, but to avoid overloading the data set only three to four individuals per species, representing the species' geographic range and genetic diversity of the ITS region, were included in the final analyses. The sequences were carefully examined for mosaic patterns, which are mainly caused by recombination of different ITS copies after hybridization. Two obvious chimeric sequences were excluded from the final analyses. The resulting data-set comprising 249 sequences was analyzed with phenetic, cladistic and model-based algorithms. Phenetic and parsimony analyses were performed in PAUP* 4.0b10 (Swofford 2002), Bayesian inference (BI) in MRBAYES 3.1 (Ronquist and Huelsenbeck 2003). MODELTEST 3.7 (Posada and Crandall 1998) was used to test different models of sequence evolution, and the GTR+Γ model was chosen as the best fitted model by the Akaike information criterion. Pairwise genetic distances were calculated on the basis of the GTR+Γ model, and the resulting distance matrix was analyzed with the neighbor joining (NJ) algorithm. For parsimony (MP) analysis the characters were treated as unordered and gaps as missing characters. The parsimony analysis was performed in two

steps (Blattner, 2004). In an initial step heuristic searches were conducted with 200 random addition sequence replicates, hold = 5, with the tree bisection–reconnection (TBR) used for branch swapping, MULPARS option set "on" and a limit of 10 trees saved per replicate. In the second step of the analysis the resulting 1880 MP trees were used as starting trees for TBR in a heuristic search limited to 80,000 trees. Statistical support of the branches was tested with 500 bootstrap re-samples in NJ and with 100,000 re-samples using the 'fast and stepwise' procedure in MP. For BI eight chains were run for 2 million generations under the GTR+Γ model of sequence evolution, sampling a tree every 100 generations. Twelve-thousand initial trees were discarded as burn-in and posterior probabilities were calculated on the basis of the remaining 28,002 trees with MRBAYES 3.1.

2.3 Results

Characteristics of the ITS region and phylogenetic analyses

After introducing the necessary gaps the alignment comprised 680 bp and revealed 474 variable sites, of which 417 were potentially parsimony informative. The sequencing of cloned ITS amplicons revealed divergent ITS alleles within single individuals only in three cases. It is unlikely that these were pseudogenes, as they did not show mutations in the 5.8S region, and matched perfectly the ITS types from other species. These additional copies were included in the final analyses except of the sequences with obvious chimeric patterns (see section 2.2).

The intrasubgeneric variation in the ITS region is high in comparison to other plant groups and reaches 21.7% within the subgenus and 33% between ingroup and outgroup taxa. The ITS variation within the species does not show a uniform pattern. In several species intraspecific diversity is quite high (*A. suworowii* 1.42%, *A. orientale* 0.3%, *A. karataviense* 0.6%, *A. sarawschanicum* 0.2%). In contrast, the relationship of several closely related species (e.g. *A. alexeianum, A. nevskianum, A. protensum*) is even beyond the resolution of the ITS region, as these taxa share identical ITS sequences in all sequenced individuals, and most species show no within-species ITS diversity although we sequenced up to ten (or more) individuals for most species. In other cases, we found that different populations of a species yielded divergent sequences (in the absence of intra-individual ITS polymorphisms) that occur within the same polytomy in the phylogenetic trees. For example, extensive sampling of *A. tashkenticum* resulted in a large polytomy placing all *A. tashkenticum* sequences along with sequences of *A. severtzovoides* and *A. sewerzowii*. It is worth mentioning that similar results were obtained from analyses of chloroplast data (*trn*L-*trn*F spacer, chapter 3). Surprisingly, I found divergent ITS sequences in different individuals of *A. darwasicum, A. hissaricum, A. akaka, A. rosenorum* and *A. stipitatum* rendering these species polyphyletic. ITS amplicons were cloned for all accessions of these species to ensure that there are no intra-individual polymorphisms in rDNA clusters. However, two divergent copies were found only in one accession of *A. akaka*. The second copy was found twice within 10 cloned amplicons of *A. akaka*

accession number 5980 (only one of them, akaka 5980II, is shown in the tree in Fig. 2b. – cluster 5, group H), and it groups with the ITS sequences of *A. tashkenticum* and *A. severtzovioides,* which are 44 mutational steps apart from the remaining eight clones of *A. akaka* 5980. From these two *tashkentikum*-like sequences, in one sequence two positions and in the other sequence seven positions were homologous (probably homogenized) to *A. akaka*-type ITS.

Phylogenetic analyses were conducted with different numbers of outgroup species (2–10). As this did not change tree topologies in the ingroup, the results of the analyses with the two most closely related outgroup taxa (*A. oreophilum* and *A. kujukense*) are presented here. The trees resulting from different phylogenetic algorithms (MP, NJ and BI) differed only slightly in their topologies, therefore the Bayesian tree, combining branch lengths and posterior probabilities (pp) from BI with bootstrap support values (bs) from MP, and the topology-differences between Bayesian and strict consensus trees is shown (Fig. 3).

In all analyses subgenus *Melanocrommyum* proved to be monophyletic. Within the subgenus species are arranged into a basal grade (8 species) and a core clade (95% bs, 1.00 pp) with all the remaining species. The divisions within the grade are statistically well supported, and two species groups (A: *A. fetisowii – A. sp. 7 " chychkanense"*, 100% bs, 1.00 pp; B: *A. robustum – A. tulipifolium – A. chelotum*, 100% bs, 1.00 pp) have high statistical support. *Allium decipiens*' sister relationship to group B is supported by 0.97 pp, and the topology is also recovered in the strict consensus tree although with low bootstrap support (57%). In the basal grade two more species are represented: *A. viridulum* and *A. zergericum.* The sistergroup relationship of *A. viridulum* to all the other taxa is statistically not well-supported (56% bs, 0.83 pp), while *A. zergericum* is placed as sister to the large core clade with 88% bs and 1.00 pp support (Fig. 3a).

The remaining species belong to the core clade, which contains seven clusters. The relationships among these clusters are not resolved, but they are retained as monophyletic in all trees (BI, strict MP and NJ). Three clusters received only moderate or low bootstrap support while the posterior probability values from BI were quite high [cluster 2 (65% bs, 0.95 pp), cluster 5 (<50% bs, 0.99 pp) and cluster 7 (60% bs, 1.00 pp)]. All other clusters were moderately or well-supported in all analyses [cluster 1 (71% bs, 1.00 pp), cluster 3 (77% bs, 1.00 pp), cluster 4 (100% bs, 1.00 pp), and cluster 6 (98% bs, 1.00 pp)].

Some groups within these clusters are retained with high statistical support, some received only low BI support (0.82-0.89 pp) and branch lengths defining the groups are mostly quite short (see discussion), which collapse in MP analysis. Therefore, only congruent groups with high support in all analyses were considered for discussion (groups C – L within the core clade). *Allium sarawschanicum, A. trautvetterianum, A. aroides, A. gypsaceum, A. lipskyanum, A. komarowii, A. koelzii, A. regelii, A. suworowii, A. karataviense* do not fall in these well supported ITS groups (A – L) and collapse in MP analysis. Therefore, their phylogenetic positions could not be resolved with this analysis and, thus, remained unclear.

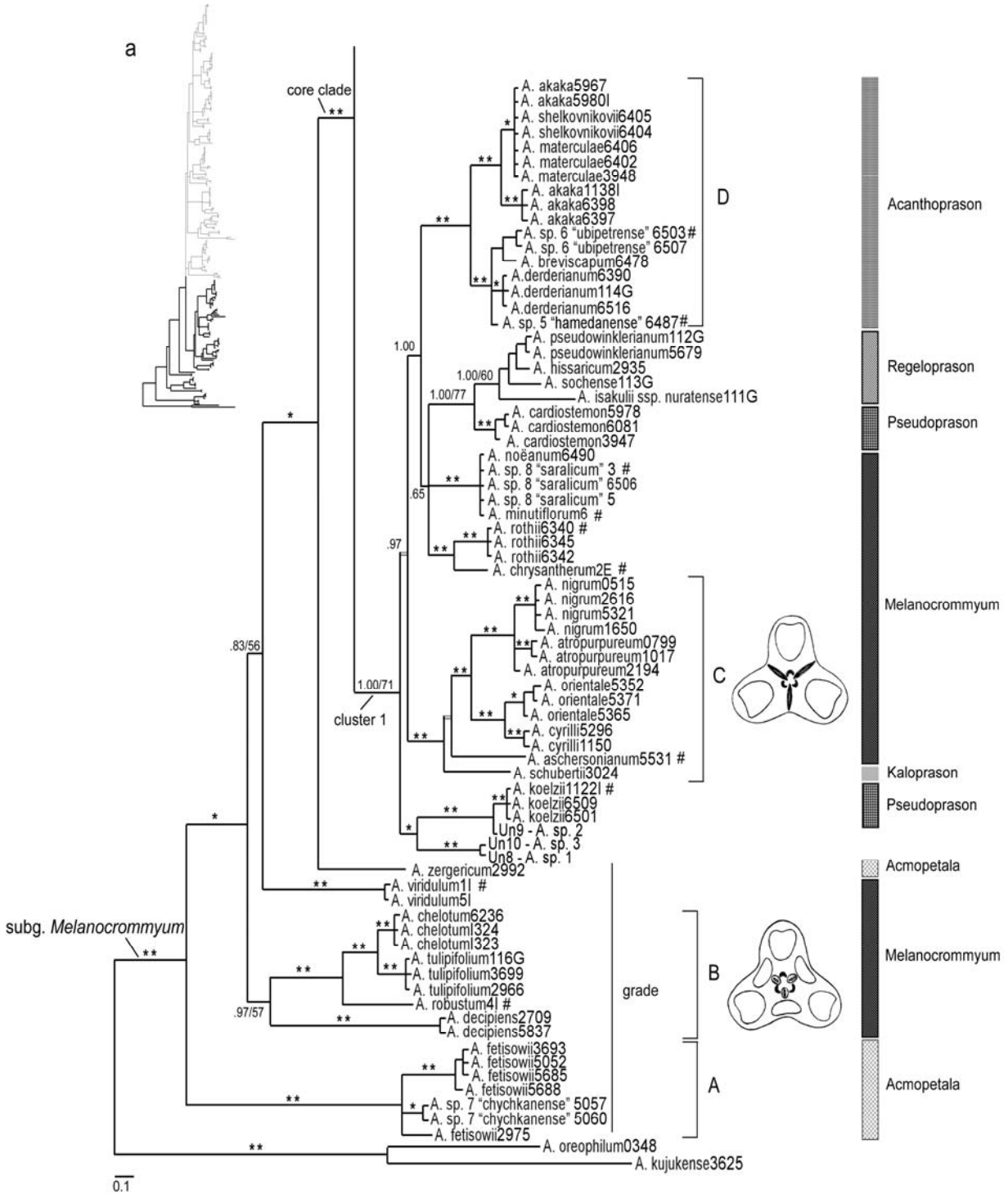


Fig. 3(a – c). Phylogenetic tree resulting from a Bayesian analysis of the ITS sequences from species of subgenus *Melanocrommyum* and two outgroup species. The sequences of hitherto unnamed new taxa (see discussion) are abbreviated with Un. The numbers along the branches depict Bayesian posterior probabilities and bootstrap support values (%). One asterisk indicates posterior probabilities ≥ 0.99 and MP bootstrap values between 80 and 90%, while two asterisks indicate posterior probabilities of 1.0, and bootstrap values ≥ 90%. Branches present in the Bayesian tree and not present in the strict consensus tree of MP are indicated by gray double-lines. Schemes of the cross-sections of ovaries indicate molecular groups where taxa possess the same type of nectaries. Species for which I was not able to study nectary anatomy are indicated by "#". The current taxonomic (sectional) affiliations of species are shown on the right side, where shaded bars correspond to names of the respective sections.

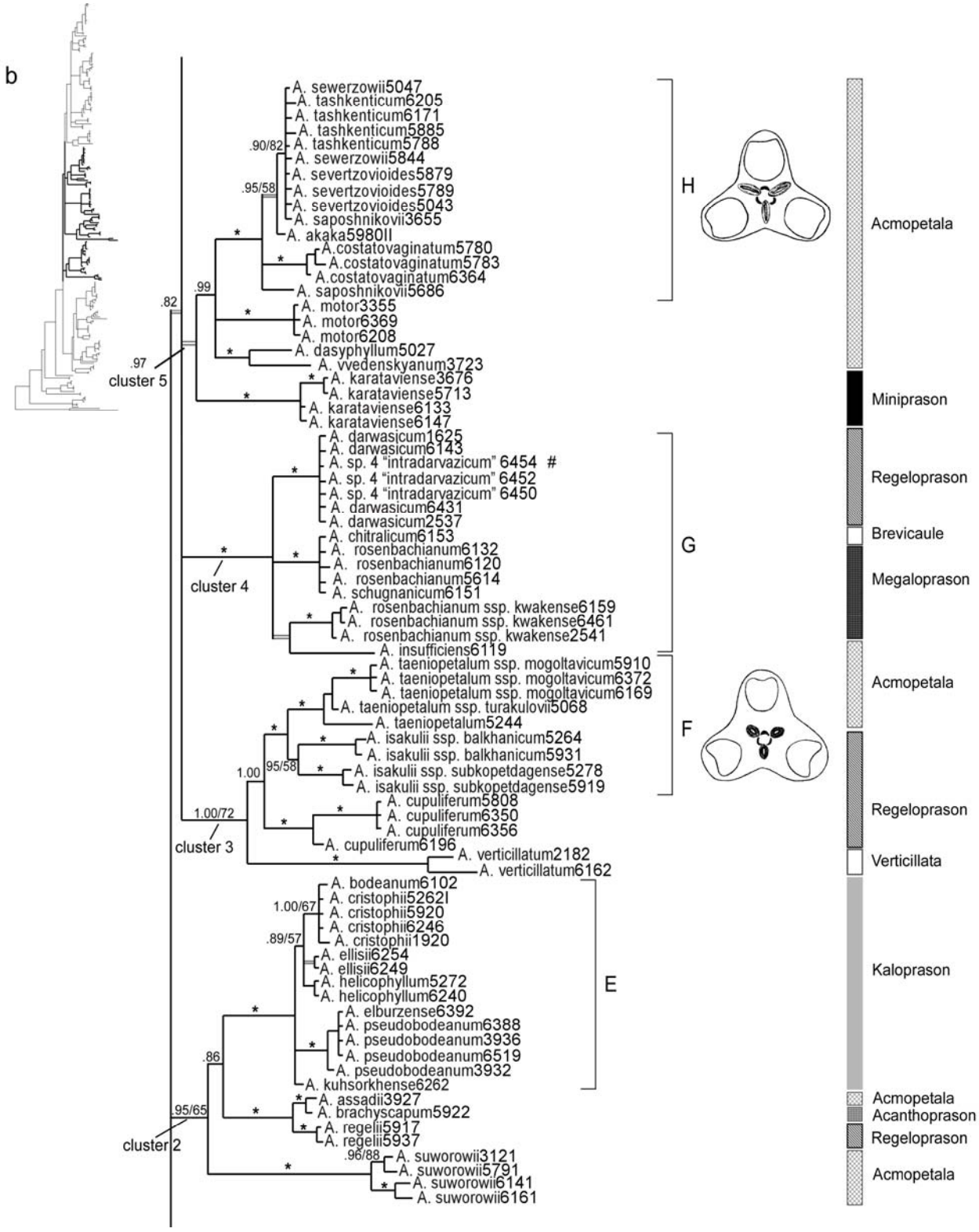


Fig. 3b. continued

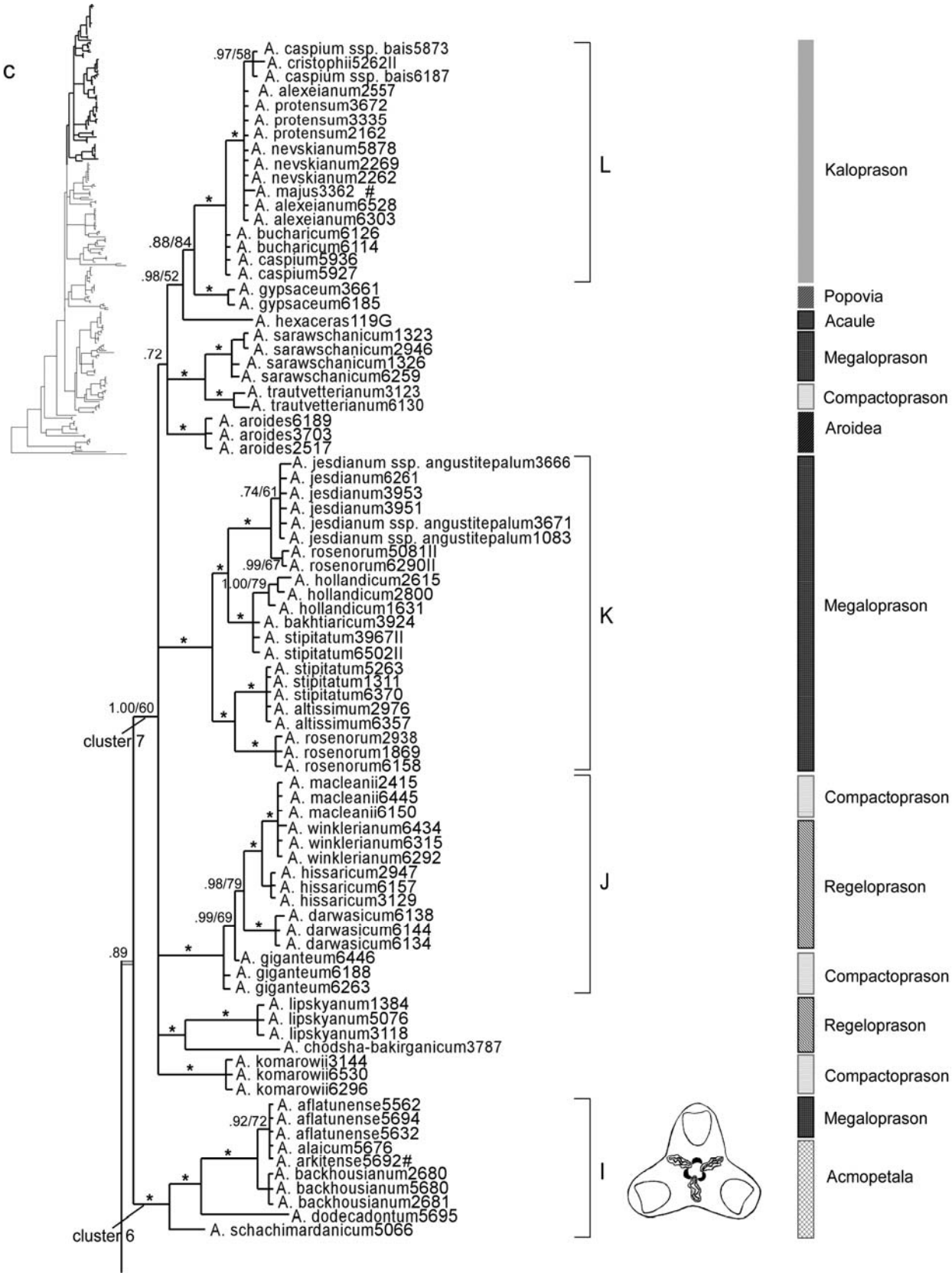


Fig. 3c. continued

A conspicuous feature of the phylogenetic analyses of ITS sequences is the splitting of almost all traditionally defined taxonomic sections into several parts belonging to different well-supported smaller groups, rendering these sections either para- or polyphyletic (Fig. 3). Only small sections (*Popovia, Acaule, Miniprason*) were confirmed as monophyletic units. Several species were found polyphyletic. In *A. isakulii*, where three subspecies were included, subsp. *nuratense* falls into cluster 1, while subsp. *balkhanicum* and subsp. *subkopetdagense* are sisters within cluster 3. Even some species without a proposed taxonomic subdivision (*A. rosenorum, A. stipitatum* and *A. darwasicum*) do not form monophyletic groups, i.e. each of these species consists of two clearly separated groups according to the ITS data. Cloning of amplicons did not reveal paralogous ITS types within even a single individual of these taxa, making recent hybridization an unlikely explanation. *Allium rosenorum* and *A. stipitatum* show strong geographical separation of the individuals belonging to the different groups in the ITS tree. However, in *A. darwasicum* even neighboring plants in the type location population showed strikingly incongruent ITS topologies, as one *darwasicum*-group is a member of cluster 7 while the other occurs in cluster 4, sharing the ITS-type with *A. sp. 4 "intradarvazicum"*. It is difficult to find out whether these morphologically indistinguishable individuals of *A. darwasicum* belong to two clearly separated phylogenetic groups, or if introgression from *A. sp. 4 " intradarvazicum"* to *A. darwasicum* and subsequent homogenization of the rDNA loci towards the introgressed type is responsible for the pattern we see today.

RAPD analysis

RAPD markers were used for accessions forming clusters 4 and 7 in the ITS tree to deduce from polymorphisms in the whole genome whether the morphologically cryptic *A. darwasicum* accessions possess also different fragment profiles compatible with ITS results. Generally, the RAPD banding profiles for six primers were clearly different for the individuals belonging to the two ITS-defined clusters. For primer B10 all profiles were species-specific and shared bands among accessions from different species were completely absent. Except for this primer, *A. darwasicum* accessions belonging to different ITS clusters showed clearly cluster-specific banding patterns in all RAPD analyses (altogether 7 primers), and *A. sp. 4 "intradarvazicum"* and *A. darwasicum* from cluster 4 have nearly identical RAPD profiles (Fig. 4), thus, fully agreeing with the results from ITS data.

Anatomy of septal nectaries

The nectaries within *Melanocrommyum* taxa are located in the lower half of the ovaries, and mostly tubular excretory canals start in the lower third of the nectaries. Only species of group B on the ITS tree possess broadened excretory canals starting in the upper part of nectaries (Fig. 3a). In five ITS-defined groups (B, C, F, H, I) specific nectary types were found, differing in shape and position of nectaries and their canals. Schemes of these types are shown in Figure 3. Splitting of large taxonomic sections according to molecular data

was confirmed by clearly distinguishable nectary types in the following cases. Within section *Melanocrommyum* two genetically quite divergent groups (B and C based on ITS sequences) possess clearly distinguishable types of nectaries (Fig. 3a). Within section *Acmopetala,* three ITS groups (F, H, I) were also discernible by their nectary anatomy (Fig. 3b – c). The species of remaining groups displayed a quite uniform overall shape of the nectaries and position of the excretory canals, similar to the nectary type of ITS group H, and therefore are not shown in the figure. The only difference of this major type compared to type H, is wider and longer nectaries, reaching the margin of the septa. While most phylogenetic groups possess this uniform type, each defined nectary type was detected only in one group, thus I could not find evidence of parallel evolution of this character.

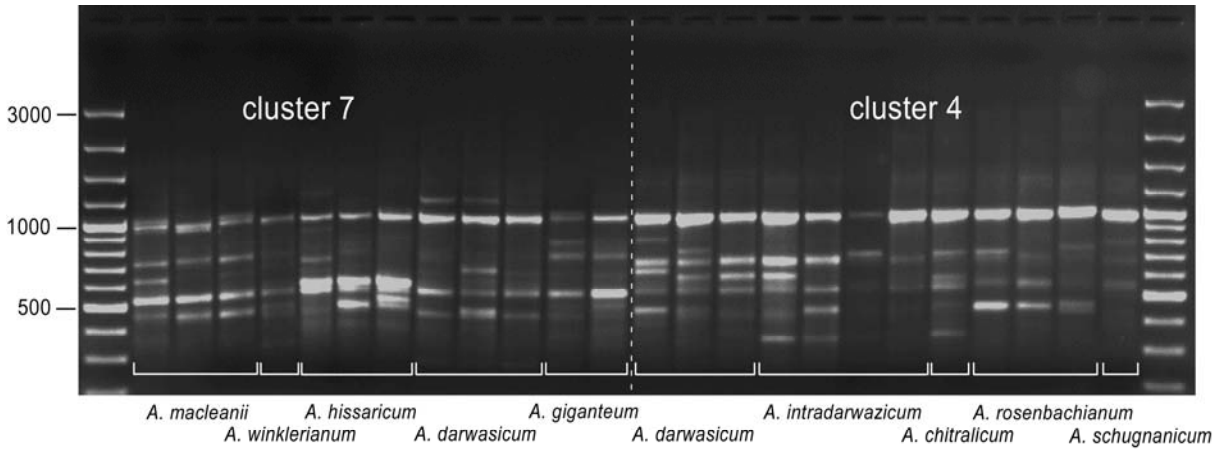


Fig. 4. RAPD reaction profile (Ethidium-Bromide-stained agarose gel) of primer C07. The dashed line separates the individuals belonging to ITS clusters 4 and 7. *Allium darwasicum* individuals from both ITS clusters have clearly different RAPD profiles, indicating that not only ITS sequences but large parts of the genomes are different.

2.4 Discussion

Variation at the ITS locus

The apparently high intrasubgeneric ITS distances (21.7% within the subgenus, which correspond to 0.0936 substitutions per site within subgenus *Melanocrommyum*) was not unexpected taking into account the high genetic distances in the entire genus *Allium* (Friesen et al. 2006). This would imply that (i) either the genus is of old age or (ii) ITS substitution rate is very high in *Allium* and also in subgenus *Melanocrommyum.*

Another conspicuous, although perhaps not unexpected, result is the different level of differentiation at the ITS locus within and among species of *Allium* subgenus *Melanocrommyum*. High genetic diversity at the ITS region within species like *A. suworowii, A. karataviense, A. sarawschanicum,* and *A. orientale* could be caused by an old age and/or large population size of those species. *Allium suworowii*, *A. karataviense,* and *A. sarawschanicum* are geographically widespread and morphologically diverse taxa.

For example, *A. suworowii* is perhaps the most variable and most widely distributed species of section *Acmopetala*, and *A. sarawschanicum* and *A. karataviense* have huge population sizes and consist of different morphotypes. A detailed population genetic analysis and sampling from the entire distribution area would be necessary to test whether high molecular diversity reflects an ongoing evolutionary differentiation correlated with geographical distribution in these widespread species. Hybridization as an alternative explanation is less plausible, because within-individual polymorphism was not detected in these taxa. However, the genetic diversity of *A. orientale* could be connected to disagreeing views concerning taxonomic circumscription of this species (see below, cluster 1, group C).

The topology of three subspecies of *A. isakulii* and two subspecies of *A. rosenbachianum* and *A. taeniopetalum,* which form monophyletic groups belonging to different clusters in ITS analyses indicate that these taxa deserve species rank, though morphological differences are not quite pronounced. On the other hand, the frequent lack of molecular differentiation among several closely related species (despite high overall ITS divergence within the subgenus) cannot be a sampling artifact, as I sequenced multiple accessions for each taxon. The low average pairwise genetic distances in terminal groups consisting of several species suggest that these groups underwent a very recent radiation. Similarly, the reason for the fact that *A. severtzovioides* and *A. sewerzowii* are placed in a polytomy with *A. tashkenticum,* a result also obtained from analyses of chloroplast data (Chapter 3), could be the very recent origin of these species. The cautious interpretation, that these taxa are recently diverged and the ITS gene tree has not yet attained reciprocal monophyly for the constituent species, needs a better estimate of the evolutionary history, which would require multiple unlinked loci and more thorough population sampling.

Divergent ITS types within species and individuals

The existence of divergent ITS types among different individuals of the same diploid species in the absence of intra-individual polymorphisms (sequencing of 8–16 clones failed to detect different ITS copies) could be explained by different phenomena. In *A. rosenorum* and *A. stipitatum,* geographically clearly separated accessions form monophyletic groups in the ITS tree. Therefore, the assumption that they represent independent evolutionary lineages, which seem to have converged on similar phenotypes, is more plausible. In contrast, the topologically different ITS-groups of *A. darwasicum* do not show any geographical structure, as plants from single populations belong to these genetically quite different clades (see discussion below in RAPD analyses). In case of *A. akaka,* the reason could be introgression followed by homogenization of ribosomal DNA, as we found in one individual of *A. akaka* an additional ITS type (see results). Until now *A. akaka* was reported to be diploid (2n = 16; Özhatay 1986) and no data are available about the likely formation of hybrids for this species under experimental conditions. But recently, we detected differences in genome size and ploidy level in some individuals of this taxon (chapter 5), which would support the hybridization hypothesis in this group (group D – cluster 1).

In general, the rare cases where more than one copy was found per individual (three out of 70 cloned accessions) indicate that either gene flow between species within the subgenus is not frequent (though exists), or very fast homogenization of rDNA loci is a common feature in *Melanocrommyum*.

RAPD analysis

Allium darwasicum is widely distributed in Tajikistan. Plants from the type location in the Darvaz range show yellowish-white flowers, while plants from Hissar Mountains have greyish-white flowers, and plants from the Vakhsh mountain range have yellowish-green flowers. The unexpected topological differences in ITS trees between two groups of *A. darwasicum* do however not correlate with the differences in flower color. Moreover, there is no geographical structure related to the ITS groups, as plants from single populations group in both different *darwasicum*-clades. The reason could be either these *darwasicum*-groups represent two different monophyletic units, which converged morphologically, or ancient introgression followed by two-directional homogenization of the ITS loci mingled the picture (Wendel et al. 1995; reviewed in Álvarez and Wendel 2003). This latter scenario would, however, require the complete absence of gene flow among these types. The aim in using RAPD markers was to see whether the genetic differentiation concerns large parts of the genome, and to ensure that the position of *A. darwasicum* in two different clades in the tree does not represent a peculiarity of only the ITS region. However, a full-fledged RAPD analysis (scoring bands and calculating a phenetic tree) was not conducted due to the well-known problems of fragment homology assessment when analyzing taxa with high genetic distances (Rieseberg 1996; Wolfe and Liston 1998), and because I was mainly interested to see if the *A. darwasicum* individuals belong to two different groups also in a genome-wide screening. The patterns in the RAPD profiles were clearly different for the subsets of taxa forming clusters 4 and 7 in the ITS trees, thus confirming that the results obtained by ITS sequencing hold for large parts of the genome of these individuals. The fact that *A. darwasicum* seems to consist of two morphologically indistinguishable different evolutionary lineages belonging to genetically quite different clades within subgenus *Melanocrommyum* adds another example of cryptic species in plants (Martínes-Ortega et al. 2004; Whittall et al. 2004). However, other studies on cryptic species are mostly restricted to sister or closely related plant species (Martínes-Ortega et al. 2004; Fernandez et al. 2006; Grundt et al 2006), while the case of *A. darwasicum* provides an example of sympatric, morphologically identical taxa with high genetic differentiation, which implies quite long reproductive isolation.

ITS phylogeny and incongruence with taxonomic classification

The groups resulting from phylogenetic analyses of the ITS region clearly contradict the existing (or any earlier) taxonomic classification of the subgenus. Most sections are split into several subgroups, and species belonging to different taxonomic sections are clustered

together. All large sections (i.e. *Megaloprason, Melanocrommyum, Compactoprason, Acmopetala, Kaloprason*) were found not to be monophyletic. The most conspicuous example is section *Regeloprason*; its members appear in nearly all well-supported groups along the tree. Below we outline highly supported groups resulting from ITS analyses according to their position on the tree. Although not all the new phylogenetic groups can be morphologically characterized, species of several groups (B, C, F, H, I) share special anatomical structures of septal nectaries.

Groups in the grade:

Group A. There are two species in the group: *Allium fetisowii* s. lat. and *A. sp. 7 "chychkanense"*, which were affiliated to section *Acmopetala* subsection *Longibidentata* (Khassanov and Fritsch 1994), but in our analyses they are clearly separated from other members of that section.

Group B unites *A. decipiens,* and the well supported clade of *A. tulipifolium, A. robustum* and *A. chelotum* from section *Melanocrommyum*. These taxa share the same type of nectaries with large excretory canals, which is unique to this group.

Groups in the core clade:

Cluster 1

Except the supported groups (C and D), several species with unresolved relationships fall in this cluster. The ITS sequence of *A. koelzii*, the type species of section *Pseudoprason*, forms a small well supported group together with the sequences of three at the moment unnamed taxa (abbreviated on the tree with Un8, Un9, Un10). These sequences were obtained from dried herbarium specimens, character combinations of which do not match any hitherto described species, and ITS sequences also differ clearly from all other sequences of known taxa. However, before describing them as new species, study of additional living material from natural populations is necessary. Another species previously included in section *Pseudoprason*, *A. cardiostemon*, forms another small group with members of section *Regeloprason*. However the relationships within this group do not have convincing support. The relationships of *A. minutiflorum*, *A. noëanum* and *A. sp. 8 "saralicum"* sharing identical ITS sequences, and *A. rothii* sister to *A. chrysantherum* (all from section *Melanocrommyum*) are also unresolved.

Group C unites five taxa from section *Melanocrommyum* and a single accession of *A. schubertii* from section *Kaloprason* s. str. (Fig. 3a). *Allium schubertii* is morphologically quite different from other species of this group, although it comes from the same geographical region. The entire group has a circum-Mediterranean distribution. However, another recently sequenced individual of *A. schubertii* (not included in the tree shown here), possesses the same sequence. Thus, the close relationship of *A. schubertii* to this group was confirmed by another individual from a different population, making ITS capture less probable. In addition, the members of this group share extremely narrow septal nectaries. It is worth mentioning that in this group, *A. orientale* is taxonomically unclear. Some authors restrict its natural occurrence to southern Turkey (where also our material has come from),

and others accept this species in a much wider sense including plants from Syria and Israel to Egypt.

Group D. Species of section *Acanthoprason* falling into this group *(A. akaka, A. materculae, A. shelkovnikovii, A. derderianum, A. sp. 6 "ubipetrense"* and *A. sp. 5 "hamedanense"*) are known to *Allium* taxonomists extremely difficult to determine, as they are morphologically diverse with intermediate characters. The fact that, via extensive cloning, different ITS types were found in one *A. akaka* accession could support the hypothesis that hybridization is not rare in this group. The genetically distant Turkish accessions (5967, 5980) of *A. akaka* are also morphologically well discernable from the Iranian ones and might represent a separate species.

Cluster 2

Allium suworowii, hitherto included in section *Acmopetala* is another species where the phylogenetic relationship could not be resolved with these analyses. Despite the fact that its sister relationship to the remaining taxa in the cluster is present in both Bayesian and MP trees, the moderate bootstrap support (65% bs) and pronounced morphological differences to other taxa in this cluster do not allow to draw any conclusion about the phylogenetic relationship of *A. suworowii*.

Group E. Taxa belonging to this group *(A. cristophii, A. elburzense, A. pseudobodeanum, A. kuhsorkhense, A. ellisii, A. helicophyllum)*, were hitherto affiliated to two different sections *Acanthoprason* and *Kaloprason*. However, these taxa share several similar morphological characters, which could support their recognition as a separate section.

Cluster 3

Beyond the well-supported group F, there are two additional species included in this cluster. *Allium cupuliferum* (section *Regeloprason*) bearing morphological similarities to *A. isakulii*, is sister to group F, however with bootstrap support below 50%. *Allium verticillatum,* the only species of section *Verticillata,* is morphologically very distinct and appears as sister to the remaining members of the cluster with relatively good statistical support. However this could be the result of long branch attraction (Bergsten 2005), taking into account the long branch leading to *A. verticillatum*.

Group F. All taxa of this group were previously assigned to either section *Regeloprason* or *Acmopetala*. As it was already mentioned in the results, the topology of the subspecies of *A. taeniopetalum* and *A. isakulii* would better correspond to recognition of these subspecies as separate species.

Cluster 4

Group G. The group unites several members of section *Megaloprason* and two species of section *Regeloprason* (several accessions of *A. darwasicum*, see detailed discussion above). All analyzed accessions of *A. rosenbachianum* subsp. *kwakense* are clearly separated from *A. rosenbachianum* subsp. *rosenbachianum*. The topology of these subspecies and high bootstrap values for the divisions favor their recognition as separate species.

Cluster 5

The monotypic section *Miniprason* (*A. karataviense*) is strongly supported as a single unit, but its inclusion in cluster 5 requires support from additional markers, as this topology was inferred only by Bayesian analysis, although with relatively high support (0.97 pp).

All remaining taxa in this cluster belong to section *Acmopetala. Allium motor* and and a clade consisting of *A. dasyphyllum* and *A. vvedenskyanum* form a polytomy with the members of group H.

Group H consists of morphologically similar, although clearly distinguishable species, naturally occurring in the western Tian Shan region. Shared ITS sequences by several members of the group (*A. severtzovioides* and *A. sewerzowii, A. tashkenticum*) together with morphological similarity could indicate a relatively recent diversification.

Cluster 6

Group I. All taxa in this group except *A. aflatunense* belong to section *Acmopetala*. This is a clear geographically defined group as all members of this clade stem from the Fergan depression, the Alai mountain range in the south, or from the contact area of Fergan and western Chatkal mountain ranges (both belong to Tien Shan range) in the north-east. This group possesses sinuous nectaries, totally different from all other types (Fig 3c).

Cluster 7

This is a large cluster uniting three groups (J, K, L) and several species forming a polytomy. The relationships among monotypic sections *Aroidea*, *Popovia* and *Acaule* remain unresolved.

Unexpected were the positions of *A. sarawschanicum* placed phylogenetically distant from the morphologically similar *A. rosenbachianum* (cluster 4, group G) and its sister *A. trautvetterianum*. But *A. sarawschanicum* differs from *A. rosenbachianum* by having outgrowths at the top of the ovaries, a character that is completely missing in the *A. rosenbachianum* alliance.

Group J. In addition to several species of section *Regeloprason*, this group comprises *A. giganteum* and *A. macleanii* from section *Compactoprason*. Unexpectedly, *A. winklerianum* and *A. macleanii* have identical ITS sequences. These species show quite pronounced morphological differences, but are growing in the same geographical region.

Group K unites the Northwestern Iranian taxa of section *Megaloprason.* This clade contains *A. rosenorum* and *A. stipitatum,* both split in two evolutionary independent lineages based on ITS data. These different lineages within both taxa are also geographically clearly separated (see above). The ornamental species *A. hollandicum* (*A. aflatunense* of bulb traders) is also embedded in this group. The molecular data presented here strongly suggest an Iranian origin of this taxon.

Group L. All species of this clade are taxonomically affiliated to section *Kaloprason.* The data shown here strongly support the separation of this group from other members of this section (*A. cristophii* and *A. elburzense*) which form another well separated clade (group C).

Even this brief outline of the clades shows clearly that most morphological characters used for previous classifications do not show a clear correlation to the groups resulting from the analyses of ITS sequence data.

Reasons for the molecular-taxonomical discordance

Incongruence between molecular data and morphology-based taxonomies has been repeatedly reported (Sytsma 1990; Wendel and Doyle 1998; Albach et al. 2004; Hörandl et al. 2005). The reason could be parallel or rapid evolution of morphological characters, ancient hybridization events, or/and rapid diversification. As a single genetic change in a regulatory region can cause dramatic morphological transformations, which typically are unaccompanied by similar levels of molecular divergence (rate differences in molecular and morphological evolution), the result will be incongruence between molecular phylogenies and morphology (reviewed in Wendel and Doyle 1998). In addition to evolutionary rate differences, convergent evolution of morphological characters may also be a reason of incongruence (Sytsma 1990; Givnish and Sytsma 1997; Albach et al. 2004).

Another obvious cause of incompatibilities between molecular and morphological data is reticulation and introgression, which turned out to be an even more widespread phenomenon in plants than earlier suspected on morphological grounds. Given the prevalence of reticulation (Comes and Abbott 2001; Schwarzbach and Rieseberg 2002; Pan et al. 2007) and polyploidization (Soltis et al. 2007), difficult to document ancient hybridization events could often be responsible for the phylogenetic incongruence.

Due to fast diversification (rapid radiation), phylogenetically inferred internodes on gene trees may be short and difficult to resolve with confidence. This short interior branch phenomenon may be also a common cause of phylogenetic incongruence (Wendel and Doyle 1998). Taking into account the unresolved ITS phylogeny of the core clade representing the majority of species of subgenus *Melanocrommyum,* this diversification could represent a rapid radiation, as rapid evolutionary radiations have been proposed to explain poorly resolved phylogenies in many groups of organisms, including birds (Cooper and Penny 1997), mammals (Murphy et al. 2001) and higher plants (Mathews and Donoghue 1999; Soltis et al. 2004). However, poor internal branch support in phylogenies can also be obtained by using molecular or morphological data that are not variable enough at the appropriate level, having data sets strongly conflicting with one another (Whitfield and Lockhart 2007), or by the loss of phylogenetic signal through substitutional saturation. Taking into consideration the high divergence values (and maybe the accelerated rates of evolution in addition) for the ITS sequences in *Melanocrommyum,* it becomes extremely difficult to distinguish whether the reason is rapid radiation or conflicting signals in the data. Homoplasy-free rare changes might be the only option to distinguish between these phenomena (Boore 2006).

All the phenomena discussed above, like parallel or rapid evolution of morphological characters, ancient hybridization events, or/and rapid diversification could to some extent be the reason for the conflict between ITS results and several characters used for sectional

delimitation in previous taxonomic treatments. But I want to emphasize that as polyploidy is rare in the subgenus and there is no indication of frequently ongoing introgression within *Melanocrommyum,* parallel evolution of morphological characters and rapid diversification seem the more likely explanation.

Nectary types

The different nectary types present in section *Melanocrommyum* (two types) and section *Acmopetala* (three types) confirm the clades recovered in the phylogenetic analyses of ITS sequences, thus, fully agreeing with molecular data. However, the anatomy of nectaries was quite uniform in the remaining groups, as the detected differences were only quantitative which changed without any discernible pattern or trend even within different accessions of the same species. As such differences could be caused by environmental factors as well as by changes during ontogenetic development, they will not be discussed here. The fact, that the majority of the phylogenetic groups show the uniform type of nectaries, allows to hypothesize its plesiomorphic state, while all other types occurring only once (each type in one ITS group) could represent derived characters, synapomorphic to the respective clades. However, the poorly resolved phylogenetic tree with a large polytomic core clade hampers testing the hypotheses of character state evolution within the subgenus.

2.5 Conclusions

The ITS data support the monophyly of the subgenus *Melanocrommyum* and indicate multiple radiations within the group. Although the relationships are not well resolved among all groups, the phylogenetic analyses favor the circumscription of smaller sections. However, a new taxonomic classification would involve the description of a number of new taxonomic units, particularly sections as well as species, which is not the part of this thesis.

Although extensive cloning of ITS amplicons could not detect evidence for frequent ongoing introgression within the subgenus, ancient hybridization events followed by homogenization of ITS repeats cannot be excluded.

The present study provides molecular evidence for distantly related cryptic species in *Allium* subgenus *Melanocrommyum*. However, the existence of morphologically identical and genetically clearly separated lineages, as well as the underlying mechanisms, need further investigation.

3. Species level phylogeny of the subgenus *Melanocrommyum* – phylogenetic and genealogical analyses of noncoding chloroplast DNA

3.1 Introduction

It is well known, that intraspecific gene evolution cannot always be represented by a bifurcating tree. Despite the abundance of methods available for phylogenetic inference (Swofford et al. 1996) few network building methods were proposed (Posada and Crandall 2001) suitable for estimating intraspecific haplotype trees, where relationships among haplotypes are often multifurcated, descendant genes coexist with persistent ancestors and recombination events produce reticulate relationships. These network approaches have recently been developed to estimate intraspecific genealogies that take into account these population level phenomena, whereas traditional phylogenetic methods used at species-level phylogenies assume bifurcating trees. However, there are plenty of reasons which could violate the assumptions of the strictly bifurcating tree not only in intraspecific (Posada and Crandall 2001), but also in interspecific relationships (Maddison 1997). Recent studies showed, that this is particularly critical for phylogenetic studies using chloroplast DNA, where ancestral polymorphism, incomplete lineage sorting, and multifurcating relationships of chloroplast haplotypes could hinder correct phylogenetic inference using phylogenetic methods assuming bifurcating trees (Jakob and Blattner 2006; Bänfer et al. 2006).

There are several reasons why the chloroplast genome is generally accepted as an important source of molecular information in plant systematic and phylogeographic studies (Olmstead and Palmer 1994; Newton et al. 1999; Hewitt 2001). First, uniparental inheritance of chloroplasts (Reboud and Zeyl 1994; Birky 1995) avoids problems caused by recombination. Although this holds for most angiosperms, groups with biparental and/or paternal inheritance exist, like *Oenothera* (Chiu et al. 1988), *Pelargonium* (Metzlaff et al. 1981), *Medicago* (Smith 1989), *Turnera* (Shore et al. 1994; Clément and Pacini 2001), and *Passiflora* (Hansen et al. 2007), which makes important to know the mode of plastid inheritance for phylogenetic analyses (Harris and Ingram 1991). Second, even within an individual the possibility of recombination between genomes from individual plastids is extremely low, and there are only a few studies describing the occurrence of multimeric chloroplast DNA (cpDNA) genomes or interchromosomal cpDNA recombination (Govindaraju et al. 1989; Dally and Second 1990). Third, land plant plastomes are known for their high degree of conservation in size, structure, gene content, and linear order of genes (Palmer 1991; Downie and Palmer 1992).

Although among the various marker systems coding sequences played an important role in phylogenetic studies, especially when molecular clock approaches and multigene data sets were assembled, noncoding DNA sequences are widely employed markers at species level phylogenetics. Among them, one of the most often used regions in plant systematics and phylogeography is the plastid *trn*L (UAA)-*trn*F (GAA) region since Taberlet et al. (1991) introduced universal primers to amplify three noncoding regions of chloroplast

DNA. The marker not only provided phylogenetic signal to resolve deep angiosperm phylogeny (e.g. the *trn*T-*trn*F region was used by Borsch et al. (2003) in basal angiosperm phylogeny) but also revealed extensive haplotype variation to elucidate speciation processes on the population level (e.g. Dobeš et al. 2004; Jakob et al. 2007). The *trn*L-*trn*F genes are co-transcribed (Kanno and Hirai 1993), and therefore it can be assumed that intron as well as spacer regions are of functional importance. However, despite the extremely high utility of this marker, we still have to learn more about the *trn*L-*trn*F and other noncoding plastid regions that are widely used in plant phylogenetics and systematics, as it was shown that deep coalescence (Hudson 1990) combined with incomplete sorting of ancient lineages can be misleading in phylogenetic inferences not only in young and rapidly speciating plant groups (Mason-Gamer et al. 1995; Comes and Abbott 2001) but also in old lineages reaching deep into the history of a genus (Jakob and Blattner 2006). Moreover, several studies report multiple independent origins of *trn*F pseudogenes in crucifers (Koch et al. 2005, 2007), and in Asteraceae (Vijverberg and Bachmann 1999; Wittzel 1999). Additionally, Pirie et al. (2007) found two copies of the *trn*L-*trn*F region in Annonaceae. They showed that *trn*L-*trn*F sequences display a phylogenetic signal conflicting with that of other chloroplast markers in Annonaceae and demonstrated that this conflict results from using different paralogous copies for phylogenetic inferences.

In the present chapter I analyzed the *trn*L-*trn*F region in *Allium* subgenus *Melanocrommyum*, sequencing multiple accessions for 100 species to get insights into the phylogeny of the subgenus. Here I discuss the utility and limitations of this region for species-level phylogenies on the example of subgenus *Melanocrommyum*, and report the occurrence of *trn*F gene duplication in the subgenus. Despite the fact that the major phylogenetic lineages resulting from analyses of the *trn*L-*trn*F region are consistent with the clades recovered from nuclear rDNA ITS in *Melanocrommyum* (Chapter 2), the analyses show that ancient chloroplast haplotypes co-occur with their descendants not only at the intraspecific, but also at the interspecific level, which together with incomplete sorting of ancestral alleles and chloroplast capture (although the latter was rarely detected in the present study) could hinder the correct species-level phylogenetic inferences from chloroplast data in *Allium*.

3.2 Materials and Methods

Taxon sampling

Totally 440 individuals were studied representing about 100 species, some of them only informally named new species. The most closely related taxa (*A. oreophilum* and *A. kujukense*) based on a nuclear rDNA ITS phylogeny (Gurushidze et al. 2008) were chosen as outgroups. Sampling of this study covers all existing taxonomic groups of subgenus *Melanocrommyum*, and their entire geographic distribution, analyzing multiple individuals per species wherever available. Leaf material for DNA extraction was obtained from the Taxonomic *Allium* Reference Collection of the IPK Gatersleben (Germany), collected in

natural populations, and sometimes extracted from herbarium specimens. Some newly collected material turned out to be new species which need to be described, and several species collected this year are not definitely determined. The undescribed species names are in quotation marks (as in Chapter 2), whereas not definitely determined species are abbreviated with *spec* followed by accession number.

Molecular methods

Genomic DNA was extracted from about 50 mg fresh or 10 mg silica-dried leaf material using the DNeasy Plant Kit (Qiagen, Hilden, Germany) according to the instructions of the manufacturer. The *trn*L-*trn*F region comprising the *trn*L(UAA) gene including a group I intron, the *trn*F(GAA) gene and the corresponding intergenic spacer of *Allium* subgenus *Melanocrommyum* was amplified using the forward primer 5'-GGAAATGGGGATATGGCG-3' that anneals in exon 1 of the *trn*L gene, and a reverse primer 5'-ATTTGAACTGGTGACACGAG- 3' annealing in the *trn*F gene (Taberlet et al. 1991). Polymerase chain reaction (PCR) was performed with a GeneAmp 9700 PCR System (PE Biosystems, Foster City, CA) with the following cycling program: a 95 °C denaturation step for 3 min was followed by 35 cycles each comprising denaturation at 95 °C for 30 sec, annealing at 55 °C for 35 sec, and elongation at 68 °C for 2 min. Afterwards a post treatment at 70°C for 15 min was conducted. The reaction was carried out with 1.5 U Taq DNA polymerase (Qiagen) in the supplied reaction buffer, 0.2 mM of each deoxynucleoside triphosphate, 25 pmol of each primer, and about 20 ng of total DNA in a final reaction volume of 50 µl. Q-Solution (Qiagen) was added to a final concentration of 20% to weaken DNA secondary structures during PCR. The amplicons of about 600 bp in length were purified using Nucleofast PCR Plates (Macherey-Nagel, Düren, Germany). For cycle sequencing a nested forward primer (5'-CGAAATCGGTAGACGCTACG- 3') was used, whereas the reverse primer was the same used for PCR amplification. PCR products were directly sequenced on a MegaBACE 1000 automatic DNA sequencer (Amersham Biosciences, Freiburg, Germany) or an ABI 3730xl automatic DNA sequencer (Applied Biosystems) using the respective dye-terminator sequencing technology.

Data analyses: phylogenetic inference and maximum parsimony network

Forward and reverse sequences of each sample were manually edited and combined into single consensus sequences. Aligning of the sequences could be done manually by introducing gaps of 1 to 10 basepairs (bp) in length. In four cases the lengths of deletions were much longer, varying from 71 to 95 bp. Identical sequences were grouped into haplotypes. The two poly- T/A mononucleotide repeat regions, which showed homoplastic length variation, and one hypervariable region adjacent to the second poly- T/A stretch were excluded from the phylogenetic analyses. The analyses were done using two different data matrices with tree building algorithms: one matrix containing all sequences of 440 individuals, and the other reduced matrix containing single sequences of each haplotype.

The reduced data matrix consists of 74 unique sequences (i.e. haplotypes) found in subgenus *Melanocrommyum*. The *trn*L-*trn*F sequences from *A. oreophilum* and *A. kujukense* served as outgroups in both analyses. The maximum parsimony optimality criterion (MP) was applied for reconstruction of phylogenies using PAUP* 4.0b10 (Swofford 2002). The heuristic search algorithm was chosen, using the random addition of taxa and the tree bisection–reconnection (TBR) option for branch swapping, gap mode was set to missing, but all indels present in the sequences were coded as additional single characters in the nexus input file. The parsimony analysis was performed in two steps (Blattner 2004), with the same settings as described in Chapter 2. A strict consensus tree was constructed from 80,000 shortest trees. Bootstrapping was carried out on 500 replicates using the heuristic search option. For Bayesian inference the model of sequence evolution was tested in MODELTEST 3.7 (Posada and Crandall 1998), and K81uf+I model was chosen as the best fitted model by the Akaike information criterion. For Bayesian analyses (BI) eight chains were run for 3 million generations under the respective model of sequence evolution, sampling a tree every 100 generations. The posterior probabilities were calculated with MRBAYES 3.1 (Ronquist and Huelsenbeck 2003) after discarding the initial 25% (non-stationary) of the resulting trees.

TCS 1.20 (Clement et al. 2000) was used for a statistical parsimony network approach. The network-building algorithm (Templeton et al. 1992) joins haplotypes according to the number of the mutational differences between sequences. For this analysis each insertion/deletion (indel) was considered as a single mutation event, and all indels were therefore reduced to single characters in a final alignment. In cases where it was not possible to code the indel variation by a single alignment position (overlapping indels or point mutations in insertions), the network was calculated with the necessary number of alignment positions to represent all variation at these sites, and mutation steps were afterward adjusted manually. To account for the presumably different mutation rates underlying base substitutions and microsatellites, a two-step strategy was followed in the construction of a statistical parsimony network from haplotype data (Bänfer et al. 2006; Jakob et al. 2007). In the first step, all length variation at mononucleotide repeats (i.e. chloroplast microsatellites consisting of poly- T/A residues) and one hypervariable indel were excluded from the analysis, and a backbone network of 74 haplotypes was constructed. In the second step, the variation at microsatellite loci was used to define subhaplotypes within the haplotypes of the backbone network.

Some closed loops in the network, caused by homoplasious mutations, could be resolved according to the criteria outlined in Crandall and Templeton (1993).

3.3 Results

Sequence variation and inference of chloroplast haplotypes

The sequenced *trn*L-*trn*F region of the *Melanocrommyum* individuals was in a range between 492 (in *A. hissaricum* accessions with the longest deletion of 95 bp - HT33) and 633 bp, resulting in an alignment of 665 bp in length after introducing the necessary gaps.

The indels in the *trn*L-*trn*F intergenic spacer were responsible for the largest amount of length polymorphism observed. Excluding the length variation at 2 mononucleotide repeats and one hyper-variable indel resulted in an alignment length of 640 bp, the coding of each indel for phylogenetic analyses resulted in an alignment of 653 bp. In this matrix 25 out of 72 variable positions were potentially parsimony informative. Treating every indel as a single mutational event shortened the alignment to 556 bp (used in network approach). The variation at the *trn*L-*trn*F region resulted in 74 chloroplast haplotypes. Fifty of them were exclusive for single species (species-specific), whereas 24 haplotypes were shared among up to 15 species. Absolute frequencies of particular haplotypes varied from singletons, found only once, to 89 in case of HT 27. Analyzed species, number of individuals and haplotype numbers found in each species are given in the Table 2.

Table 2. Haplotype distribution among *Melanocrommyum* species.

Species (subspecies)	No of individuals	Haplotype No
Allium aflatunense B. Fedtsch.	4	24, 25
Allium akaka S.G. Gmelin ex Schult. et Schult. f.	6	1d, 4, 74a, 72
Allium alaicum Vved.	1	24
Allium alexeianum Regel	4	50, 55, 57
Allium altissimum Regel	2	27
Allium aroides Popov & Vved.	2	47
Allium atropurpureum Waldst. et Kit.	3	1a, 5
Allium backhousianum Regel	4	25
Allium bakhtiaricum Regel	1	27
Allium brachyscapum Vved.	1	38
Allium breviscapum Stapf	1	65b
Allium bucharicum Regel	2	55
Allium cardiostemon Fischer et Meyer	6	1b, 1c, 2, 3
Allium caspium (Pall.) M. Bieb. subsp. *baissunense* (Lipsky) F.O. Khass. et R.M. Fritsch	2	55
Allium caspium (Pall.) M. Bieb. subsp. *caspium*	2	55
Allium chelotum Wendelbo	2	9, 10
Allium chitralicum Wang et Tang	1	46
Allium chrysantherum Boiss. et Reut.	1	15
Allium costatovaginatum Kamelin et Levichev	4	20, 21, 22
Allium cristophii Trautv.	4	36, 37
Allium cupuliferum Regel subsp. *cupuliferum*	4	55
Allium cyrilli Tenore	3	1a, 13, 14
Allium darwasicum Regel	7	38, 42, 43
Allium decipiens Fischer ex Schult. et Schult. f.	2	6
Allium derderianum Regel	3	65a
Allium dodecadontum Vved.	2	25
Allium elburzense Wendelbo	3	36
Allium ellisii J.D.Hooker	3	36
Allium fetisowii Regel	4	7a
Allium giganteum Regel	8	27, 33
Allium gypsaceum Popov et Vved.	2	53, 54
Allium haemanthoides Boiss. et Reut. ex Regel s. str.	1	65c
Allium helicophyllum Vved.	1	38
Allium hissaricum Vved.	3	33, 38
Allium hollandicum R.M. Fritsch	5	27, 31
Allium hooshidaryae Mashayekhi, Zarre et R.M. Fritsch	1	70
Allium insufficiens Vved.	1	38
Allium isakulii R.M. Fritsch et F.O. Khass. subsp. *balkhanicum* R.M. Fritsch et F.O. Khass.	2	55
Allium isakulii R.M. Fritsch et F.O. Khass. subsp. *subkopetdagense* R.M. Fritsch et F.O. Khass.	2	60
Allium jesdianum Boiss. et Buhse subsp. *angustitepalum* (Wendelbo) F.O. Khass. et R.M. Fritsch	10	55

Allium jesdianum Boiss. et Buhse subsp. *Jesdianum*	7	27, 55
Allium jesdianum Boiss. et Buhse subsp. remediorum R.M. Fritsch ined.	2	27
Allium karataviense Regel	15	17, 18, 19
Allium kazerouni Parsa	1	27
Allium keusgenii R.M. Fritsch ined.	1	68
Allium koelzii (Wendelbo) K.Persson et Wendelbo	8	65c, 67, 72, 74a
Allium komarowii Lipsky	8	49, 50
Allium kuhsorkhense R.M. Fritsch et Joharchi	1	45
Allium lipskyanum Vved.	3	35
Allium macleanii J. G. Baker	13	27, 33, 34
Allium majus Vved.	1	55
Allium materculae Bordz.	3	65a, b
Allium materculae Bordz. subsp. *graveolens* R.M. Fritsch ined.	5	65a, 65c, 66
Allium motor Kamelin	10	17, 20, 21
Allium nevskianum Vved. ex Wendelbo	7	27, 48, 55
Allium nigrum L.	3	1a, 16
Allium noëanum Reut ex Regel	1	1a
Allium orientale Boiss.	4	1a, 12
Allium protensum Wendelbo	6	55
Allium pseudobodeanum R.M. Fritsch et Matin	2	36
Allium regelii Trautv.	3	38, 39
Allium robustum Kar. et Kir.	1	7c
Allium rosenbachianum Regel subsp. rosenbachianum	6	38
Allium rosenbachianum Regel subsp. *kwakense* R. M. Fritsch	3	38, 42
Allium rosenorum R. M. Fritsch	27	27, 30, 32, 56
Allium saposhnikovii E. Nikitina	3	20
Allium sarawschanicum Regel	13	51, 55, 61, 62, 64
Allium schachimardanicum Vved.	1	1a
Allium schubertii Zucc.	4	16
Allium schugnanicum Vved.	1	38
Allium scotostemon Wendelbo	1	36
Allium sergii Vved.	1	18
Allium severtzovioides R.M. Fritsch	8	17, 20, 21,23
Allium sewerzowii Regel	1	20
Allium shelkovnikovii Grossh.	2	65b
Allium sp. 10 "bisotunense"	1	65c
Allium sp. 11 "chloranthum"	1	16
Allium sp. 12 "farsianum"	1	73
Allium sp. 4 "intradarvazicum"	3	41
Allium sp. 5 "hamedanense"	1	65c
Allium sp. 6 "ubipetrense"	7	65a, c
Allium sp. 7 "chychkanense"	2	8
Allium sp. 8 "saralicum"	1	11
Allium sp. 9 "albokurdicum"	1	65c
Allium stipitatum Regel	42	27, 28, 29, 35
Allium suworowii Regel	7	38, 40, 44
Allium taeniopetalum Popov & Vved subsp. *mogoltavicum* (Vved.) R.M. Fritsch et F.O. Khass.	3	55
Allium taeniopetalum Popov & Vved subsp. *taeniopetalum*	4	55, 58
Allium taeniopetalum Popov & Vved subsp. *turakulovii* R. M. Fritsch et F.O. Khass.	1	63
Allium tashkenticum F.O. Khass. et R. M. Fritsch	8	17, 20, 21
Allium trautvetterianum Regel	4	55
Allium tulipifolium Ledeb.	3	7b, c
Allium ubipetrense R.M. Fritsch ined.	1	65c
Allium verticillatum Regel	3	47, 52
Allium viridiflorum Pobedim.	1	26
Allium viridulum Ledeb.	2	7b
Allium vvedenskyanum Pavlov	1	20
Allium winklerianum Regel	3	38
Allium zagricum R.M. Fritsch ined.	3	65a, c

Adding the sequence variations at two mononucleotide repeat loci and one hypervariable region at the second step of network construction resulted in 80 chloroplast haplotypes. These microsatellite based haplotypes are indicated by original haplotype numbers extended with lowercase letters (e.g. 1a, 1b) in Fig. 7.

*Trn*F duplication

A duplication of the *trn*F gene was found in all *Melanocrommyum* accessions and also in both outgroup taxa. The differences between the duplicated copy and the functional *trn*F gene, which is quite conserved among all land plants, are not very pronounced. There are length differences caused by three single-nucleotide indels: two insertions (one Adenine, one Thymine), and one Cytosine deletion. In addition, there are three transversions in the anticodon domain, and one Adenine → Guanine transversion in the T-domain. All the above mentioned mutations are identical in duplicated (divergent) *trn*F copies of all *Melanocrommyum* individuals, and all these mutations are absent in the *trn*F gene of *Melanocrommyum, Triticum, Nicotiana,* and *Lactuca*. Therefore I assume that the divergent *trn*F copy is a pseudogene, which is placed in the intergenic spacer between *trn*L and *trn*F gene in all sequences.

In contrast to other studies reporting about *trn*F pseudogenes (in Brassicaceae and Asteraceae), where the authors found different number of tandemly repeated pseudogene copies in close 5' proximity of the functional gene (Koch et al. 2005; Vijverberg and Bachmann 1999), no difference was found in *trn*F (pseudo)gene copy number in *Melanocrommyum.*

Phylogenetic analyses

The phylogenetic analyses of the *trn*L-*trn*F sequences of 440 individuals resulted in unresolved tree topologies with high number of zero-length branches. Multiple taxa with zero-length branches at the basal nodes of the clades sharing ancestral sequence types are a typical feature of the entire tree. A part of the neighbor joining tree of 440 individuals, which exemplifies the zero-length branches, is presented in Fig. 5. As up to 89 individuals (representing up to 15 species) are sharing one haplotype, thorough phylogenetic analyses of the data-matrix comprising 74 unique haplotypes (each haplotype represented by a single sequence) plus two sequences of outgroup taxa, was carried out as described in Materials and Methods (section 3.2). The strict consensus tree of 80,000 most parsimonious trees (89 steps in length; consistency index = 0.888; retention index = 0.929) was poorly resolved (data not shown). The Bayesian tree using the 74 haplotypes was also poorly resolved, with most groups assembled along a large polytomy. Although Bayesian analyses recovered more species groups (6 lineages) compared to parsimony analyses (3 lineages), statistical support for the groups was mainly low. However, the monophyly of the subgenus got high statistical support (1.0 pp and 97% bs) in both analyses. The Bayesian tree with branch

lengths, bootstrap support from MP, posterior probabilities from BI, and depicting the groups also present in a strict consensus tree of MP is shown in Fig. 6.

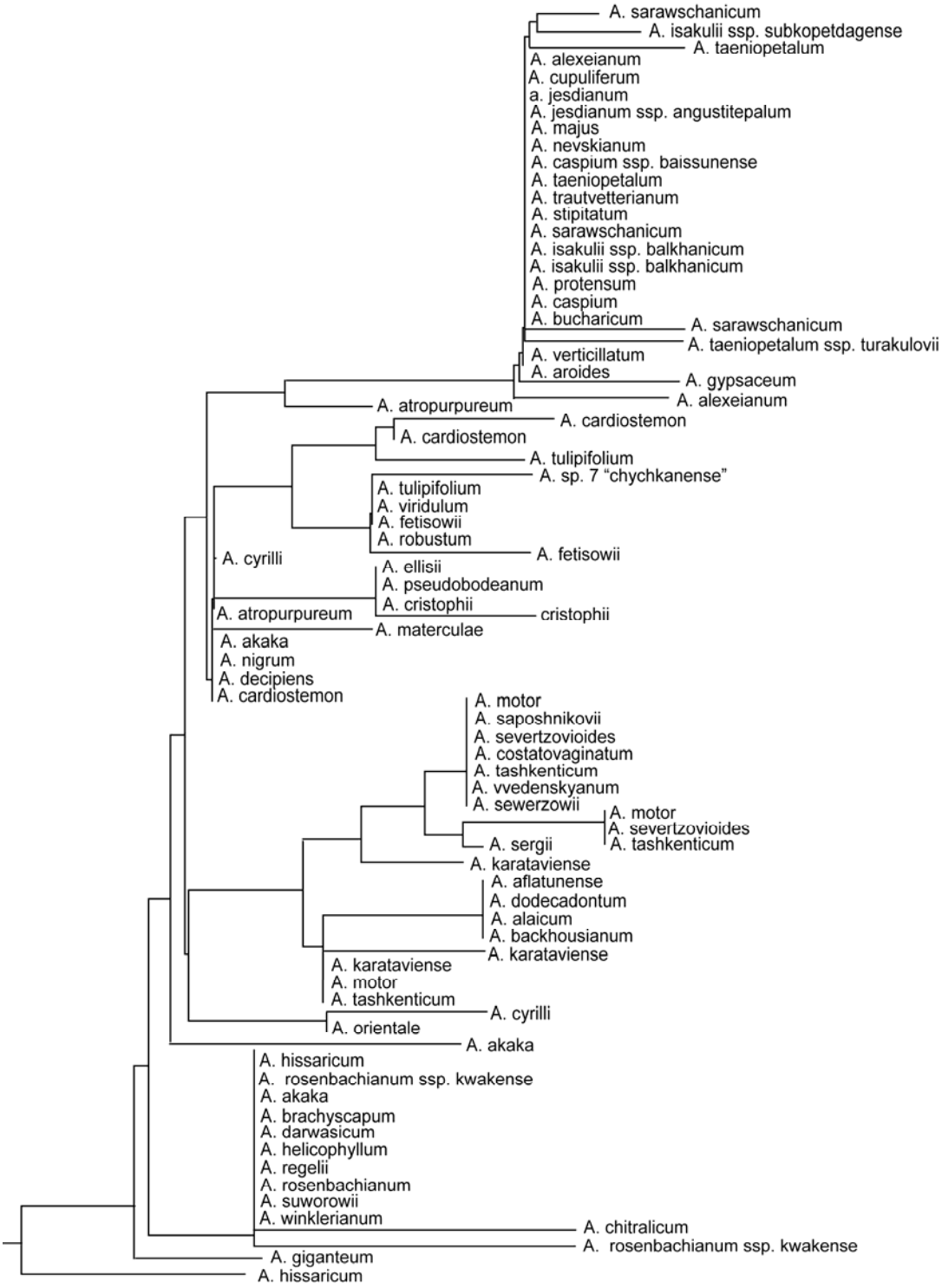


Fig. 5. A part of the Neighbor-Joining tree of *trn*L-*trn*F sequences of 440 *Melanocrommyum* individuals calculated using K81uf+I distances. Zero-length branches at the base of the nodes of this tree correspond to the sequences (haplotypes) which are inferred as ancestral haplotypes (occupying the interior positions within the lineages) by TCS network.

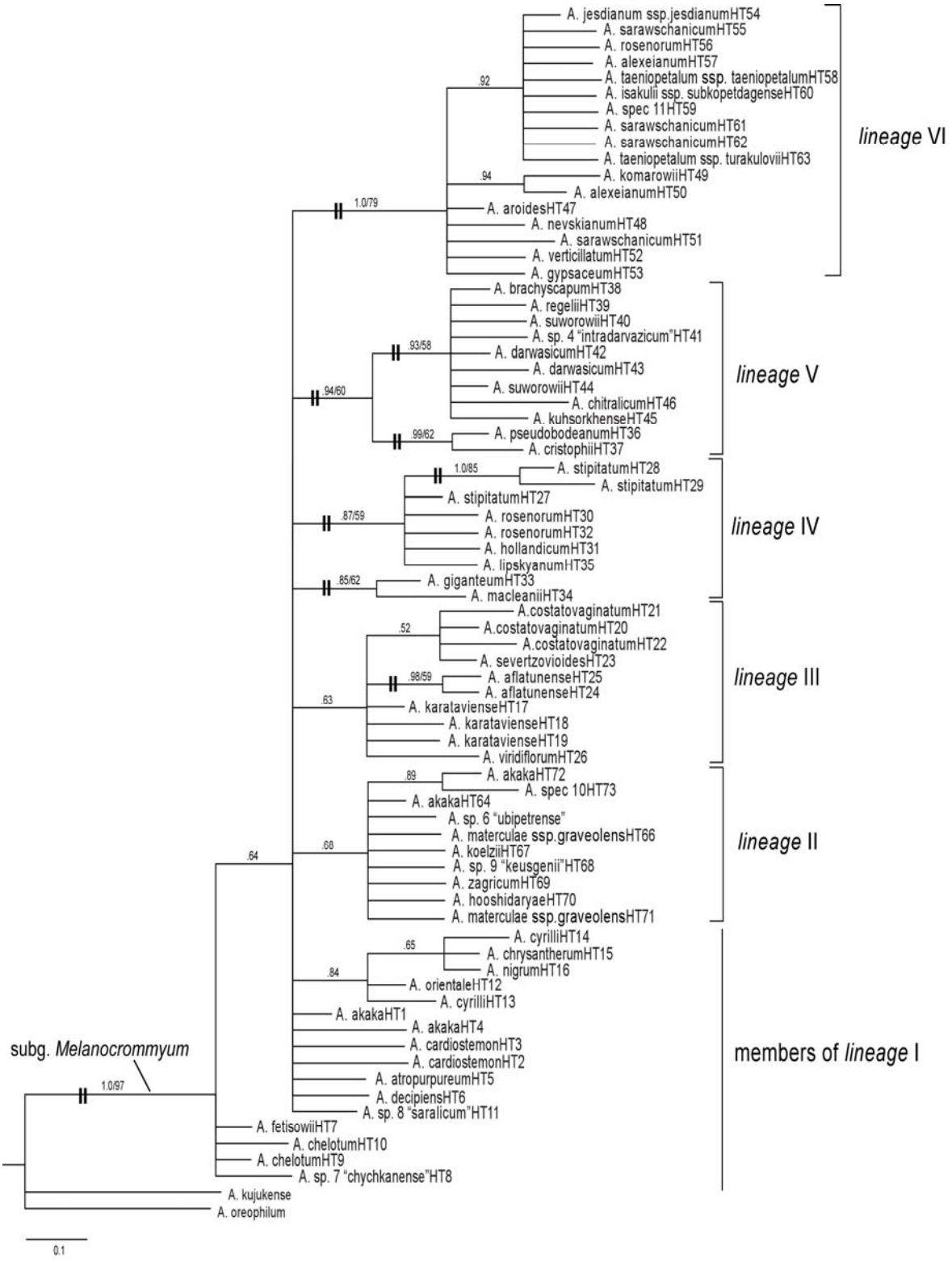


Fig. 6. Bayesian tree of 74 chloroplast haplotypes and two outgroup species. The black bars show the nodes, which also exist in the strict consensus tree of MP. The numbers above branches depict the posterior probabilities from BI and bootstrap support from MP. The lineage numbering corresponds to the lineages in TCS network represented in Fig. 7.

Statistical parsimony network

The network-building algorithm implemented in TCS calculated a 95% parsimonious connection limit of 10 mutational steps. The network resulting from the two-step procedure is complex, but relatively well-resolved (Fig. 7). Only a few loops indicative of homoplasy are present. One of the closed loops concerns the connection of the *Melanocrommyum* haplotype network to the outgroup taxa. The more likely root of the network according to criteria given by Crandall and Templeton (1993) is shown with a larger asterisk in Fig. 7, while the less probable root is depicted by a small asterisk. Some closed loops could be resolved according to the above mentioned assumptions (Crandall and Templeton 1993). These less probable ways of chloroplast sequence evolution are shown with dashed gray lines in Fig. 7. Three positions in the TCS network revealed closed loops caused by single homoplastic alignment positions that could not be resolved unambiguously.

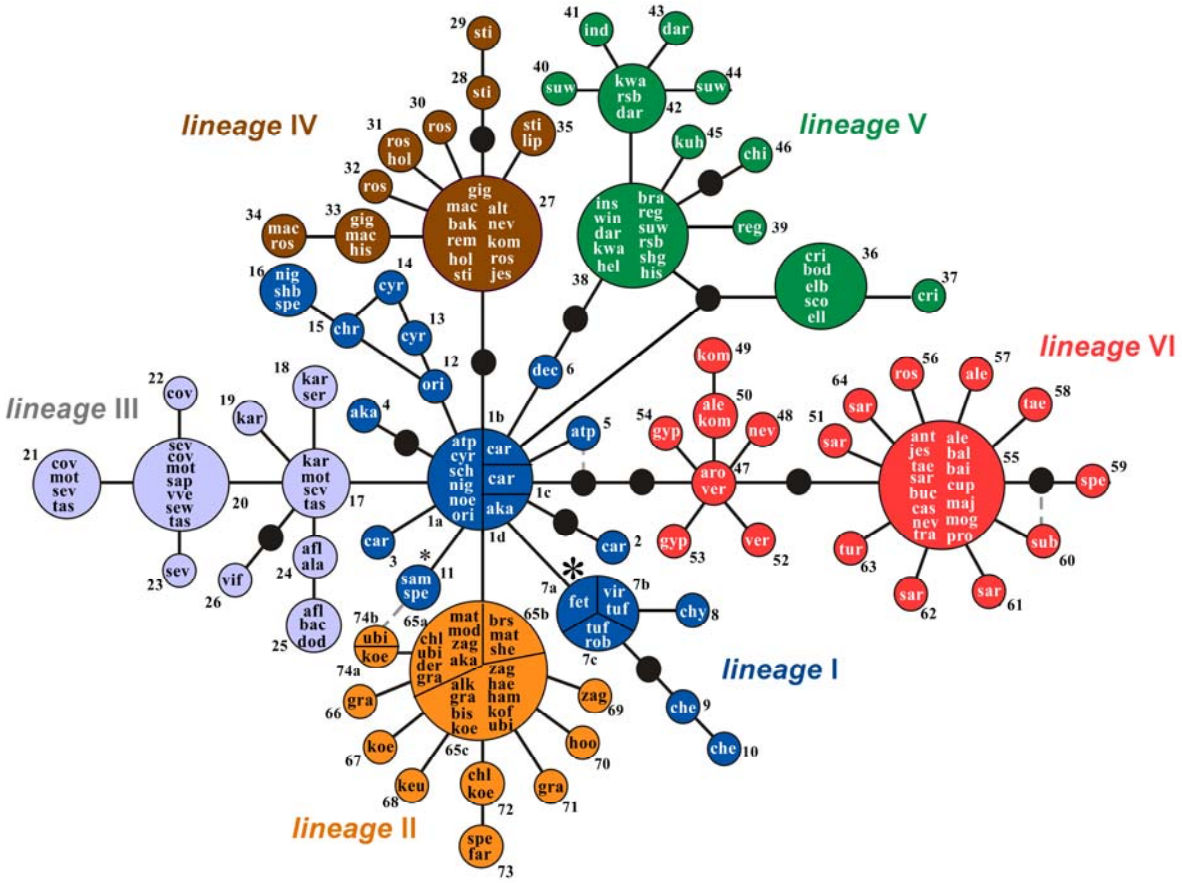


Fig. 7. Statistical parsimony network of 80 chloroplast haplotypes found in subgenus *Melanocrommyum*. Circle sizes correspond to the number of taxa possessing the haplotype. Taxon names are abbreviated by the first three letters. Asterisks depict possible connections to the outgroup taxa (root), where the larger asterisk represents the more probable connection. Inferred haplotypes (not present in the data set) are depicted as black dots. Dashed grey lines represent less probable connections according to criteria suggested by Crandall and Templeton (1993). The haplotypes of the backbone network (1-74) are arbitrarily numbered. Extensions with lower case letters (a to c) indicate haplotype division resulting from variation at two mononucleotide repeats and one highly variable indel.

The haplotypes with interior positions in the network connecting other derived haplotypes, are assumed to be ancestral haplotypes giving rise to descendant types. Interior haplotypes and their descendants were combined into "lineages". Within the network derived from 74 chloroplast haplotypes of subgenus *Melanocrommyum*, 6 lineages were defined; the internal lineage I, where also the outgroup connection was inferred by the TCS, corresponds to the ancient (ancestral) lineage. Inferred but non-observed haplotypes (missing intermediate sequences) were much more frequent at the base of the network. Seven out of fourteen inferred missing intermediates are restricted to the oldest lineage 1 (Fig. 7) near the root of the network.

The topology of the haplotype network was congruent with the Bayesian tree of 74 chloroplast haplotypes. All major lineages present in the network (II, III, IV, V, VI) are also present in the Bayesian tree (Fig. 7 and 6, respectively). The only exception is lineage I according to the TCS network, which was not recovered in the Bayesian tree (Fig. 6), as some members of this lineage are arranged at the base of the tree and along the major polytomy. Thus, relationships among the haplotypes, forming the ancient lineage I in statistical parsimony network, are unresolved in the Bayesian tree (see discussion in this chapter).

Phylogenetic and taxonomic relationships of chloroplast haplotypes

The conspicuous result of *trn*L-*trn*F sequence analyses of subgenus *Melanocrommyum*, is that usually in closely related species (based on ITS data and partly morphology) the chloroplast haplotype distribution is not species-specific. Often, more than one haplotypes was found in a single species, whereas several closely related species often share the same haplotype. For instance, five haplotypes were found in *A. sarawschanicum* (lineage VI), four haplotypes in *A. severtzovioides* (lineage III) and *A. akaka* (lineage I), and 3 haplotypes in *A. cardiostemon* (lineage 1). It is worth mentioning that the phenomenon (having several haplotypes within one species) is not restricted to the above mentioned taxa. The detailed information concerning each species can be found in Table 2. Mostly the different haplotypes present in single species are quite closely related to each other and belong to the same lineage. However, in several species occurred individuals possessing haplotypes which are many mutational steps away from each other and belong to two different lineages. These are single individuals of *A. nevskianum, A. hissaricum,* and *A.rosenorum*. *Allium nevskianum* accession number 5451 according to chloroplast data is assigned to the lineage IV, whereas all other A. *nevskianum* individuals bear haplotypes of the lineage VI. The same is true for *A. hissaricum* 2947, which different from conspecific individuals (HT33, lineage IV) possesses HT38, belonging to lineage V. *Allium rosenorum* 5081 bears a haplotype belonging to lineage VI, whereas all remainder 26 analyzed individuals of this taxon have chloroplast types of lineage IV.

On the other hand, the accessions of the closely related species [based on nrDNA ITS sequences and morphology (Chapter 2)] *A. severtzovioides, A. costatovaginatum, A. motor, A. tashkenticum, A. saposhnikovii, A. vvedenskyanum,* and *A. sewerzowii* share the same

haplotype (HT20), and several of the above mentioned taxa share two more haplotypes (HT17 and HT21). However, species-specific haplotypes are also found in *A. costatovaginatum* (HT22) and in *A. severtzovioides* (HT23). Similarly, members of a circum-Mediterranean species group comprising *A. nigrum, A. atropurpureum, A. cyrilli, A. orientale, A. schubertii*, share chloroplast HT1, and most of the taxa of this group in addition possess species-specific haplotypes (for details see Fig. 7).

3.4 Discussion

Network construction and comparison of different analyses methods

Although only few data are yet available on absolute mutation rates of chloroplast microsatellites (Provan et al. 1999), the assumption that the number of the residues within a mononucleotide repeat is more likely to be homoplasious than other types of cpDNA sequence mutations seems more probable. The study of Collevatti et al. (2003) on the endemic Brazilian tree *Caryocar brasiliense* is one of such examples. The authors found strong topological differences between a median-joining network based on 21 cpDNA microsatellite haplotypes and a network based on point mutations and larger indels in the *trn*T-*trn*F spacer region. The authors concluded that the observed incongruence was most probably caused by several independent mutations (i.e. homoplasy) of the microsatellites. Therefore, I followed a two-step procedure of network construction similar to Bänfer et al. (2006), in order to account for the presumably different mutation rates at microsatellite loci vs. base substitutions or larger indels. The two-step procedure of network construction used here allowed to more reliably exploit sequence information at fast mutating microsatellite loci, without introducing many homoplastic sequence positions into the primary data set.

In this study I used two approaches (tree-based and network-based) to analyse the relationships among the haplotype sequences of subgenus *Melanocrommyum*. The maximum parsimony analysis resulted mainly unresolved trees (black double bars along the branches of the Bayesian tree in Fig. 6 show the groups present in the strict consensus tree), whereas low resolution and low statistical support are common features of Bayesian tree, although more groups are recovered in this analysis. The genealogical (network-based) approach is designed to deal with intraspecific molecular data. As within one species descendant alleles coexist with persistent ancestors, intraspecific genealogies are often multifurcating and cannot be represented by a bifurcating tree (Posada and Crandal 2001). However, the same situation (i.e. coexistence of older haplotypes along with their multiple descendants) is not rare also at inter-species level (Jakob and Blattner 2006). In addition, the persistence of ancestral chloroplast alleles through several speciation events was shown in several plant groups including *Arabis* (Dobeš et al. 2004) and *Hordeum* (Jakob and Blattner 2006). The fact that the well-resolved chloroplast haplotype network of subgenus *Melanocrommyum* reveals multifurcating relationships, indicates that ancient chloroplast types, co-occurring with their multiple descendants, are also common in subgenus *Melanocrommyum*. This can easily explain why the oldest chloroplast lineage (lineage I in

the network based on *trn*L-*trn*F region, which unites ancestral sequences) is not recovered in tree based analyses methods, which cannot deal with situations where ancestral sequence types coexist with their derivatives. Thus, network-based approaches can better describe the relationships among haplotypes and be more informative even at the species level in subgenus *Melanocrommyum*.

Overall congruence with the nuclear rDNA ITS phylogeny

The chloroplast data (*trn*L-*trn*F sequences) support the general outcome of nuclear rDNA ITS analysis. The groups found in a Bayesian tree and statistical parsimony network of 74 haplotypes (lineages I-VI) generally agree with the groups inferred from the phylogenetic analyses of the ITS region, and both nuclear and chloroplast phylogenies strongly contradict morphology-based taxonomic sections. The species clustering together on the ITS tree share generally the same chloroplast lineages (Fig. 8).

One of the obvious examples is section *Regeloprason*, its members are scattered along the ITS tree. The split of the section is confirmed by chloroplast data presented here, as members of the section belong to three well-separated chloroplast lineages (IV, V, VI) in both statistical parsimony network and Bayesian tree.

Another conspicuous example, where *trn*L-*trn*F data are in full accord with ITS results, is the group G in ITS tree (Fig 8, and Fig. 3b in Chapter 2) uniting several morphologically clearly diverse taxa belonging to several taxonomic sections. Surprisingly, all the taxa from this group exclusively share closely related haplotypes of lineage V.

However, despite the overall congruence of *trn*L-*trn*F with ITS data at higher level relationships (i.e. level of lineages) within the subgenus, there are also few cases of incongruence. Based on ITS sequences *A. jesdianum* (subsp. *jesdianum* and subsp. *angustitepalum*) belongs to group K (cluster 7 in ITS phylogenetic tree). All other members of this ITS group possess *trn*L-*trn*F haplotypes of lineage IV, but *A. jesdianum* subsp. *angustitepalum* accessions show chloroplast types of lineage VI. Another case of incongruence between ITS and chloroplast data is *A.winklerianum*, which according to ITS data belongs to group J, uniting species predominantly with chloroplast types belonging to lineage IV. However, all accessions of *A.winklerianum* bear haplotype 38 from lineage V. As the "wrong" haplotypes are derived haplotypes and do not belong to an ancestral (old) lineage the above mentioned conflict between chloroplast and nuclear data could not be explained by ancestral polymorphism and subsequent incomplete lineage sorting. In these cases the incongruences between *trn*L-*trn*F and nuclear ITS results could be more likely caused by chloroplast capture after hybridization.

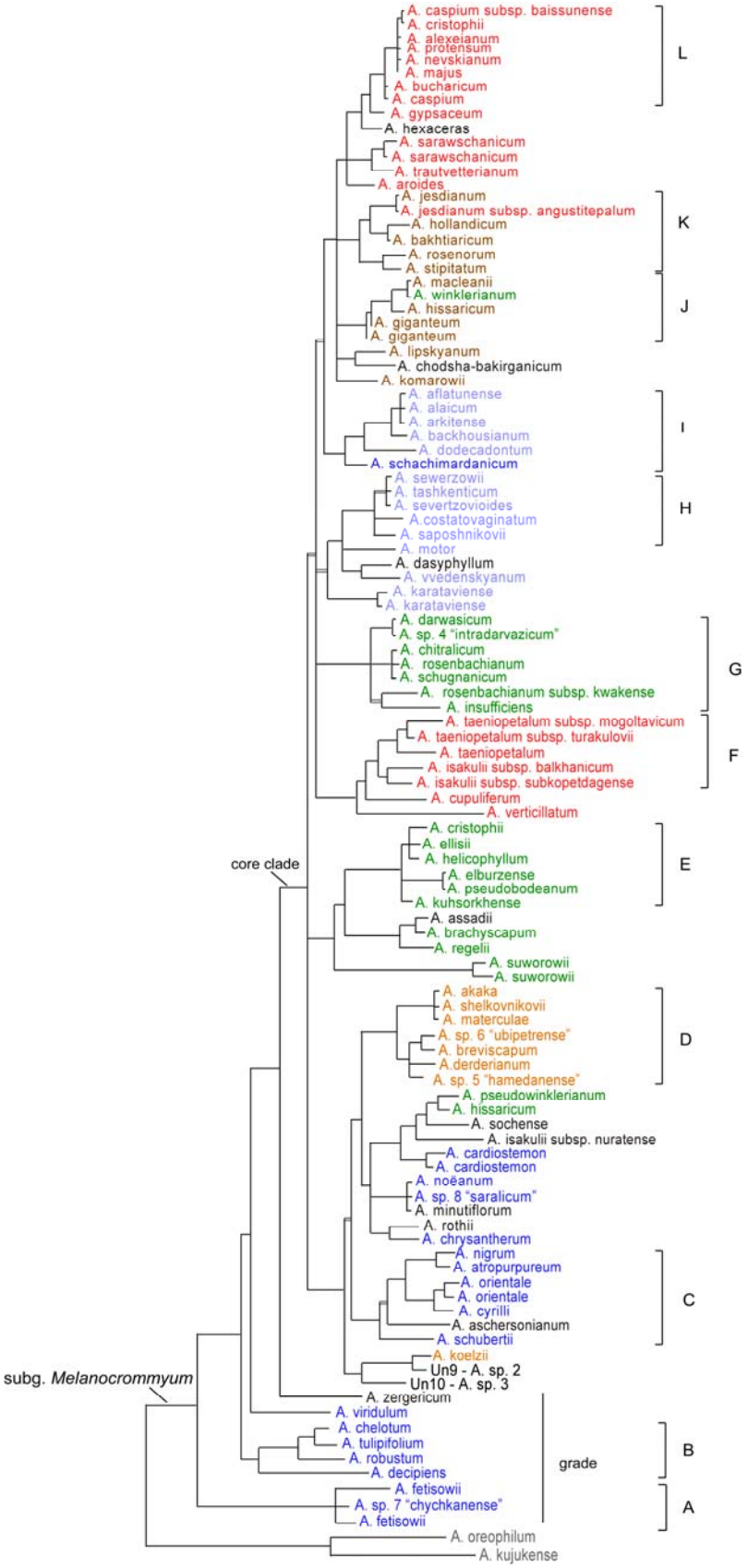


Fig. 8. A schematic representation of ITS phylogenetic tree from the chapter 2. The color codes correspond to the colors of respective chloroplast lineages from Fig. 7, and thus, show the distribution of chloroplast lineages among the "ITS groups". The taxa for which the *trn*L-*trn*F region could not be amplified are given in black.

Because all sequenced individuals of *A. jesdianum* subsp. *angustitepalum* and *A.winklerianum* possess the "wrong" chloroplast type, these taxa seem to be of hybrid origin. While in *A. jesdianum* subsp. *angustitepalum* the sampling is good (10 individuals sequenced) the relatively few number of analyzed individuals in *A.winklerianum* (3 individuals) does not allow to draw safe conclusion about a hybrid origin. In contrast, in *A. nevskianum, A. hissaricum,* and *A. rosenorum* only single individuals (*A. nevskianum* 5451, *A. hissaricum* 2947, and *A. rosenorum* 5081) are introgressed.

Phylogenetic implications, lineage sorting and hybridisation

Despite the fact that chloroplast haplotype distribution does not strictly follow species circumscription and several haplotypes were found within single species, sharing distantly related haplotypes (belonging to different lineages) was not a common phenomenon in subgenus *Melanocrommyum*. The incongruence between the species boundaries and the genealogy of their cpDNAs is further illustrated by sharing of a single chloroplast haplotype among up to fifteen species. This finding is in accord with other studies in different genera, which included multiple individuals per species (e.g. Dobeš et al 2004; Jakob and Blattner 2006; Bänfer et al. 2006). However, an association of chloroplast haplotypes with geographically circumscribed regions rather than with taxonomic boundaries, a phenomenon observed in numerous plant taxa, including, e.g. oaks (e.g. Dumolin-Lapègue et al. 1997), eucalypts (McKinnon et al. 2001), junipers (Terry et al. 2000), birches (Palme et al. 2004), *Lithocarpus* (Cannon and Manos 2003), and *Macaranga* (Bänfer et al. 2006), was not found in *Melanocrommyum*. This finding together with extremely rare cases of obvious chloroplast capture (e.g. *A. nevskianum* 5451 and *A. hissaricum* 2947), supports that hybridization is not frequent in the subgenus, a result which is consistent with the outcome of the extensive analyses of the cloned and directly sequenced nrDNA ITS region (chapter 2). However, a few cases of incongruence between chloroplast and nuclear data, like *A. nevskianum* 5451 and *A. hissaricum* 2947 could better be explained by chloroplast capture.

Incongruence between phylogenetic data sets derived from the different genomes in plants was mostly explained by cryptic hybridization and introgression (Rieseberg and Soltis 1991; Rieseberg et al. 1996). However, there is increasing evidence that incomplete or differential lineage sorting, i.e. the persistence of ancestral polymorphisms through speciation events, could also contribute to phylogenetic incongruence (Mason-Gamer et al. 1995; Wendel and Doyle 1998; Comes and Abbott 2001; Linder and Rieseberg 2004). As a result, it becomes often impossible to distinguish between these two processes in phylogenetic analyses (Wendel and Doyle 1998; Comes and Abbott 2001). Considering (i) the overall congruence of nuclear and chloroplast data in subgenus *Melanocrommyum*, (ii) major discrepancies occurring at closely related species level, without indication of current gene flow (no within-individual ITS polymorphism detected), and (iii) that these discrepancies are caused by sharing of haplotypes which are only few mutational steps away from each other, could be better explained by a lineage sorting scenario. However,

ancient hybridization followed by homogenization of rDNA loci to the maternal type can not be completely ruled out.

The fact that ancestral haplotypes (HT7, HT1, and HT11) can still be found within the subgenus support earlier observations that chloroplast types could survive several speciation processes, which can impede phylogenetic inference of species-level relationships when chloroplast data are used.

Reasons for current chloroplast haplotype distribution

According to the assumptions of coalescent theory, older alleles should prevail in populations and be characterized by a higher number of descending lineages and a geographically wider distribution than younger alleles (Crandall and Templeton 1993; Posada and Crandall 2001). In interspecific phylogenies, however, these parameters depend much on sample design and differences in the history of the taxa. However, the chloroplast haplotype distribution in *Melanocrommyum* is in accord with this assumption of coalescent theory. In each lineage older (ancestral) haplotypes are more frequently found, more widespread and gave rise to multiple lineages. Most inferred haplotypes, which either are extinct or are not recovered in sampled individuals, are mostly near the base of the network. A similar result was obtained in a phylogeographic study of three *Arabis* species by Dobeš et al. (2004). This outcome is also in agreement with the observation that in contrast to intraspecies genealogies older haplotypes often become rarer due to stochastic lineage sorting processes at the higher taxonomic levels (Jakob and Blattner 2006). The fact that in all lineages the ancestral haplotypes are prevailing and widely distributed coupled with a very low number of missing intermediate sequences would allow to speculate that *Melanocrommyum* taxa have not experienced severe bottlenecks and have relatively constant or even growing population sizes. This would fit with the inferred recent radiations within the subgenus (Chapter 2; Gurushidze et al. 2008), and a relatively young age estimated via molecular dating (Chapter 4; Gurushidze et al. in prep.) for the subgenus *Melanocrommyum*, which is probably still expanding its distribution range.

*Trn*F duplication

Structural mutations in chloroplast DNA concern mostly indels of up to 10 bp and, to a lesser extent, larger mutations (up to 1200 bp) in noncoding regions (Palmer 1991). Large rearranged cpDNAs are present in only a few plant groups, including conifers (Strauss et al. 1988) and two tribes of legumes that lack the inverted repeat (IR; Palmer et al. 1987), grasses that contain three inversions (Quigley and Weil 1985; Doyle et al. 1992), and some Geraniaceae that show a largely extended IR (Palmer et al. 1987). A summary of structural mutations in the chloroplast genome is provided by Vijverberg and Bachmann (1999). These microstructural changes have been shown to be extremely useful even in resolving deep phylogenies (Graham et al. 2000; Löhne and Borsch 2005) and have been analyzed in more detail in the chloroplast genome of *Silene* (Ingvarsson et al. 2003). Structural

mutations such as gene duplications among higher plant plastomes are relatively rarely described. Those examples involve tRNA genes (e.g., Hipkens et al. 1995; Vijverberg and Bachmann 1999; Drábkova et al. 2004, Pirie et al. 2007; Koch at al. 2005, 2007), *rpl2* and *rpl23* (Bowman et al. 1988), *psb*A (Lidholm et al. 1991), and *psa*M (Wakasugi et al. 1994).

The *trn*F pseudogenes described in several Brassicaceae genera show high variability, as well as differences in pseudogene copy number and are shown to have been originated several times independently in crucifers (Koch et al. 2007).

The divergent copy directly adjacent to the 5' end of *trn*F gene of *Melanocrommyum* taxa, showing length differences together with several transversions, is probably a pseudogene, which is further supported by the fact that the *trn*F copy of *Melanocrommyum,* directly adjacent to the assumed pseudogene, is identical to the *trn*F gene of *Triticum, Nicotiana,* and *Lactuca.* The fact that neither differences in copy number, nor the high variability of the *trn*F pseudogenes was detected in subgenus *Melanocrommyum* could be caused by a relatively young age of the duplication. However, as I cannot speculate about the mechanism causing this duplication, the constant copy number in the subgenus could also have structural reasons limiting additional duplications. It is unclear at the moment, whether the *trn*F duplication is restricted to the subgenus and related subgenera (as both outgroup species are also bearing the duplication), or present in the entire genus *Allium*. Thus, the available data do not allow to draw any conclusion about pseudogene evolution and the age of the observed duplication, which would require further work extending sequencing of the *trn*F gene to other taxonomic groups of *Allium* and maybe the inclusion of other members of Alliaceae.

4. Dating diversification events in the genus *Allium* and its subgenus *Melanocrommyum* is impeded by low *rbc*L and extremely high nrDNA ITS substitution rates

4.1 Introduction

The genus *Allium* L. is widely distributed over the warm-temperate and temperate zones of the northern hemisphere (Friesen et al. 2006). A region of especially high species diversity stretches from the Mediterranean basin to Central Asia and Pakistan. A second smaller diversity center is in western North America. Molecular phylogenetic research on *Allium* has been quite active over the past years. The molecular investigations focused on the infrageneric differentiation and classification of the entire genus *Allium* (Linne von Berg et al. 1996; Mes et al. 1997; Dubouzet and Shinoda 1999; Friesen et al. 2006), the phylogenetic position of North American (mainly Californian) taxa and their adaptation to serpentine soils (Nguyen et al. 2008) or phylogenetic relationships of particular subgenera (Dubouzet et al. 1997; Gurushidze et al. 2008). Other studies dealt with wild relatives of common onion, with the emphasis to reveal the progenitor of *A. cepa* (Havey 1992; Lilly and Havey 2001; van Raamsdonk et al. 2003; Gurushidze et al. 2007 and references therein).

A recent phylogenetic study by Friesen et al. (2006) includes both New and Old World members of subgenus *Amerallium*, the only subgenus with North American – Eurasian disjunct distribution. Based on phylogenetic analyses of rDNA ITS sequences *Amerallium* turned out to be a monophyletic group (Friesen et al. 2006). Furthermore, the subgenus is divided into two well supported sister clades, and New World and Old World *Amerallium* species are each monophyletic (although sampling was far from complete). However, the main biogeographic question about the relationships of the Old World and North American taxa, and timing the major diversification events in *Allium* was never addressed using molecular markers.

Based on the current distribution of *Allium* species and high species numbers in both continents, it was hypothesized: (i) that *Allium* is of an early Tertiary origin (Hanelt 1992), (ii) its existence in North America dates back at least to the middle Tertiary (Raven and Axelrod 1978), and (iii) that migrations between Europe and America occurred using Tertiary land bridges (Hanelt 1992). Although the majority of *Allium* species grow in open, rather dry sites in arid or moderately humid climates, species have adapted to other ecological niches (different types of forests, subalpine pastures, and moist subalpine and alpine grasslands). The species of subgenus *Melanocrommyum*, however, are exclusively adapted to arid conditions. Based on synapomorphies and specializations as well as due to distinct ecological and geographical restrictions the subgenus was hypothesized to be a phylogenetically derived group, which rapidly diversified during the late Tertiary when after closing of the Tethys Sea enormous regions could be newly colonized (Hanelt 1992). Phylogenetic analyses of nrITS and chloroplast sequences support the idea of several radiations occurring at different times within the subgenus (Chapter 2).

Since the beginning of the Tertiary period, Earth's climate has undergone significant and complex changes (reviewed by Zachos et al. 2001). The most pronounced warming trend occurred early in the Cenozoic, from the middle Paleocene to early Eocene, reaching the peak in the Early Eocene Climatic Optimum (EECO; 52 to 50 Myr ago). With the onset of glaciation and expanding massive ice-sheets in Antarctica (~34 Myr ago), the northern continents cooled and dried (Bowen 2007). This global climate change in conjunction with tectonic processes, (uplift of the Tibetan plateau and retreat of the Paratethys associated with the collision of India and Asia) resulted in aridification and the intensification of the monsoons in Asia (An et al. 2001; Dupont-Nivet et al. 2007). This glaciation was followed by climate warming, peaking with the Middle Miocene Climatic Optimum (MMCO; 17 to 15 Myr ago). The next trend was cooling of the climate and small-scale ice-sheet expansion on west-Antarctica and in the Arctic (10 and 6 Myr ago), while the Early Pliocene is marked by a subtle warming trend until the Late Pliocene. In conjunction to these climatic changes, since the middle Neogene [middle Miocene (~15-11 Myr) – Pliocene (~5.3-1.8 Myr)] the central part of Eurasia (Central Asia) has progressively become drier. The Northern Hemisphere Glaciation (NHG) started in the late Pliocene (Zachos et al. 2001) and resulted in the establishment of the Arctic ice cap before the beginning of the Quaternary (Webb and Bartlein 1992).

Considering the above mentioned major climatic changes and aridificaton events in the northern hemisphere during middle to upper Tertiary raise the following questions: 1. Does the age of the genus *Allium* fit hypothesized early Tertiary origin? 2. Is its disjunct distribution pattern Laurasian, or the result of later colonization(s)? 3. Is subgenus *Melanocrommyum* a relatively young arid-adapted lineage? 4. When did the majority of diversification within the subgenus occur, and is this diversification correlated with the episodes of aridification during the Cenozoic? 5. Is there any evidence of increased species-level diversity with the onset of semiarid and arid conditions since the late Miocene?

Recent years have witnessed how molecular phylogenetics contributed to answer long-standing biogeographic questions. For instance, age estimation from molecular sequences has emerged as a powerful tool for inferring when a plant lineage arrived in a particular area (Davis et al. 2002; Won and Renner 2006), or the reasons of the disjunct distribution patterns between eastern Asia and eastern North America, which have fascinated botanists and biogeographers since the Linnaean era (Xiang et al. 2000; Nie et al 2006; for review see Wen 1999, 2001). Whereas other studies revealed the rapid radiations in diverse array of organisms including ferns (Schneider et al. 2004), neotropical rain forest trees (Richardson et al. 2001), or in drought adapted plant lineages having spread in open habitats, such as desert, shrubland, and fynbos in response to aridification (Goldblatt 2002; Moore and Jansen 2006).

In the present study I address the questions about origin and disjunct distribution of *Allium* with a special focus on major diversification processes within subgenus *Melanocrommyum.* To answer such questions require both a well-supported phylogeny as well as reasonable calibration points to date diversification events in *Melanocrommyum.* In

many molecular-based studies that apply dates to the nodes of phylogenetic trees, well-dated ingroup fossils with reasonably clear affinities to modern taxa are used as constraints on the ages of nodes. However, such an approach is impossible for *Allium* because of the complete absence of a known fossil for the entire genus. Such a situation is common in herbaceous plant taxa of arid regions, as the conditions required for fossilization are usually not encountered in such areas (Axelrod 1979; Moore and Jansen 2006). Another reason could be the filtering of herbaceous plants (i.e. they are less likely to be preserved), which is probably also the general reason for scarcity of monocot fossils, particularly of the Early and Mid Cretaceous (Anderson and Janssen, in press). Where close relatives have no fossil record, dating approaches require the inclusion of phylogenetically distant outgroups simply because they do have a fossil record (Won and Renner 2006). However, the inclusion of distant outgroups demands to use conserved DNA regions which might not be variable among closely related species whose divergence one wishes to date. The only way to proceed then is to analyze separate data sets and to bridge them by using one of the nodes dated in the first analysis as a calibration point in the second analysis (secondary calibration, e.g. Berry et al. 2004; Zhou et al. 2006). Nevertheless, in the absence of fossils, it is still possible, although less desirable, to use other types of information, such as fossil-based ages from outgroups (unfortunately, there are no fossils also in Alliaceae s. lat. or even Asparagales) as well as known ages of past climatic or geologic events to constrain the taxon ages in the phylogeny. This approach has been used in other studies involving the radiation of Hawaiian silverswords (Baldwin and Sanderson 1998) and in the drought-adapted plant group *Tiquilia* Pers. (Boraginaceae) by Moore and Jansen (2006).

Although phylogenetic research on monocots has been extraordinarily active over the past years, and there is a rather well supported phylogenetic hypothesis for the whole clade which is quite well resolved at the family level (see review in Chase 2004), only a few studies are available about age inferences in monocots. The studies of Janssen and Bremer (2004) and Anderson and Janssen (in press) are probably the best efforts at dating the divergence times in the entire clade. They estimated the ages of major monocot groups based on *rbc*L sequences of 878 genera from 77 families representing all ten orders. The near-complete sampling of Asparagales families enables to use the age of Alliaceae s. lat. as a secondary calibration for the present study.

In this chapter I report the attempt to use the *rbc*L and nrDNA ITS sequences to test if diversification in *Allium* and radiations within subgenus *Melanocrommyum* are in accord with global aridification processes on the northern hemisphere. Because there are no fossils known from the Alliaceae, estimates of divergence times are made with a molecular clock approach from published rates of ITS evolution and from a secondary calibration based on the above mentioned studies. It is worth mentioning that, while ITS turned out useful to circumscribe intrageneric groups and resulted in a revision and new classification of the genus (Friesen et al. 2006), the relationships among major clades were not confidently resolved in the ITS trees because of extremely high genetic distances at the ITS locus (up to 50% within the genus). Therefore, the rationale to use *rbc*L sequences was also to resolve

the deeper nodes in the phylogeny, which would be necessary for dating of diversification events, and would avoide successive bridging via inclusion of *Allium* sequences in an Asparagales *rbc*L matrix.

4.2 Materials and Methods

Sampling

The selection of *Allium* as well as outgroup species for *rbc*L analyses was based on an *Allium* ITS phylogeny and respective taxonomic classification (Friesen et al. 2006), and included 157 individuals covering 122 taxa. Leaf material for DNA extraction was collected either from the *Allium* living collections of the IPK Gatersleben and Osnabrück botanical garden (Germany), or collected in natural populations. For several species living plants were not available and herbarium specimens were used. Outgroup sequences were obtained from GenBank. Accession numbers, classification (subgenera), and the provenances of the plant material, sequences of which were included in a final data-matrix, are listed in appendix Table 2. Voucher specimens were deposited in the herbarium of the IPK Gatersleben (GAT) or the University of Osnabrück (OSBU).

PCR and sequencing

For the present study *rbc*L was amplified and sequenced in 157 individuals of 122 *Allium* species. Total genomic DNA-isolation was carried out with the DNeasy Plant Kit (Qiagen) according to the instructions of the manufacturer. DNA concentrations were estimated on 1% agarose gels stained with ethidium bromide. The primer sequences used to amplify and sequence *rbc*L are listed in Table 3. Amplification from silica-dried material was carried out in one reaction using primers: *rbc*L-1F and *rbc*L-1384R. For herbarium material the gene was amplified in two overlapping pieces using two internal primers (see Table 3) in addition to the two mentioned above. Amplification primers 706R and 1384R were designed for the present study because commonly used primers 724R and 1460R failed to work in *Allium*. Additionally I designed sequencing primer *rbc*L-615F because of differences in primer binding site within *Allium* taxa. PCR was performed in 50 µl reaction volume containing 1 U *Taq* DNA polymerase (Qiagen), the supplied reaction buffer, 0.2 µM of each dNTP, 50 pmol of each primer and approximately 20 ng of total DNA. PCR was carried out in a GeneAmp 9700 PCR System (PE Biosystems) with the following cycling program: 95°C – 3 min, 38 cycles of [95°C – 30 sec, 56°C – 1 min, 68°C – 1.5 min], and a post-treatment at 68°C for 8 min. The amplicons of approximately 1400 base pairs (bp) in length were purified using Nucleofast 96 PCR plates (Macherey-Nagel) and resuspended in 35 µl TE buffer. Direct sequencing was performed with the dye-terminator technology on an ABI 3730xl automatic DNA sequencer (Applied Biosystems). Forward and reverse sequences were edited and combined in single consensus sequences in CromasPro version 1.41 (Technelysium Pty Ltd), and aligned manually. When different

accessions of one species possessed identical sequences only one sequence was included in a final data-matrix.

Table 3. PCR and sequencing primers for *rbc*L.

Primer name	sequence
*rbc*L-1F	ATGTCACCACAAACAGAAACTAAAGC
*rbc*L-636F	GCGTTGGAGAGATCGTTTCT
*rbc*L-1384R[1]	GAACTCGAATTTGATCTCTTTCC
*rbc*L-706R[1]	AGTTGCATTCAAGTAATGCCCTTT
*rbc*L-615F[1]	TGAACTCCCAGCCTTTTATGCG

[1] *Allium* specific primers designed for this study, others are the same used in Chase et al. (1993).

Data sets and phylogenetic analyses

The Asparagales *rbc*L data matrix was the same used in the study of Janssen and Bremer (2004). Each genus is represented by one sequence in the data matrix, and all sequences are available from GenBank (for detailed information see Janssen and Bremer 2004). The reason to use Asparagales data set was to have several calibration points to obtain the crown age for Alliaceae *sensu stricto* (s. str.) and to date splits within *Allium.* Therefore in this study I used the same topological constraints described by the authors, calibration points for all nodes in Asparagales were used according to the study of Anderson and Janssen (in press), which reports stem ages for the families as resulted from both methods, nonparametric rate smoothing (NPRS) from the study of Janssen and Bremer (2004), and penalized likelihood (PL) from Anderson and Janssen (in press). Thus, the Asparagales data matrix including *rbc*L sequences from representative *Allium* taxa was used to obtain the Alliaceae s. str. and *Allium* crown ages. These dates resulting from the Asparagales *rbc*L tree were used to calibrate the splits in the *Allium* ITS phylogenetic tree, which was obtained by Friesen et al. (2006), as *rbc*L turned out extremely conserved and therefore not useful to date diversification events in *Allium*. At the next step the age for the respective node from the *Allium* ITS tree was taken to calibrate the splits in *Melanocrommyum*, using the ITS tree of *Melanocrommyum* from Gurushidze et al. (2008; Chapter 2).

Phylogenetic analyses (phenetic and parsimony) of *Allium rbc*L sequences were performed in PAUP* 4.0b10 (Swofford 2002), Bayesian inference (BI) in MRBAYES 3.1 (Ronquist and Huelsenbeck 2003). The K81uf+I+G model was chosen as the best fitted model by the Akaike information criterion after testing of different models of sequence evolution was performed in MODELTEST 3.7 (Posada and Crandall 1998). The parsimony analysis was performed in two steps (Blattner 2004), with the same settings as described in Chapter 2. For BI eight chains were run for 3 million generations under the respective model of sequence evolution, sampling a tree every 100 generations. The posterior

probabilities were calculated after discarding the initial non-stationary (burnin = 15000) trees.

The number of substitutions at synonymous (Ks) and non-synonymous sites (Ka) of *rbc*L was calculated using MEGA version 4 (Tamura et al. 2007). Analyses included the base positions 61 to 1305 of *rbc*L.

Estimation of divergence times

As noted above, because there are no fossils in Alliaceae, I used two approaches to calibrate the molecular clock in the genus *Allium* and subgenus *Melanocrommyum*. First, a range of mutation rates of ITS sequences from herbaceous plant lineages was taken from the literature. Independently calibrated and published ITS substitution rates across herbaceous angiosperms range from 1.72×10^{-9} substitutions per site per year in *Saxifraga* estimated by Richardson and colleagues (2001) from data in Vargas et al. (1999) to 8.34×10^{-9} substitutions per site per year in *Soldanella* (Zhang et al. 2001). For a detailed survey see Kay et al. (2006). Second, the age for Alliaceae s. lat. (Janssen and Bremer 2004; Anderson and Janssen, in press) was used to estimate the age of Alliaceae s. str. and the genus *Allium.* Besides good sampling of all orders, the studies cover also most genera from all families recently transferred (APG II 2003) to Alliaceae s. lat. (Alliaceae s. str., Agapanthaceae, and Amaryllidaceae, and Xanthorrhoeaceae s. lat. including Asphodelaceae, Hemerocallidaceae, and Xanthorrhoeaceae s. str.). Although the resolution within *Allium* was low for *rbc*L, the *rbc*L sequences were useful to obtain the ages for Alliaceae s. str. and the genus *Allium*. For that purpose, *Allium rbc*L sequences representing all taxonomic groups (15 subgenera) were included in the Asparagales data-matrix, and the resulting tree (the constraints for the families were the same as described in Janssen and Bremer 2004) was calibrated using the reported stem node ages from the above mentioned monocot dating studies (Janssen and Bremer 2004; Anderson and Janssen, in press). As *rbc*L is too conserved to resolve the relationships among the major lineages of *Allium*, the only way to proceed was to analyze separate datasets and to bridge them by using one of the nodes dated in the first analysis as a calibration point in the second analysis (secondary calibration). I used this approach bridging the Asparagales *rbc*L and *Allium* ITS data-matrices.

In the above mentioned monocot dating studies the two dating methods penalized likelihood (PL; Sanderson 2002) and nonparametric rate smoothing (NPRS; Sanderson 1997) resulted in highly similar ages for most groups, while the third method (PATHd8) estimated generally younger ages, particularly for Asparagales (50-70 Myr younger compared to PL and NPRS), and in several cases the ages estimated by PATHd8 were too young and clearly in conflict with the fossil record, e.g. for Arecaceae (Anderson 2007). Therefore, in this study the calibration dates resulting from PL and NPRS were used. The ages obtained for Alliaceae s. str. using PL and NPRS differed slightly (76 Myr and 79 Myr respectively). These dates were used as minimum and maximum ages to constrain the *Allium* ITS tree (Friesen et al. 2006). This study covers all morphological and geographical

groups of the genus *Allium*, together with the outgroup genera *Ipheion* Raf., *Tulbaghia* L., *Dichelostemma* Kunth., and *Nothoscordum* Kunth. The age for subgenus *Melanocrommyum* resulting from *Allium* ITS dating analyses was applied to the detailed phylogenetic tree of the subgenus as a new calibration point.

A likelihood ratio test of molecular clock-constrained and unconstrained trees was done in PAUP* 4.0b10 (Swofford 2002), which indicated a significant level of rate heterogeneity ($P<<0.001$). Therefore, I applied a rate smoothing approach using penalized likelihood (PL) with the TN algorithm as well as nonparametric rate smoothing (NPRS) to both *Allium* and *Melanocrommyum* ITS trees. The cross-validation procedure for PL was performed to determine the appropriate smoothing levels and was set to S = 10 for *Allium* and S = 750 for *Melanocrommyum*. Both methods (PL and NPRS) are implemented in the computer program R8S v.1.70 (Sanderson 2003).

4.3 Results

Phylogenetic analyses of *Allium rbc*L sequences

The final alignment of *Allium rbc*L sequences is 1305 bp in length, and contains no gaps. The intrageneric pairwise genetic distances at *rbc*L is extremely low and ranges from 0 to 2.67% within the ingroup. The averages and standard deviations of the observed *Ka/Ks* ratios were 1.21 ± 0.63. From 174 variable positions 114 were potentially parsimony informative. We kept only one sequence per species for final analyses except *A. wallichii* and *A. fasciculatum*, where sequenced individuals possess quite divergent sequences; these taxa are shown on the figure with respective accession numbers (Fig. 9). The genus is monophyletic (100% bs and 1.0 pp), using all phylogenetic methods. Bayesian and parsimony analyses resulted in similar tree topologies, whereas the phenetic analysis (neighbor-joining) yielded a slightly better resolved phylogeny, although all short branches collapse during bootstrap analyses. Here I present the Bayesian tree with posterior probabilities (pp) and bootstrap supports from MP (Fig. 9). The conspicuous result of all methods is a mostly unresolved *rbc*L phylogeny. Both parsimony and Bayesian analyses resulted in a large polytomy, in which mostly single *Allium* taxa and only some small taxon groups occur, most of them without statistical support. A few well-supported small groups are recovered from *rbc*L sequence analyses; among them are two groups each uniting three species belonging to subgenus *Reticulatobulbosa*, two groups of subgenus *Rhizirideum* and a small group of taxa classified in subgenus *Allium* (Fig. 9). The only larger group resulting from Bayesian analysis unites the members of several well-defined subgenera: *Melanocrommyum*, *Amerallium*, *Reticulatobulbosa*, *Rhizirideum*, *Cepa*, and *Anguinum.* These subgenera are not separated from each other in *rbc*L analyses, as their members are intermingled in this group with extremely short or even zero-length branches. The group got low BI support (0.82 pp), and no bootstrap support ≥50 in MP. Thus, in contrast to high divergence values at the ITS region, *rbc*L was too conserved for resolving deeper splits and dating divergence times within *Allium*.

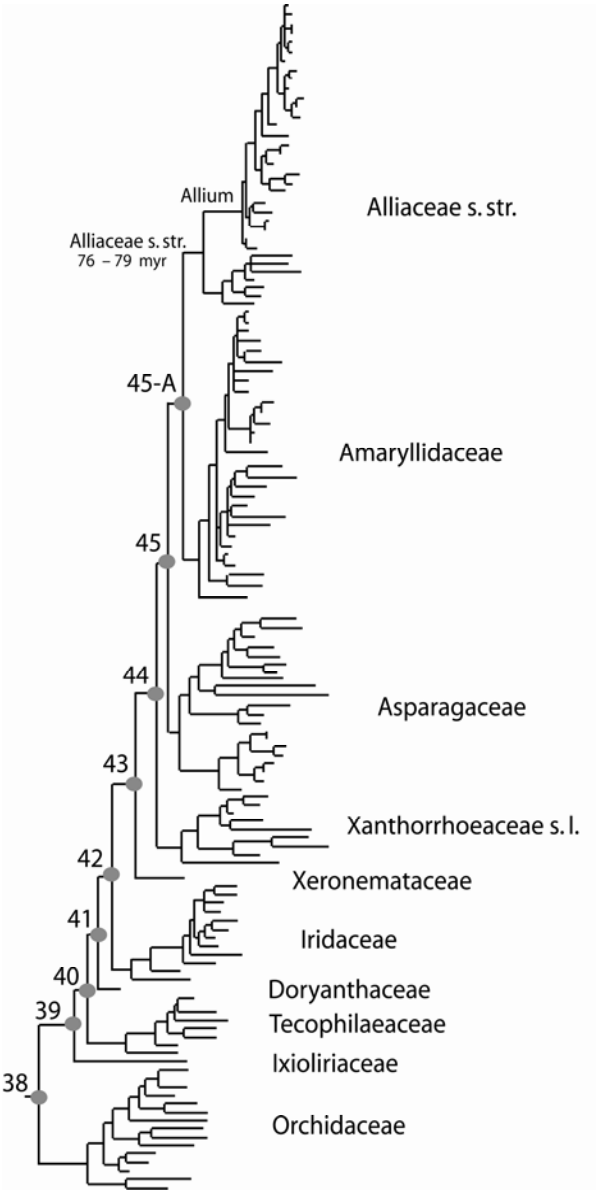
Fig. 9. Phylogeny of *Allium* based on *rbc*L sequences and Bayesian analysis. The numbers above the branches depict posterior probabilities from BI, below the branches bootstrap supports from parsimony analyses are given. The taxonomic affiliations (at subgeneric level) of the groups are shown on the right side.

Estimation of divergence times

ITS substitution rates. The maximum genetic distances within *Allium* and *Melanocrommyum* are high, reaching 53% and 21.7%, respectively. Assuming a molecular clock, maximum likelihood branch lengths in *Allium* correspond to the sequence divergence of 0.255 substitutions per site in ITS since the basal divergence in *Allium*, 0.0936 substitutions/site within *Melanocrommyum* (crown group) and 0.135 substitutions/site since the divergence of the subgenus from its closest outgroup. Based on the published lowest (1.72×10^{-9} substitutions/site/yr) and highest (8.34×10^{-9} substitutions/site/yr) ITS substitution rates for herbaceous annuals/perennials and the pairwise distances in *Allium* at the ITS locus, the approximate age for the genus *Allium* is between 147 and 31 Myr. Applying the same rates to the genetic distances of subgenus *Melanocrommyum* results in a range between 57 and 11 Myr – the time span from early Paleocene to middle Miocene. Thus, the resulting ranges are too high to allow testing any biogeographic hypothesis.

Secondary calibration. The second dating approach using the Asparagales data matrix and phylogenetic tree calibrated with published dates for monocots yielded an age of c. 76 (PL) to 79 (NPRS) Myr (late Cretaceous) for Alliaceae s. str. (Fig. 10) and 45 (PL) to 50 (NPRS) Myr for the *Allium* crown node. As the main interest of this study was to estimate the ages within *Allium*, I proceeded with the analyses with the phylogenetic tree of Friesen et al. (2006. Fig. 11), which contains nearly all Alliaceae s. str. genera as outgroups, and will be later referred to as "*Allium* ITS tree". The obtained ages of 76 (PL) to 79 (NPRS) Myr, were applied to the *Allium* ITS tree as minimum and maximum constraints. The age estimation for the *Allium* crown group using the ITS data is 62 (PL) to 67 (NPRS) Myr, which is close to the age calculated by the average herbaceous plant ITS substitution rate (4.13×10^{-9} substitutions/site/yr), but is older than the age estimated by *rbc*L sequences. The chronogram (Fig. 11) shows the result of the estimated divergence times within *Allium*. Only for subgenus *Butomissa* PL and NPRS yielded quite different results with 21 and 37 Myr, respectively. The ages of subgenera *Allium*, *Cyathophora*, *Rhizirideum* fall all within a range of 28-30 (PL) and 30-36 Myr (NPRS). Subgenus *Amerallium*, which has a disjunct distribution on the northern hemisphere and represents one of the earliest branching lineages, is estimated to be 33 (PL) to 38 (NPRS) Myr old (crown node age).

The most surprising result was an extremely young age estimated for subgenus *Melanocrommyum*, with 7 (PL) to 9 (NPRS) Myr, making it even slightly younger than calculated with the highest up to now reported ITS substitution rates. This would indicate a higher than published rate of ITS evolution within the subgenus (see discussion of this chapter). As I was interested in dating the diversifications within this arid-adapted subgenus, where molecular data suggest at least two rapid diversifications, one relatively early, and a second later one with parallel radiations in several monophyletic groups (see chapter 2), this date of 7 to 9 Myr was applied to the *Melanocrommyum* ITS tree.



Node Number	NPRS	PL
38	119	117
39	112	110
40	108	105
41	107	104
42	103	101
43	100	97
44	93	91
45	91	89
45-A	87	85

Fig. 10. A schematic representation of Asparagales *rbcL* tree including *Allium* sequences from this analyses. The names of the taxa are replaced by the family names, calibration points for the node ages, as well as node numbers are taken from the work of Anderson and Janssen (in press). The only difference with their study is the node 45-A, which does not have node number in their paper, and is referred as crown node age for Alliaceae-Amaryllidaceae clade (Anderson 2007). The resulting age from the present study for Alliaceae s. str. is shown on the scheme in Myr.

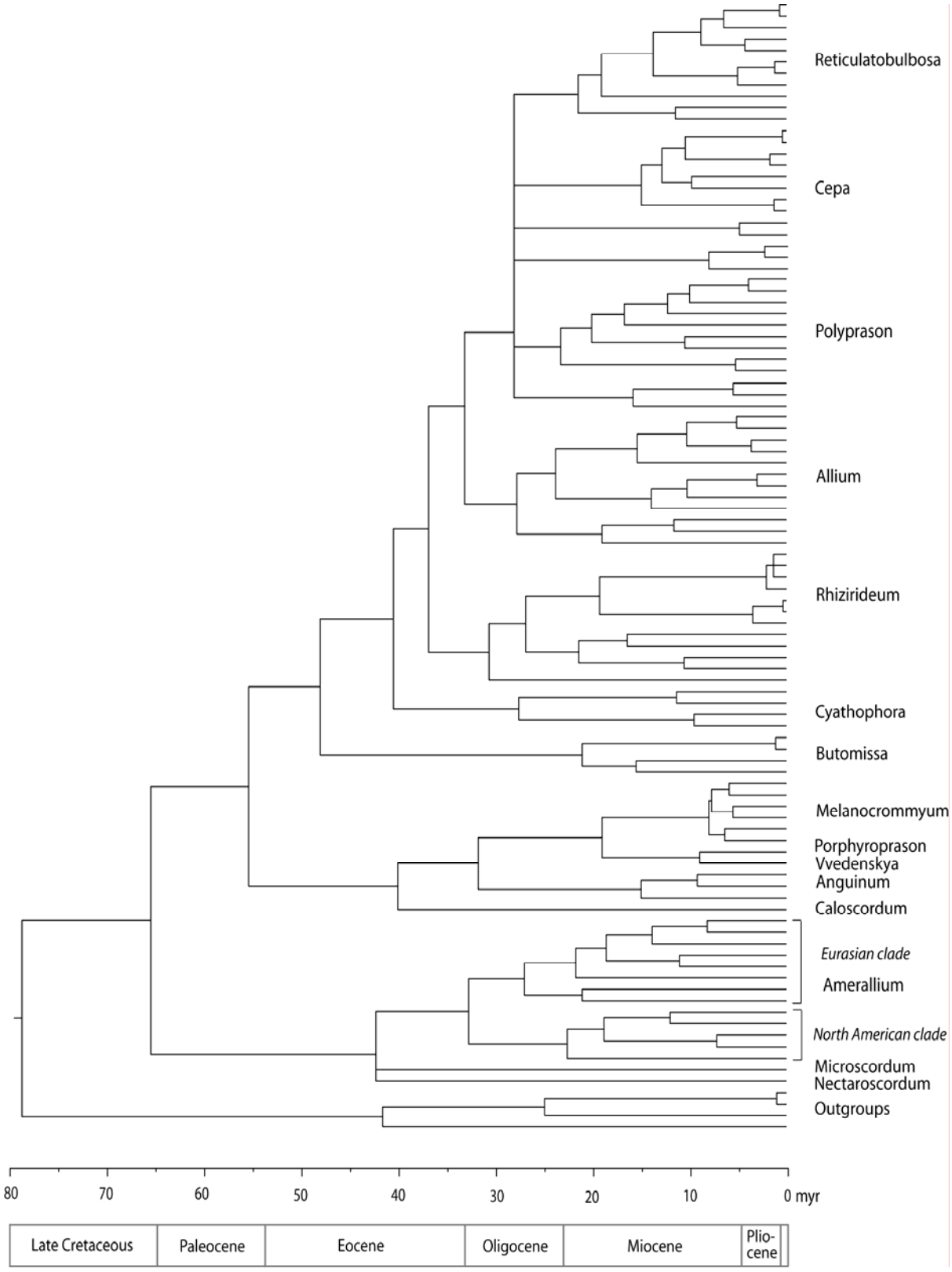


Fig. 11. Chronogram of *Allium* inferred by penalized likelihood (PL) based on the ITS tree from Friesen et al. (2006), which was calibrated with the root node using 76 (Pl) – 79 (NPRS) Myr as minimum and maximum ages (see text for details). The species names are replaced by the names of the subgenera according to Friesen et al. (2006).

The approximate ages for these major diversification events range between 3 – 4.5 Myr for the initial diversification (named core clade in the ITS tree in Chapter 2) and during the last 1 – 0.4 Myr for the terminal groups (Fig. 12), reflecting nearly simultaneous and relatively recent diversification. Although the ages resulting from two different methods are largely congruent, the NPRS dates yield slightly older ages, which was also detected by other studies (Anderson et al. 2005). As for subgenus *Melanocrommyum* these minor differences fall within the above reported interval, the detailed comparison of the outcomes from these two methods for each node are not provided here. Moreover, Sanderson (2002) suggests that PL should be chosen over NPRS when possible, because the latter has a tendency to produce rapid fluctuations in groups with short internal branches.

4.4 Discussion

*Rbc*L variation in *Allium*

Although *rbc*L sequences are mostly used to infer the relationships at higher phylogenetic levels, the low phylogenetic signal at the *rbc*L locus in *Allium* was unexpected taking into consideration large genetic distances in ITS as well as *rbc*L-*atp*B spacer sequences (Friesen, pers. communication). Furthermore, considering that in some genera *rbc*L is almost as variable as noncoding regions (e. g. *Moraeae*; Goldblatt et al. 2002), and the fact that ITS divergence in the genus *Allium* is comparable to the distances in other families (Friesen et al. 2006), it seemed plausible that *rbc*L should be useful to resolve relationships in the genus.

The result that several well-defined subgenera are monophyletic in ITS analyses but occur mixed in the *rbc*L phylogenetic tree (although with low support), could be caused by different reasons. Although this could be the result of shared ancestral alleles, which is a well-known phenomenon for chloroplast loci (chapter 3), it seems more probable that the reason is low resolution, when the amount of variable sites is too low to overcome noise created by a few homoplasious characters. Considering the very low divergence values at the *rbc*L locus, faster evolving chloroplast regions would be necessary to answer this question. Likewise, the reason for intraspecific polymorphism detected in two species (*A. wallichii* and *A. fasciculatum*) remains an open question. As in these cases a lineage sorting scenario would imply quite deep coalescence (at the initial differentiation in *Allium*), ancient introgression might be a more plausible explanation.

Phylogenetic analyses of *rbc*L sequences resulted in a long polytomy, thus, not resolving the phylogenetic relationships among the deeper splits in *Allium*. Taking into account the poorly resolved deep nodes in both, the ITS and *rbc*L phylogenies of *Allium*, this diversification could represent an ancient rapid radiation (Cooper and Penny 1997; Soltis et al. 2004). However, the same result could be caused by using molecular data that are not variable enough at the appropriate level (Whitfield and Lockhart 2007), or by the loss of phylogenetic signal which is often the case when using markers with high substitution rates. The low genetic distances of *rbc*L, accompanied by high divergence

values at the ITS locus in *Allium* indicate that additional markers might be necessary to prove or reject an initial rapid radiation in the genus.

The timing of diversification events in the genus *Allium* and subgenus *Melanocrommyum*

Despite the fact that several studies calibrated the phylogenetic trees based on ITS substitution rates to test biogeographic or some other hypotheses (Kay et al. 2005 in *Costus*; Baldwin 2007 in *Deinandra*), the genetic distances in *Allium* and the variation of substitution rates among angiosperms are too high to arrive at safely interpretable results. Although some studies (Comes and Abbott 2001; Baldwin 2007) use phylogenetic relatedness or more commonly cited rates in choosing ITS substitution rates to get a more reasonable interval, I could not give preferences to any substitution rates because of the following reasons: (i) there are no rates available from the literature for groups phylogenetically closely related to *Allium*, and (ii) extensive survey of ITS substitution rates (Kay et al. 2006) found that phylogenetic relatedness does not adequately account and predict the nature and extent of variation in ITS rates.

The differences in the *Allium* crown node age estimates using Asparagales *rbc*L [45 (PL) – 50 (NPRS)] and Alliaceae s. str. ITS [62 (PL) to 67 (NPRS)] could result from a taxon-density effect in molecular dating (Linder et al. 2005), which means, that more densely sampled clades generally yield older ages. The ITS tree used in the analyses contains more *Allium* taxa than represented in the Asparagales *rbc*L data matrix.

On the other hand, although studies from where the calibration points were taken are the best-sampled studies on dating divergence times in monocots, the ages obtained are generally older than those obtained by Bremer (2000), and in some cases predate (e.g. 10 Myr older for Arecaceae and 5 Myr for Zingiberales) the fossil ages. Although here one can always argue that fossil ages are minimum ages, the reason might be relatively poor fossil record in monocots, and/or available molecular dating methodologies need improvement. In addition, the fact that the clade of Alliaceae is many nodes away from the nearest reference fossil, demands that the results should be taken with caution. Moreover, successive bridging is likely to compound the estimated error (Shaul and Graur 2002), and therefore extrapolation and how much it is acceptable is one of the most controversial issues surrounding the use of molecular clocks (Graur and Martin 2004; Renner 2005). Another conflicting result is the age of subgenus *Melanocrommyum*, where the age obtained by secondary calibration falls outside the quite impressively large range calculated from observed ITS substitution rates. Some authors argue that the lineages which radiate fast have also generally high substitution rates (Barraclough and Savolainen 2001; Xiang et al. 2004), and thus accelerated mutation rates of subgenus *Melanocrommyum* compared to other large subgenera in *Allium* could be another example of such a correlation. However, this challenging but highly speculative explanation requires additional analyses.

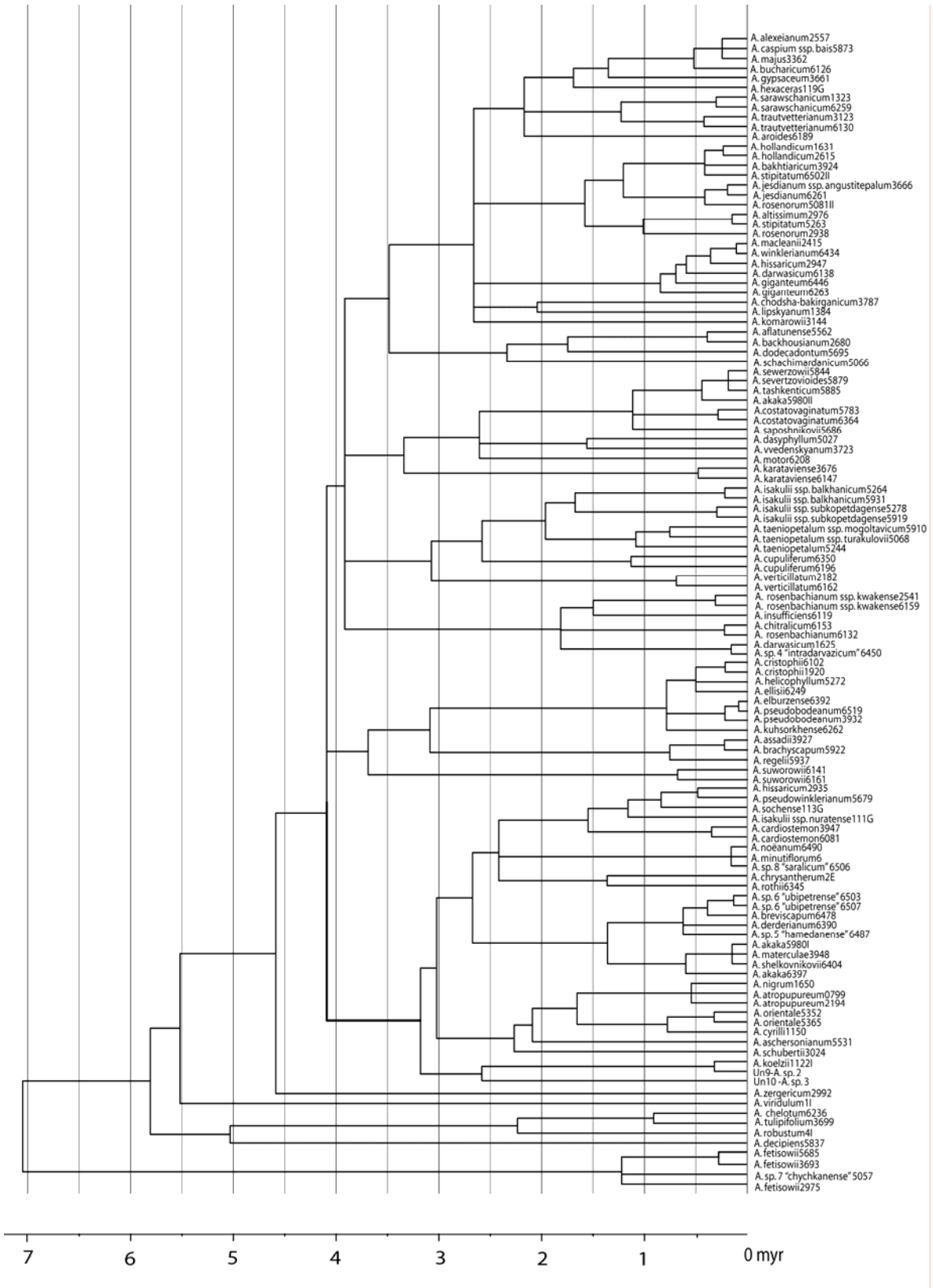


Fig. 12. Chronogram of subgenus *Melanocrommyum* based on ITS data, calculated using penalized likelihood (PL). The age of the crown node of subgenus *Melanocrommyum* was taken from the dated *Allium* ITS tree (chronogram, Fig. 11) as a calibration point. Scale below is in million years.

Biogeographic implications

Any attempt to apply dates to the radiation within a family such as Alliaceae, where fossil records are missing, must be interpreted with caution. Here I make a first effort at dating of diversification events in *Allium* using smoothing methods, which relax a strict assumption of a molecular clock. The time of origin of the *Allium* lineage (stem age) is estimated to fall in the Cretaceous – Tertiary boundary, while the first diversifications within the genus are inferred to have occurred during the Paleocene and early Eocene. However, both dates are consistent with the hypothesis that the early lineages of *Allium* already existed before the global aridification at the Eocene – Oligocene transition. Interestingly, the divergence times of the major *Allium* lineages fall within the time frame of aridification in continental Asia associated with the onset of glaciation in Antarctica about 34 million years ago (Dupont-Nivet 2007), and is consistent with the hypotheses of Hanelt (1992) that the main diversifications in the genus are connected to the opportunity occupying vast territories following aridification in the Tertiary.

The stem age of subgenus *Amerallium*, which is a monophyletic group with disjunct distribution in North America and Eurasia falls in the middle Eocene [42 (Pl) – 48 (NPRS) Myr old], whereas the division into New World and Old World lineages is dated to occur at the Eocene – Oligocene boundary. As the opening of the Atlantic Ocean broke up the connection between North America and Eurasia in the upper Eocene, plant migration between these two continents was possible during the Oligocene till early Miocene only via stepping-stone islands in the North Atlantic (Tiffney 1985a, b). These dates allow the following hypothesis to explain the extant distribution of *Amerallium* taxa. Members of the stem group of this subgenus had a western Laurasian distribution and split when the North Atlantic opened. Afterwards Old and New World lineages became reciprocally monophyletic via lineage sorting. As no other *Allium* lineages show this New World – Old World disjunction (apart from from *A. schoenoprasum*) it is tempting to assume that the stem groups of the other subgenera were distributed mainly in the eastern part of Laurasia, and were therefore not affected by the opening of the North Atlantic and the split of the continents. Otherwise one would expect similar disjunctions in several other *Allium* groups (like in other plant groups, e.g Papaveraceae, Blattner and Kadereit 1995) or has to postulate extinction of all early *Allium* lineages except *Amerallium* in North America. However, the latter scenario would imply that none of other *Allium* lineages survived in North America, which considering generally similar conditions required by the members of the genus seems much less likely. I can, of course, not exclude Oligocene or Miocene migration along island chains (stepping stones) or long-distance dispersal of an *Amerallium* taxon to North America. The coincidence of the date of the split between Old and New World *Amerallium* and opening of the North Atlantic seems however to favour the vicariance hypothesis. This hypothesis relies on the monophyly of subgenus *Amerallium*, which still has to be proven.

The apparently young crown age estimation for subgenus *Melanocrommyum* (7 – 9 Myr) falling in late Miocene (although the stem age is 18 Myr) would imply a relatively

recent differentiation. This implies that the early radiation within *Melanocrommyum* is connected to progressive aridification in western Asia from middle Miocene till Pliocene, and climate changes associated to the Messinian salinity crisis after the Mediterranean Sea dried out and created arid habitats in the surrounding area including western Turkey where the second diversity center of *Melanocrommyum* species lies. The age of the terminal groups implies that their members radiated in the upper Pleistocene. Although the influence of Quaternary climatic oscillations on European biota is well studied (Comes and Kadereit 1998; Hewitt 2000; Kadereit and Comes 2005) little is known about the changes in Central Asian plant diversity. Particularly studies using molecular methods are largely missing. This study is one of the first to test biogeographic hypotheses in *Allium*, as well as the hypothesis of rapid diversification in an arid-adapted Eurasian plant lineage. Considering the low resolution at *rbc*L and high substitution rates in nrITS sequences, additional markers seem necessary to resolve the relationships (or prove that the unresolved topologies represent rapid radiations). However, in the absence of fossil record or safe geographical calibration dates molecular makers can only provide rough estimates of divergences times and, thus, are a starting point for the development of biogeographic hypotheses in *Allium*.

5. Genome size variation and evolution in the subgenus *Melanocrommyum*

5.1 Introduction

The angiosperms vary tremendously in nuclear DNA content (from 0.063 to 127.4 pg for 1C DNA, Leitch et al. 2005; Greilhuber et al. 2006), although most of this variation is not associated with differences in gene number or gene size. While it is well documented that, in addition to polyploidization (Leitch and Bennett 1997; Soltis and Soltis 2000), (retro)transposon amplification (Kalendar et al. 2000) has been a major cause of genome expansion, the lack of evidence for counterbalancing mechanisms that curtail unlimited genome growth has prompted the already outdated idea, that angiosperms might have a "one-way ticket to genomic obesity" (Bennetzen and Kellogg 1997). Several studies addressed the issue of genome contraction and proposed mechanisms responsible for genome size decrease, which include unequal crossing over, illegitimate recombination, a higher overall rate of deletions than insertions, and selection against transposable elements (Morgan 2001; Devos et al. 2002; Petrov 2002; Wendel et al. 2002; Ma et al. 2004; Bennetzen et al. 2005). The biological and evolutionary significance of the intensive variation in DNA content is an unsolved puzzle of evolutionary biology. Most theories about genome size evolution emphasize the effects of DNA amount (mass and volume) on phenotype. Such effects include an increase in nucleus and cell size, prolonged duration of both mitosis and meiosis, and possible consequences of these changes (Bennett 1972, 1987; Wyman et al. 1997). While these theories focus on the adaptive importance of genome size variation, Petrov (2001) proposed that the question of genome size evolution should be viewed in terms of population genetics. The principle is simple: a variety of mechanisms create DNA-length mutations either expanding or contracting the genome size. Some of these mutations affect the phenotype and undergo natural selection, while some might have negligible selective effects and are governed primarily by genetic drift. The combined interplay of all these forces affects genome size (Petrov 2001). Within this framework, it is obvious that different evolutionary forces can be important in different organisms and across different time-frames. In other words, long term evolution of genome size and differences among closely related taxa can be influenced by different forces. For example, the fixation of changes in genome sizes (particularly caused by neutral or effectively neutral length mutations) would primarily be influenced by drift in small populations. However, population genetic studies revealing the mode of speciation and sizes of speciating populations are largely missing to check if arising differences in genome size between closely related species are influenced by population-level phenomena.

Given the large range in DNA amounts encountered in plants, any attempt to investigate the directionality of genome size evolution, which itself is closely linked to the elucidation of the biological significance of genome size changes, requires that data are viewed within a rigorous phylogenetic context. Recent accumulation of C-value data (e.g. Plant DNA C-values database; 2003) together with a sound phylogenetic framework in many plant groups has opened up the perspective to study genome size evolution in angiosperms. A few

studies have followed this approach at the family or genus level (Bennetzen and Kellogg 1997; Cox et al. 1998; Jakob et al. 2004; Johnston et al. 2005; Price et al. 2005), while others (Leitch et al. 1998; Soltis et al. 2003) have conducted broader analyses across the entire angiosperm phylogeny or analyzed available C-value data for all land plants (Embryophyta) within a robust phylogenetic framework (Leitch at al. 2005). These studies showed that genome size evolution in all land plants, as well as in angiosperms underwent several independent increases and decreases. However, rigorous analyses of plant genome size variation on the species level in a phylogenetic framework are still largely missing for most groups. Several studies found a correlation of genome size with life history, breeding system and ecology within several taxonomic groups (Albach and Greilhueber 2004; Weiss-Schneeweiss et al. 2006), even within single species (Kalendar et al. 2000; Knight and Ackerly 2002; Jakob et al. 2004) and suggested that genome size in fact has important evolutionary effects (Achigan-Dako et al. 2008). Although intraspecific variation was found in samples from distant populations (Jakob et al. 2004; Schmuths et al. 2004; Smarda and Bures 2006) genome size differences are mostly correlated with ecological differences of the habitat, life form (e.g., Kalendar et al. 2000; Knight and Ackerly 2002; Jakob et al. 2004) or differences in plant phenotype (Knight et al. 2005; Murray 2005; Beaulieu et al. 2007; Achigan-Dako et al. 2008). However, in addition to these factors phylogenetic dependence can play a major role, i.e. closely related species can have similar C values due to their phylogenetic relatedness (Jakob et al. 2004; Weiss-Schneeweiss et al. 2006). Therefore, near-complete species-level phylogenies can contribute to understand to which extent genome size change depends on shared phylogenetic history (phylogenetic dependence), and on the adaptation to the environment (or to other factors noted above).

All *Melanocrommyum* species are specifically adapted to arid conditions, have an extremely short developmental period (ephemeroidal growth form, Hanelt et al. 1992), the species are reported to be mostly diploid with a basic chromosome number of x = 8 and possess quite uniform karyotypes (Fritsch and Astanova 1998), with the exception of *A. karataviense* where x = 9 and *A. chelotum* where x = 10 (R. Fritsch, pers. communication). Tetraploid cytotypes were rarely reported in several species known as diploids, and were mainly neglected due to seldom occurrence (in single accessions of *A. atropurpureum*, Ohri et al. 1992; *A. nigrum*, Shopova 1972, and in *A. akaka*, Vakhtina 1969). This karyological uniformity makes this group particularly interesting to study genome size evolution. In addition, there are no pronounced differences in life form, ploidy level, and even the habitat and ecology are relatively uniform among the members of this subgenus. The only factor about which there is no information in the literature is breeding system, however as deduced from observations in the IPK living collection, outbreeding should be predominant in *Melanocrommyum* taxa (R. Fritsch, pers. communication). Thus, in the absence of the most influential factors, the genome size variation within the subgenus (if present) should have clearly been defined by phylogenetic constraints and, probably should have been influenced by the number of speciation events, selection pressure, and population-level phenomena (e.g. genetic drift, population size). Genome size in

Melanocrommyum was reported to be quite variable according to the study of genome size variation in the whole genus *Allium* by Ohri et al. (1998). Interestingly, 2C DNA content among only 16 *Melanocrommyum* taxa, which were included in this study, varies from 26 to 50 pg (Ohri et al. 1998). In the present investigation genome size variation among 70 taxa (species and subspecies) of subgenus *Melanocrommyum* was determined using flow cytometry, and genome size variation was analyzed in a phylogenetic framework. This chapter will address: (1) the range of genome size variation in *Melanocrommyum*, if there is evidence (2) for alterations in genome sizes (both up and down) in the phylogenetic context, (3) correlation among genome size variation and phylogenetic groupings, and (4) provide additional evidence that diploid and tetraploid cytotypes could occur within single species of subgenus *Melanocrommyum*.

5.2 Materials and Methods

Plant material

Fresh leaf material of 160 accessions (70 taxa) representing all current taxonomic sections and phylogenetic groups of subgenus *Melanocrommyum* was obtained in early spring from the *Allium* Reference Collection of the IPK Gatersleben (Germany). For most taxa at least two accessions (often several individuals per accession) were analyzed. The species names, accession numbers, genome sizes, and the number of individuals measured for each accession are listed in Table 4.

Nuclear Genome size Estimation

For flow cytometric measurements of the genome size, samples of fresh young leaves were co-chopped with a sharp razor blade together with *Vicia faba* (2C = 26.21 pg) (Genebank Gatersleben accession number: FAB602) as internal standard in a Petri dish containing 1 ml nuclei isolation buffer (Galbraith et al. 1983) supplemented with DNAse-free RNase (50 µg/ml) and propidium iodide (50 µg/ml) and filtered through a 35 µm mesh. For species (accessions) with genomes larger than 60 pg/2C *Allium cepa* (2C = 33.69 pg, Genebank Gatersleben accession number: ALL47) and with genomes smaller than 30 pg/2C *Secale cereale* subsp. *cereale* (2C = 16.01 pg, Genebank Gatersleben accession number: R737) were used as internal standard. The relative fluorescence intensities of stained nuclei were measured using a FACStar^{PLUS} (BD Biosciences) flow sorter equipped with an argon ion laser INNOVA 90C (Coherent). Usually, 10,000 nuclei per sample were analyzed. Accessions with the most contrasting values were measured several times. The absolute DNA amounts of samples were calculated based on the values of the G1 peak means. All the flow cytometric measurements were done in cooperation with Dr. Jörg Fuchs (working group – Karyotype Evolution, IPK). Several examples of the histograms of relative fluorescence intensity are shown in Fig. 13.

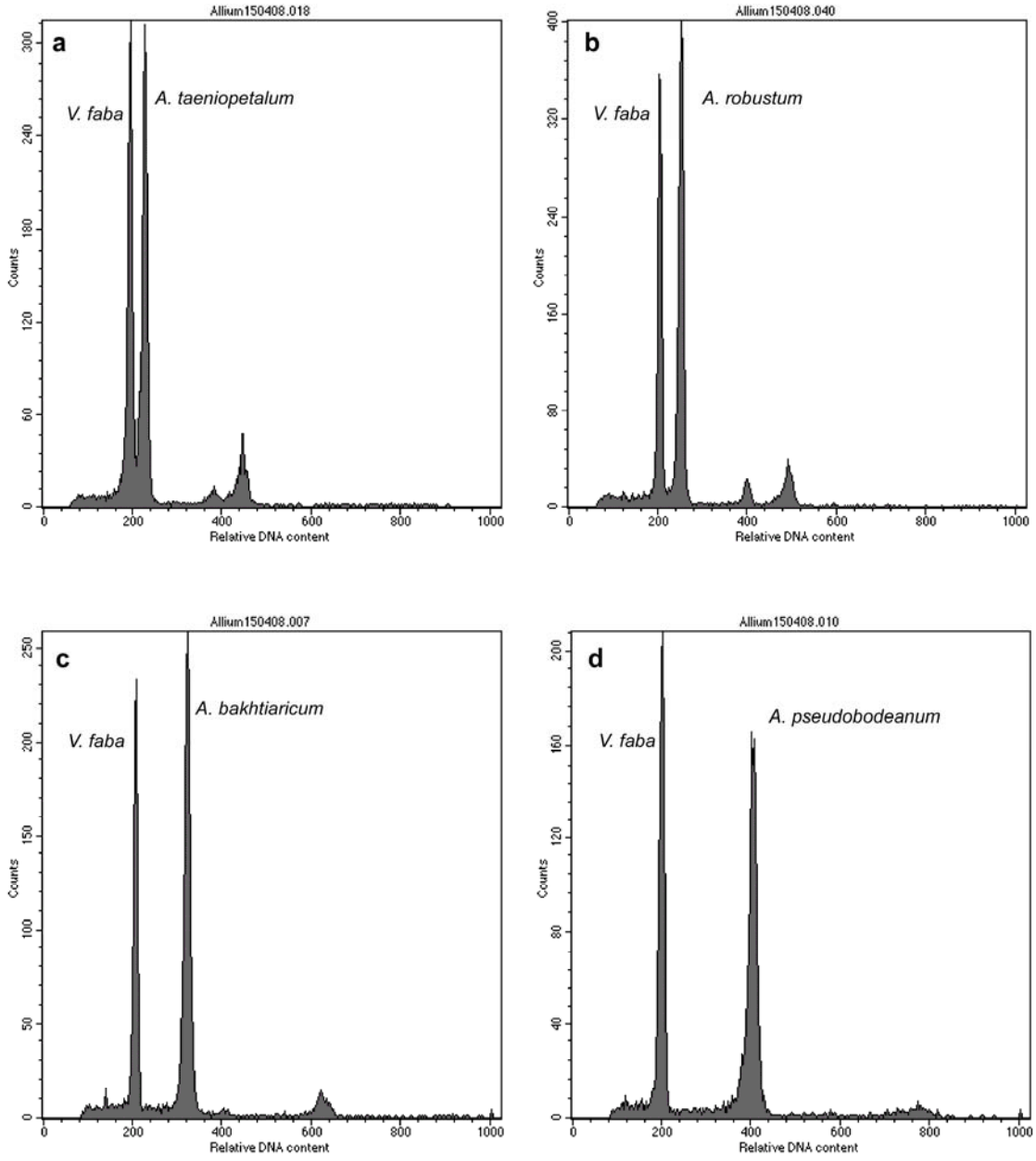


Fig. 13. Flow-cytometric genome size measurements in subgenus *Melanocrommyum*. Examples of histograms of relative fluorescence intensity obtained after analysis of propidium-iodide stained nuclei of the species: *A. taeniopetalum*, *A. robustum*, *A. bakhtiaricum*, *A. pseudobodeanum* together with the internal standard *Vicia faba*.

Chromosome numbers

Chromosome numbers were determined from root tip meristems. The growing root tips were pre-treated in ice-cold (0° C) tap water for 24 hours. Afterwards the root tips were fixed in acetic acid - ethanol mixture (1:3) for 24 hours. After washing in distilled water, and hydrolyzing in 0.2N HCl for 20 minutes at room temperature, the roots were stained with carmine-acetic acid solution. The root tips were squashed in 45% acetic acid.

Table 4. List of the accessions, ploidy level (wherever available) and genome size of the *Melanocrommyum* taxa used in this study.

Species and subspecies	accession No.	Ploidy (2n)	2C DNA amount (pg)	Species' average ± SD	N
***Allium aflatunense* B. Fedtsch.**	5562	16	43.45	43.74 ± 0.36	3
Allium aflatunense B. Fedtsch.	5694		43.94		2
***Allium akaka* S.G. Gmelin ex Schult. et Schult. f.**	6397	32[3]	75.71	74.71[2] ± 0.49	3
Allium akaka S.G. Gmelin ex Schult. et Schult. f.	6398	16	39.39	39.39[2] ± 0.10	3
***Allium alexeianum* Regel**	6291	16	34.18	34.18[2] ± 0.29	3
Allium alexeianum Regel	6303		31.63	31.63[2] ± 0.13	3
***Allium altissimum* Regel**	2976	16	44.76	44.79[1]	1
Allium altissimum Regel	6357		44.81		1
***Allium aroides* Popov & Vved.**	3703	16	38.10	38.14 ± 0.15	1
Allium aroides Popov & Vved.	3707		38.01		2
Allium aroides Popov & Vved.	6189		38.30		2
***Allium atropurpureum* Waldst. et Kit.**	1017	16; 32	52.17	52.24 ± 0.05	2
Allium atropurpureum Waldst. et Kit.	2194		52.29		2
***Allium backhousianum* Regel**	616	16	44.19	44.34 ± 0.21	2
Allium backhousianum Regel	2680		44.53		3
***Allium bakhtiaricum* Regel**	3924		40.96	41.45 ± 0.44	1
Allium bakhtiaricum Regel	6623		41.80		2
***Allium breviscapum* Stapf**	6478	16	45.20	45.95 ± 0.92	2
Allium breviscapum Stapf	6480		46.67		2
***Allium cardiostemon* Fischer et Meyer**	3947	16	41.14	41.14[1,2]	2
Allium cardiostemon Fischer et Meyer	6081		38.83	38.83[2] ± 0.41	3
***Allium caspium* (Pall.) M. Bieb. subsp. *baissunense* (Lipsky) Khassanov et R.M. Fritsch**	5873	16	34.92	34.92 ± 0.35	4
Allium caspium* (Pall.) M. Bieb. subsp. *caspium	5927	16	31.54	31.54[1]	2
***Allium chelotum* Wendelbo**	6236	20	39.74	39.74 ± 0.24	3
***Allium chitralicum* Wang et Tang**	6153	32	68.69	68.69[1]	1
***Allium costatovaginatum* Kamelin et Levichev ex Krassovskaja et Levichev**	5783		34.17	34.37 ± 0.20	4
Allium costatovaginatum Kamelin et Levichev ex Krassovskaja et Levichev	6364		34.41		2
***Allium cristophii* Trautv.**	5920	16	46.10	46.10[1]	2
***Allium cupuliferum* Regel**	6350	16	41.09	41.77 ± 0.43	3
Allium cupuliferum Regel	6354		41.18		2
***Allium cyrilli* Tenore**	1150	32	80.35	78.64 ± 1.73	3
Allium cyrilli Tenore	5296		77.77		3
Allium cyrilli Tenore	1550		79.78		3
Allium cyrilli Tenore	5349		76.65		3
***Allium darwasicum* Regel**	6134	16[3]	34.36	34.11[2] ± 0.35	2
Allium darwasicum Regel	6138		33.87		2
Allium darwasicum Regel	6431		37.56	37.54[2] ± 0.44	3
Allium darwasicum Regel	6553		37.50		1
***Allium decipiens* Fischer ex Schult. et Schult. f.**	2709	16; 20	42.14	42.28 ± 0.13	1
Allium decipiens Fischer ex Schult. et Schult. f.	5837		42.31		3
***Allium derderianum* Regel**	6390	16	43.08	43.51 ± 0.48	2
Allium derderianum Regel	6394		43.91		3
***Allium dodecadontum* Vved.**	5695	16	37.76	37.76[1]	1
***Allium elburzense* Wendelbo**	6519		52.35	51.58 ± 1.22	2
Allium elburzense Wendelbo	6658		50.17		1
***Allium ellisii* J.D.Hooker**	6249		43.43	44.73 ± 1.08	2
Allium ellisii J.D.Hooker	6254		45.01		2
Allium ellisii J.D.Hooker	6255		45.77		2

***Allium fetisowii* Regel**	2975	16	26.33	26.33[1]	2
Allium fetisowii Regel	3693		28.74	28.49 ± 0.19	2
Allium fetisowii Regel	5052		28.33		2
Allium fetisowii Regel	5688		28.52		1
***Allium giganteum* Regel**	6122	16	42.15	42.81 ± 0.77	2
Allium giganteum Regel	6148		42.53		2
Allium giganteum Regel	6258		43.86		2
***Allium gypsaceum* Popov et Vved.**	5669	16	33.82	33.82[1]	2
***Allium hollandicum* R.M.Fritsch**	1631	16	44.66	44.09 ± 0.63	1
Allium hollandicum R.M.Fritsch	2615		43.41		3
Allium hollandicum R.M.Fritsch	2800		44.73		3
***Allium isakulii* R.M. Fritsch et F.O. Khass. subsp. *balkhanicum* R.M. Fritsch et Khassanov**	5932		41.56	41.56[1]	1
***Allium isakulii* R.M. Fritsch et F.O. Khass. subsp. *subkopetdagense* R.M. Fritsch et Khassanov**	5259	16	41.23	41.32 ± 0.27	2
Allium isakulii R.M. Fritsch et F.O. Khass. subsp. *subkopetdagense* R.M. Fritsch et Khassanov	5919		41.61		1
Allium jesdianum* Boiss. et Buhse subsp. *jesdianum	3951	16[3]	38.45	38.27 ± 0.20	2
Allium jesdianum Boiss. et Buhse subsp. *jesdianum*	6261		38.13		2
***Allium jesdianum* Boiss. et Buhse subsp. *angustitepalum* (Wendelbo) Khassanov et R.M. Fritsch**	1083	16[3]	38.62	38.54 ± 0.39	2
Allium jesdianum Boiss. et Buhse subsp. *angustitepalum* (Wendelbo) Khassanov et R.M. Fritsch	3671		38.03		3
***Allium karataviense* Regel**	5040	18	43.05	42.71 ± 0.40	3
Allium karataviense Regel	5793		42.48		2
Allium karataviense Regel	6133		42.20		2
Allium karataviense Regel	6366		42.86		2
***Allium koelzii* (Wendelbo) K.Persson et Wendelbo**	6501		54.77	54.74 ± 0.73	2
Allium koelzii (Wendelbo) K.Persson et Wendelbo	6509		54.48		3
Allium koelzii (Wendelbo) K.Persson et Wendelbo	6491		55.92		2
***Allium komarowii* Lipsky**	3144	16[3]	41.42	40.71 ± 0.57	2
Allium komarowii Lipsky	6282		40.12		1
Allium komarowii Lipsky	6296		40.43		3
***Allium kuhsorkhense* R.M. Fritsch et Joharchi**	6262		46.45	46.45[1]	2
***Allium lipskyanum* Vved.**	1384		34.25	33.64 ± 0.43	3
Allium lipskyanum Vved.	3118		33.52		3
Allium lipskyanum Vved.	5076		33.43		3
***Allium macleanii* J. G. Baker**	2218	16	42.27	42.53 ± 0.71	3
Allium macleanii J. G. Baker	2415		41.85		3
Allium macleanii J. G. Baker	6150		43.81		2
Allium macleanii J. G. Baker	6445		42.86		3
***Allium majus* Vved.**	5073		36.32	36.32[1]	1
***Allium materculae* Bordz.**	3948	16	42.77	42.73 ± 0.10	4
Allium materculae Bordz.	6402		42.82		3
Allium materculae Bordz.	6406		42.61		1
***Allium materculae* Bordz. subsp. *graveolens* R.M. Fritsch ined.**	6612		45.89	45.89[1]	1
***Allium motor* Kamelin**	3355	16	36.39	36.46 ± 0.06	2
Allium motor Kamelin	6208		36.49		2
***Allium nevskianum* Vved. ex Wendelbo**	5078	16	36.51	36.60 ± 0.21	3
Allium nevskianum Vved. ex Wendelbo	5451		36.80		3
***Allium nigrum* L.**	515	16	53.26	54.64 ± 1.60	2
Allium nigrum L.	2616		57.15		2

Allium nigrum L.	5321		54.40		3
***Allium orientale* Boiss**.	5352	16	43.27	43.76 ± 0.79	1
Allium orientale Boiss.	5365		43.08		2
Allium orientale Boiss.	5371		44.68		2
***Allium protensum* Wendelbo**	2162	16	62.18	62.18[1, 2]	1
Allium protensum Wendelbo	3672		32.49	32.49[1, 2]	1
***Allium pseudobodeanum* R.M. Fritsch et Matin**	3932		50.10	51.65 ± 1.09	2
Allium pseudobodeanum R.M. Fritsch et Matin	3936		52.26		3
Allium pseudobodeanum R.M. Fritsch et Matin	6388		52.81		1
***Allium regelii* Trautv.**	5917	16	42.29	42.95 ± 0.69	2
Allium regelii Trautv.	5937		43.32		3
***Allium robustum* Kar.et Kir.**	6565	16	32.24	32.24 ± 0.10	3
***Allium rosenbachianum* Regel subsp. *kwakense* R. M. Fritsch**	6159	16	47.43	47.70 ± 0.61	2
Allium rosenbachianum Regel subsp. *kwakense* R. M. Fritsch	6461		47.38		3
Allium rosenbachianum* Regel subsp. *rosenbachianum	6120	16	41.90	41.48 ± 0.35	3
Allium rosenbachianum Regel subsp. *rosenbachianum*	6132		40.94		3
***Allium rosenorum* R. M. Fritsch**	1869	16[3]	41.19	41.08[2] ± 0.79	3
Allium rosenorum R. M. Fritsch	2938		40.25		2
Allium rosenorum R. M. Fritsch	6158		41.81		2
Allium rosenorum R. M. Fritsch	6290		39.53	39.18[2] ± 0.34	2
Allium rosenorum R. M. Fritsch	6298		38.86		2
Allium rosenorum R. M. Fritsch	5081		39.16		1
***Allium saposhnikovii* E. Nikitina**	6276	16	60.15	60.15[1]	2
***Allium sarawschanicum* Regel**	1323	16	34.20	34.20[2] ± 0.29	3
Allium sarawschanicum Regel	2946		34.58		2
Allium sarawschanicum Regel	6248		36.49	36.49[2]	2
***Allium schachimardanicum* Vved.**	5066		33.94	33.94 ± 0.05	3
***Allium severtzovioides* R.M. Fritsch**	5043		31.34	32.09 ± 0.60	2
Allium severtzovioides R.M. Fritsch	5879		32.73		2
***Allium sewerzowii* Regel**	5047	16	57.04	57.04 ± 0.45	3
***Allium shelkovnikovii* Grossh.**	6404		42.26	42.26[2] ± 0.06	3
Allium shelkovnikovii Grossh.	6405		39.73	39.73[2] ± 0.17	3
Allium sp. 4 "intradarvazicum"	6452	16[3]	37.55	37.92 ± 0.26	2
Allium sp. 4 "intradarvazicum"	6454		38.08		2
Allium sp. 5 "hamedanense"	6487		46.28	46.28 ± 0.57	3
Allium sp. 6 "ubipetrense"	6475		43.55	43.54 ± 0.35	3
Allium sp. 6 "ubipetrense"	6507		43.53		3
Allium sp. 7 "chychkanense"	5057		34.71	34.27 ± 0.67	3
Allium sp. 7 "chychkanense"	5060		33.37		2
Allium sp. 8 "saralicum"	6506		42.80	42.80 ± 0.24	4
***Allium stipitatum* Regel**	3246	16[3]	43.93	43.58 ± 0.79	3
Allium stipitatum Regel	3958		42.01		2
Allium stipitatum Regel	3962		42.83		2
Allium stipitatum Regel	3967		43.13		2
Allium stipitatum Regel	5263		44.31		3
***Allium suworowii* Regel**	5791	16	37.45	37.57 ± 0.22	3
Allium suworowii Regel	6141		37.67		2
***Allium taeniopetalum* Popov & Vved subsp. *mogoltavicum* (Vved.) R.M. Fritsch et Khassanov**	5910	16	30.67	30.62 ± 0.12	2
Allium taeniopetalum Popov & Vved subsp. *mogoltavicum* (Vved.) R.M. Fritsch et Khassanov	6372		30.46		2
Allium taeniopetalum* Popov & Vved subsp. *taeniopetalum	5249	16	31.04	31.30 ± 0.47	3

Allium taeniopetalum Popov & Vved subsp. *taeniopetalum*	6358		31.62		2
Allium taeniopetalum Popov & Vved subsp. *taeniopetalum*	6531		30.71		2
***Allium tashkenticum* Khassanov et R. M. Fritsch**	6171	16	30.28	30.51 ± 0.38	3
Allium tashkenticum Khassanov et R. M. Fritsch	6373		30.95		3
***Allium trautvetterianum* Regel**	6130	16	41.32	41.32 ± 0.51	3
***Allium tulipifolium* Ledeb.**	6599	16	32.16	32.16 ± 0.28	4
***Allium verticillatum* Regel**	6162	16	33.80	33.34 ± 0.48	2
Allium verticillatum Regel	6525		33.05		3
***Allium vvedenskyanum* Pavlov**	3723	16	34.05	34.05[1]	1
***Allium winklerianum* Regel**	6434	16[3]	32.20	33.30 ± 0.89	2
Allium winklerianum Regel	6564		33.69		4

Note: N refers to the number of analyzed individuals from each accession; SD refers to standard deviation calculated from measurements of all individuals of each species. The species names in quotation marks are new taxa, which are not yet formally described.
[1] measurements were based on one or two individuals (thus SD-s are not available)
[2] two distinct 2C values available within one taxon (for details see text)
[3] chromosome counts for present study

Data analyses

C-value statistics. All statistical tests were conducted with SPSS 12.0 (SPSS Inc., Chicago, Illinois, USA). Nuclear DNA-values (2C) for each accession were calculated as mean values from at least three measurements. The analyses were done both with and without known tetraploid accessions due to well known differences in genome evolution after polyploidization. Frequency distribution curves were performed for the variable "genome size" to test for normal distribution. I used Kolmogorov-Smirnov test for every phylogenetic clade to check if the genome sizes are normally distributed. The entire data set was normally distributed, but as the data did not show normal distribution for each phylogenetic clade, Kruskal-Wallis test was used to see whether DNA-value differences among phylogenetic clades were significant, defining clades resulting from ITS phylogenetic analyses as grouping variable. Kruskal-Wallis test is a nonparametric alternative to one-way ANOVA. It is used to compare three or more samples, and it tests the null hypothesis that the different samples in the comparison were drawn from the same distribution or from distributions with the same median (Hill and Lewicki 2006). Thus, the interpretation of this test is basically similar to that of parametric one-way ANOVA, except that it is based on ranks rather than means.

Phylogenetic analyses. Phylogenetic relationships of *Melanocrommyum* species were inferred from sequences of the nrDNA ITS region using distance (neighbor-joing), cladistic (maximum parsimony) and model based (Bayesian inference) methods (see chapter 2). Here a simplified Bayesian tree (Fig. 14) is shown retaining only a single sequence per species. Only in cases when genome sizes differed markedly within a single species, sequences of two individuals were kept to represent the detected genome size variation. For this study I re-calculated a Bayesian tree for ancestral genome size estimation using only

the taxa for which genome sizes were available. However, for brevity, the resulted ancestral genome sizes as well as available information about 2C DNA values of present taxa are plotted on the simplified phylogenetic tree with all available species (Fig. 14).

Evolution of genome size. Ancestral genome sizes within *Melanocrommyum* were assessed using Phylogenetic Generalized Least Squares (PGLS) method implemented in the computer program Compare version 4.6b (Martins 2004), which tests the evolution of continuous characters along the phylogenetic tree from extant taxa and predicts each ancestral state as the weighted average of other taxa on the phylogeny (Martins and Hansen 1997). This GLS model belongs to the directional GLS models, which are particularly suitable for continuous traits, and examine the correlation between the species' trait values and total phylogenetic distance or path length from the root of the tree (Pagel 1999). As the program requires that the phylogeny should be a completely resolved binary tree, the polytomies had been resolved introducing very short branch lengths according to the authors' recommendation (http://compare.bio.indiana.edu/). The predicted genome sizes with standard errors (SE) for major lineages are plotted on the phylogenetic tree (Fig. 14).

5.3 Results

Genome size correlation with phylogenetic clades

The phylogenetic tree based on ITS sequences of 110 species of subgenus *Melanocrommyum* consists of a basal grade (8 species) and a core clade (95% bs, 1.00 pp) with all the remaining species (Chapter 2). In total, 12 well supported groups (A-L) were circumscribed according to the ITS phylogeny (Fig. 14). The variation of genome size among the defined phylogenetic clades is represented in a box plot (Fig. 15). Table 5 shows the mean genome sizes, as well as the number of the species, maximum, minimum values and standard deviations for each phylogenetic group excluding the polyploid accessions. The genome size values were not normally distributed for every phylogenetic clade, which justified the use of a nonparametric test for comparison of clades. Kruskal-Wallis test showed that genome size differences among phylogenetic clades were significant. The ranking of the groups differs from one clade to another (Kruskal-Wallis $\chi^2 = 40.19$, df = 11, $p < 0.001$). The genome sizes in the grade are generally lower than in the core clade (33.6 ± 5.7 pg DNA and 42.0 ± 8.7 pg DNA, respectively). The basal-most group A possess the lowest (29.7 ± 4.11) and group C the highest average DNA (50.217 ± 5.71) values. The latter group unites morphologically similar circum-Mediterranean taxa, all belonging to section *Melanocrommyum,* also containing reported tetraploid *A. cyrilli.* However, the result (the highest 2C DNA value for this group) holds also excluding this tetraploid species.

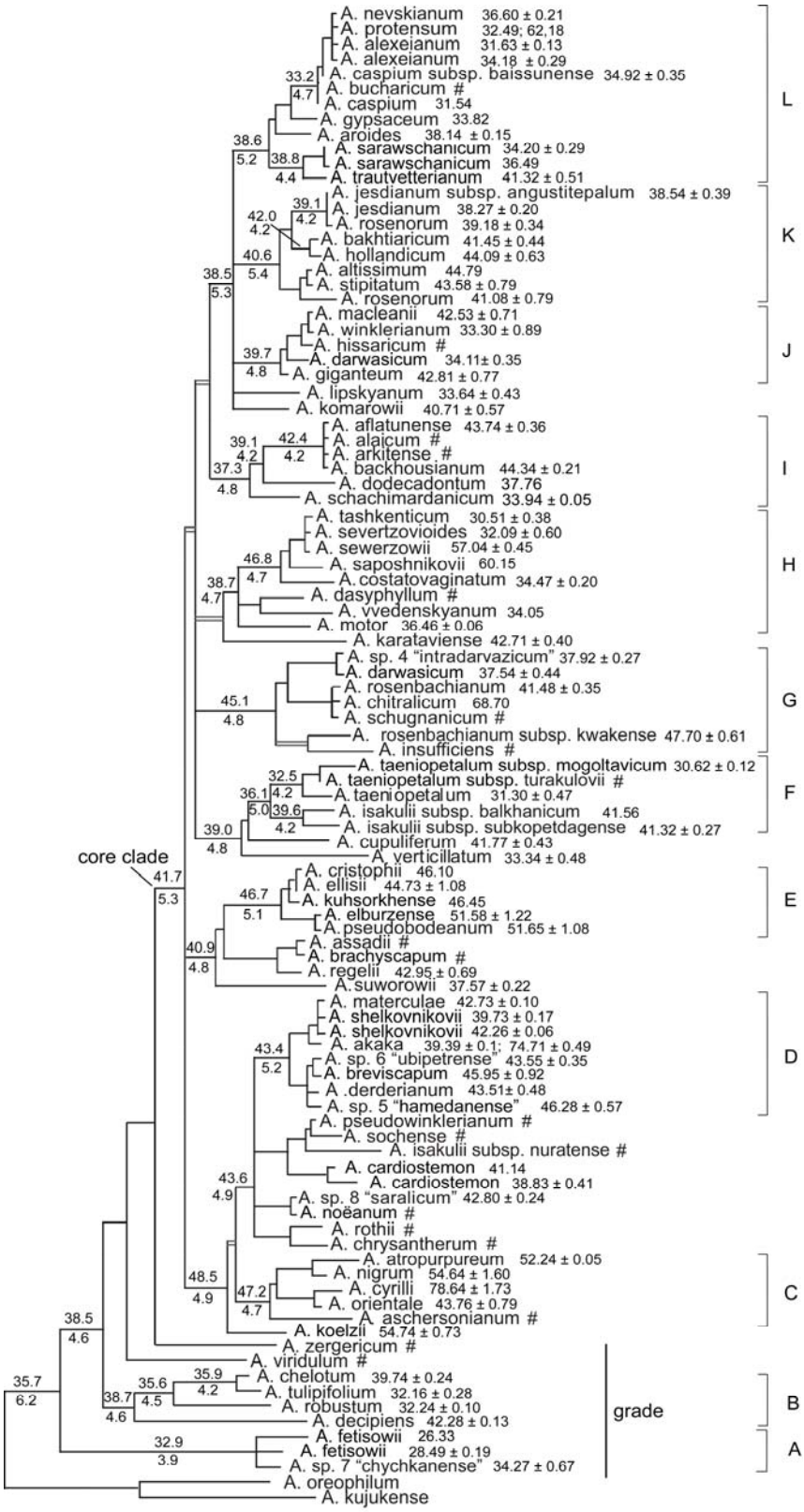


Fig. 14. Phylogenetic tree resulting from a Bayesian analysis of ITS sequences of subgenus *Melanocrommyum* species and two outgroups. Gray double-lines indicate branches not present in the strict consensus (MP) tree. The numbers next to the species names depict 2C DNA values and standard deviation (SD), and the numbers along the branches show estimated ancestral genome sizes with standard errors (SE). Species for which genome sizes are not available, are depicted by *#*, the well-supported groups (according to Chapter 2) are indicated by Latin uppercase letters on the right side.

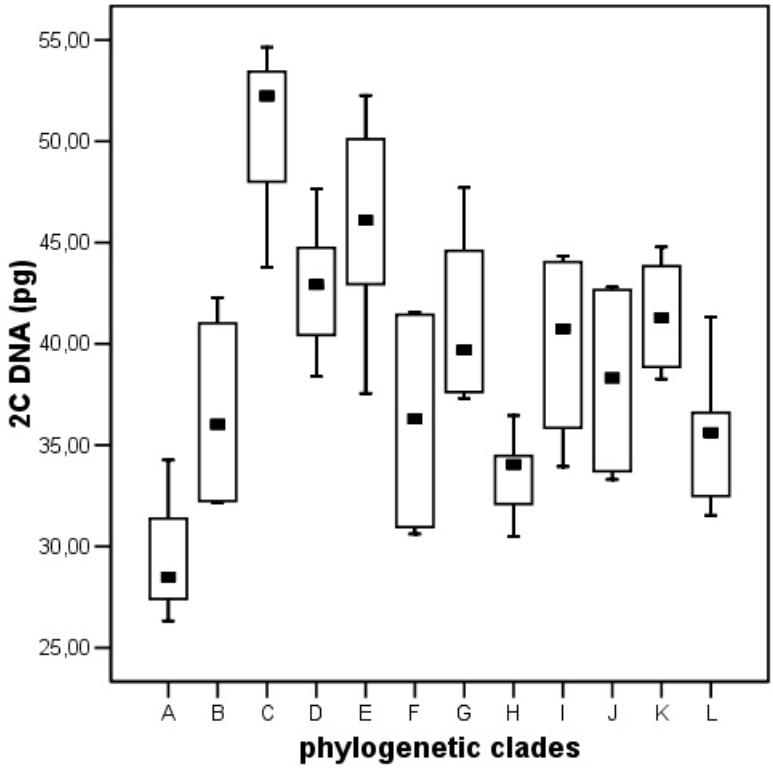


Fig. 15. Box plot displaying the distribution of 2C DNA values (Y – axis) across the defined phylogenetic clades (X – axis), resulting from ITS analyses. ■ represent the median, whiskers the standard deviation.

Table 5. Genome size amount in phylogenetic clades.

Clade	Number	Mean	Minimum	Maximum	SD
A	3	29.70	26.33	34.27	4.11
B	4	36.62	32.16	42.28	5.171
C	3	50.21	43.76	54.64	5.71
D	12	42.76	38.38	47.65	2.93
E	9	45.83	37.56	52.26	4.95
F	4	36.20	30.62	41.56	6.06
G	4	41.11	37.32	47.70	4.77
H	5	33.52	30.51	36.46	2.29
I	4	39.95	33.94	44.34	4.98
J	4	38.19	33.30	42.81	5.18
K	8	41.38	38.27	44.79	2.58
L	10	35.43	31.54	41.32	3.16

Genome size variation

The 2C DNA values of all measured species of subgenus *Melanocrommyum* varies from 26 to 78 pg and 26 to 62 pg when the known polyploids were excluded (Table 4). As chromosome numbers are not available for approximately one-third of the taxa, for which genome size measurements were done, I here report only 2C values and not monoploid genome sizes – Cx (*sensu* Greilhuber et al. 2005).

Genome size measurements from single species were quite similar, as the coefficient of variation was below 2.5%. Although within single species the genome sizes were highly uniform, within-species genome size variability was detected in following taxa: 6% in *A. shelkovnikovii*, 6% in *A. cardiostemon*, 7.8% in *A. fetisowii*, 8% in *A. darwasicum*, and 8.8% in *A. alexeianum*. Moreover, in the two species *A. akaka* and *A. protensum* nearly two-fold differences in genome size were observed: 74.71 pg DNA and 38.38 pg DNA in *A. akaka*, 62.18 pg DNA and 32.48 pg DNA in *A. protensum.* Until now *A. akaka* was reported to be mostly diploid (2n = 16, Pedersen and Wendelbo 1966; Pogosian 1983; Özhatay 1986). However tetraploid and even hexaploid individuals were reported in one study (Vakhtina 1969). Chromosome counting was conducted of *A. akaka* individual with doubled 2C DNA value for the present study, which could confirm tetraploidy of this accession by chromosome numbers 2n = 4x = 32 (Table 4). Despite the fact, that *A. protensum* chromosome counts are available only for one accession (2n = 16, Pedersen and Wendelbo 1966), double difference in DNA amount could indicate the occurrence of tetraploid cytotypes in this species.

In the species *A. saposhnikovii* (60.15 pg) and *A. sewerzowii* (57.04 pg) genome sizes are high, corresponding to a doubled amount compared to closely related taxa (*A. tashkenticum* 30.51 pg and *A. severtzovioides* 32.09 pg). However, both species were reported to be diploids (Pogosian and Seisums 1992; Fritsch and Astanova 1998).

Ancestral genome size estimation and genome size evolution

Ancestral genome size estimation showed that the alterations during phylogeny occurred in both direction (increase and decrease) along the lineages leading to the extant species. However, within the subgenus genome size increase prevailed over decrease. In the basal grade most taxa show the lowest DNA amount, and PGLS predicted 2C DNA amount for the root of the subgenus *Melanocrommyum* to be 35.68 ± 6.21, which is smaller than genome size of most contemporary taxa. However, it is obvious, that both gain and loss can occur in closely related lineages. For example, in the clades F, I, and K genome size increases in one sister lineage, while it is decreasing in another (Fig. 14).

5.4 Discussion

Genome size variation within the subgenus *Melanocrommyum*

Generally high genome size values (2C ranging from 26 to 78 pg DNA) in *Melanocrommyum* are in agreement with the observation/hypothesis that the plants with largest genomes have short periods of growth followed by long periods of dormancy (e.g geophytes in the orders Asparagales and Liliales; Chase et al. 2005). Although the underlying mechanisms are unknown, subgenus *Melanocrommyum* is another example of plants having extremely short developmental mode and high DNA values. The genome size differences across species are more than two-fold. Unfortunately insufficient information about the ploidy of many taxa with high amount of DNA does not allow conclusions if the genome sizes among species are correlated to their ploidy level. For example, the high 2C DNA value in *A. saposhnikovii* (60.15 pg) and *A. sewerzowii* (57.04 pg) could indicate that tetraploid cytotypes occur in this species, taking into account that the closely related species have only half of their genome sizes (30 to 34 pg). However, the tetraploidy of these individuals needs to be confirmed by chromosome counts, as well as hypothesized differences in ploidy level in *A. protensum* where DNA content differs two-fold between the different accessions. In case of *A. akaka*, detected doubled DNA amount is correlated with a difference in ploidy level. Thus, the presence of different cytotypes explains taxonomic complexity of this morphologically highly variable species, and once more proves the hypothesis proposed based on morphological and molecular grounds (Gurushidze et al. 2008, Chapter 2) that hybridization is not rare in the lineage comprising *A. akaka* and a few closely related species.

Taxonomic relevance of genome size variation

Pronounced differences in genome size within plant species were often reported (e.g., Price et al. 1981; Graham et al. 1994; Bennett and Leitch 1995), correlated with ecological differences of the habitat (e.g., Kalendar et al. 2000; Jakob et al. 2004) or differences in plant phenotype (Knight et al. 2005; Murray 2005; Beaulieu et al. 2007). However, as Obermayer and Greilhuber (2005) point out, intraspecific variation in genome size must be taken with utmost caution, as many earlier studies probably reported differences caused by technical artifacts. However, intraspecific genome size variation may also indicate micro-evolutionary differentiation and could be taxonomically significant (Murray 2005; Achigan-Dako et al. 2008). This was clearly demonstrated in *Juncus biglumis* populations (Schönswetter et al. 2007), where both, genome size differences and molecular markers yielded the congruent main groups.

Although the aim of this paper was not analyzing within-species genome size variation in *Melanocrommyum* taxa, interestingly in some cases within-species differences in DNA content were reflected in molecular differentiation at the ITS region. The within-species differences in genome size were 6% of measured DNA content in *A. cardiostemon*, 7.8% in *A. fetisowii*, and 8% in *A. darwasicum*. As it was discussed in Chapter 2, in *A. darwasicum*

ITS sequences and RAPD markers showed genetically clearly separated lineages, and the above mentioned genome size difference within this species correlates with cryptic taxa detected by molecular data. Although *A. cardiostemon* and *A. fetisowii* are not polyphyletic according to ITS sequences, the genetic differentiation at the ITS locus of these species is quite high. Thus genetic divergence coupled with genome size differences within *A. darwasicum* is another support for separately evolving lineages within morphologically indistinguishable species, while in *A. cardiostemon* and *A. fetisowii* genome size differences could indicate the process of ongoing micro-evolutionary differentiation. Furthermore, there are at least several examples (*Hordeum spontaneum*, Kalendar et al. 2000; *Dactylis glomerata*, Reeves et al. 1998; *Juncus biglumis*, Schönswetter et al. 2007) where intraspecific genome size variation might be indicative of incipient speciation. Although one might argue, as Greilhuber and Speta (1985) did, that whether such cases represent intraspecific variation or not, depends on the species concept applied, and if a narrow concept is applied then the intraspecific variation disappears.

Another explanation for this result (congruent differences in genome size and genetic differentiation) could be that differences in both ITS sequence and DNA content are caused by ancient interspecific hybridization. As after introgression ITS could get partially homogenized, hybridization could be responsible for genetic distance. In this case also genome size differences in hybrid offspring could be caused by the contribution of species with different genomes. Thorough population genetic analyses with sampling multiple individuals and analyzing additional unlinked loci would be required to distinguish between these two explanations (ongoing micro-evolutionary structure versus ancient hybridization). Furthermore, in *A. shelkovnikovii* (6% difference in 2C DNA content) and *A. alexeianum* accessions with the highest genome size variation (8.8%), ITS sequences were identical. The range of variation seems quite high compared to the "normal" range of variation (up to 3% of measured genome size for most taxa); however, it is possible that the genetic differentiation between these individuals is too young for the resolution of ITS marker.

Genome size changes in relation to phylogeny

Based on comparison of the genome size between the basal grade and core clade, and estimated ancestral and contemporary genome size (see results), DNA increase seems to be the predominant process in *Melanocrommyum*. However genome size changes do not always show this trend, as both DNA gain and loss were detected even within closely related lineages.

The result of ancestral genome size estimation is in agreement with other evolutionary reconstructions showing that both increases and decreases have occurred in different taxonomic groups (e.g. Wendel et al. 2002; Jakob et al. 2004).

The fact that genome sizes are quite conserved among closely related species and there are significant differences of genome size among defined phylogenetic groups of subgenus *Melanocrommyum* indicate that genome size is not distributed randomly, and rather the distribution of genome size among species is strongly dependant on phylogenetic

relationships. Intraspecific stability of genome size (Greilhuber 1998) and reports of rapid genome reorganization after speciation in diploids (Rieseberg et al. 1995; Rieseberg 1998) and polyploids (Song et al. 1995; Özkan et al. 2001) imply that genome size differences are not correlated with time since divergence or rate of mutation accumulation. Therefore, some authors suggested a speciation model with most changes occurring immediately after speciation to be more realistic (Gold and Amemiya 1987; Albach and Greilhuber 2004; Jakob et al. 2004; for a review see Gregory 2004). Under these assumptions one should expect to find conspicuous differences within species-rich groups, which underwent many speciation events (Jakob et al. 2004). Interestingly the variation within most species-rich phylogenetic groups (D, E, K, L) is not higher than within species-poor groups (A, C, F, G, I, J), and also comparisons of DNA values among pairs of sister taxa do not show conspicuous differences in subgenus *Melanocrommyum*. This is also in agreement with the limited genome size variation in several species-rich animal groups (reviewed in Gregory 2004) making clear that extensive speciation can occur without substantial changes in genome size. The above mentioned prompts the hypothesis that probably not only the number of speciation events and ecological constraints, but also the population size of speciating lineages is crucial for genome size changes. For example, during founder event, when a small population gets separated, different ecological conditions, or just some stochastic influence like genetic drift can fix the differences in DNA content. Unfortunately, population genetic studies in non-model groups of organisms, like *Allium*, are largely missing. Thus without knowing population size differences during speciation the hypothesis remains non-testable.

To summarize, the analyses showed significant differences in 2C DNA content among phylogenetic clades, while the closely related taxa show highly similar genome sizes. The alterations in genome size occurred in both direction (increase and decrease) in the phylogeny along the lineages leading to the present species, and provide evidence that diploid and tetraploid cytotypes could occur within species. Finally, genome size differences correlated with genetic differences of several taxa, which even display topological differences on phylogenetic trees (Chapter 2), could indicate the process of ongoing micro-evolutionary differentiation and might therefore be of taxonomic importance.

6. Summary

Studying diversification processes in subgenus *Melanocrommyum* is interesting due to several aspects: (i) it is a species-rich group with species occurring in relatively similar ecological habitats, without major morphological differentiation within the group (qualitative morphological characters for taxonomic classification are largely missing), and thus could be speculated to have evolved in a non-adaptive radiation, (ii) the group is taxonomically complicated with different and contradictory taxonomic treatments, and was thought to include a considerable number of hybrid species, as taxa often show an admixture of morphological characters used for delimitation of infrasubgeneric groups.

Ecological restrictions, a relatively narrow distribution range compared to other large subgenera of *Allium*, special synapomorphies and specializations raised the hypothesis that subgenus *Melanocrommyum* is a highly derived group within *Allium*, which diversified in the emerging arid areas after closing of the Tethys Sea (Hanelt 1992).

The main goal of this thesis was to evaluate the relationships within the subgenus, and determine the underlying mechanisms of its diversification. Nearly complete species sampling and analyzing of multiple individuals for widespread taxa to control for intraspecific variation, combining information from nuclear and chloroplast genomes, as well as extensive survey of anatomical (septal nectary) and genome size variation, undertaken in this study contributed to understanding the major aspects of evolution and diversification of the group.

6.1 Phylogeny and diversification of the subgenus *Melanocrommyum*

Based on detailed nuclear rDNA ITS phylogeny covering most taxa, all existing taxonomic groups and their entire geographic distribution, analyzing multiple individuals of species with broad geographic ranges, extensive cloning of ITS amplicons allowed to arrive at following conclusions:

The monophyly of the subgenus is well-supported, whereas most current sections proved either para- or polyphyletic, thus molecular data (both nuclear and chloroplast) demand a new treatment at sectional level. The splits of the large sections got support by differences in the anatomy of flower nectaries.

In the phylogenetic tree inferred by ITS sequence analyses, most taxa of subgenus *Melanocrommyum* are assigned to groups forming a polytomic core clade. The unresolved relationships among them could indicate an initial rapid radiation, followed by diversifications within different monophyletic groups (named as C – L in Chapter 2) of the core clade. The timing of diversification processes within the genus *Allium* with the main focus on estimating the age of inferred radiations in subgenus *Melanocrommyum* was assessed using molecular clock approaches. The attempt to estimate the dates of these diversifications were hampered by unavailability of the adequate marker system in *Allium*, and the absence of the fossils or/and safe geographical or climatic calibration points (Chapter 4). However, using dated nodes from monocots allowed to roughly approximate

the diversification times of the inferred radiations. The resulting age estimates for an initial diversification in *Melanocrommyum* following the Miocene/Pliocene aridification are consistent with the hypothesized scenario of evolution after closure of the Tethys Sea. According to this approach, the later radiations resulting in most extant taxa, seems to have occurred during the Pleistocene, probably connected to major climatic oscillations. Synthesis from nuclear and chloroplast data, and results from molecular dating infer that subgenus *Melanocrommyum* is a relatively young group of *Allium*, which underwent at least two radiations at different time scales. The group probably had not faced severe bottlenecks in the past and is still expanding its distribution range.

6.2 Eurasian – North American disjunct distribution of *Allium*

Despite several factors noted above, impeding more precise age estimations for particular lineage splits, all available data clearly indicate that the genus *Allium* already existed before the global aridification at the Eocene–Oligocene transition. The crown age of *Allium* subgenus *Amerallium*, which is apart from chives (*A. schoenoprasum*) the only group of the genus with a North American and Eurasian distribution, was inferred to fall within the late Eocene, while several other lineages of *Allium* already existed by the middle Eocene. The North American and Eurasian species of subgenus *Amerallium* form sister clades, making each of these groups monophyletic. Coincidence of the split between these groups and the brake-up of Laurasia indicates a vicariance event for this group, i.e. an early lineage of this group was probably subdivided by the opening of the North Atlantic. Alternatively it is also possible that North America was reached from Eurasia via a North Atlantic land bridge or stepping-stone islands sometime during the upper Eocene or early Miocene. Late Paleocene and early Eocene lineages of *Allium* probably were distributed in more eastern parts of Eurasia and therefore never reached North America. They started to diversify with the onset of aridification in Central and Southwest Asia, where the main diversity centers are located at present.

6.3 Hybrids versus cryptic species

Based on the ITS data, within each of the following species *A. stipitatum, A. rosenorum* and *A. darwasicum* two groups were separated in the absence of intra-individual ITS polymorphisms (Chapter 2). The first idea to explain the result was a hybridization scenario. Due to concerted evolution (Arnheim et al. 1980) of ribosomal DNA loci, ITS can be inherited uniparentally. If in offspring of hybrid taxa ITS homogenization worked towards the different parental ITS types (bidirectional homogenization, Wendel et al. 1995), this would result in the division of such a hybrid taxon into two “ITS groups” after several generations in the absence of gene flow among them. As one major force in plant speciation is (allo)polyploidization, the next step was to determine the ploidy level of these species. However, chromosome counts of multiple individuals of the three taxa (Chapter 5) confirmed that the individuals of all inferred cryptic species are diploids, thus, showing that

allopolyploidy did not play a role in the described pattern. While in two cases (*A. stipitatum* and *A. rosenorum*) these ITS lineages were close relatives, in *A. darwasicum* they fall in quite different clades in the phylogenetic tree. Genetic fingerprinting was used for *A. darwasicum* individuals falling in different ITS clusters to see if the result is confirmed by arbitrary screening of large parts of the genome. The RAPD profiles show that this result is not due to introgression of rDNA (ITS capture) but that genome-wide differences between both lineages exist. Additional support for separated lineages in *A. darwasicum* was provided by differences in genome size between individuals forming the distinct molecular groups (Chapter 5). Thus, the case of *A. darwasicum* represents one of the rare cases in plants where morphologically indistinguishable diploid species occurring in mixed populations are non-sister cryptic species.

6.4 Reticulations within the subgenus *Melanocrommyum*

Combining information from nuclear and chloroplast molecular data pointed to putative hybrid taxa (*A. jesdianum* subsp. *angustitepalum* and *A. winklerianum*), as "cyto-nuclear discordance" (Wendel and Doyle 1998) was present in all analyzed individuals. To see if the taxa are of allopolyploid origin, chromosomes were counted in multiple individuals of both species. Interestingly, the chromosome counts proved both species diploids (Chapter 5). This, of course, does not preclude the hybrid origin of these taxa, as homoploid hybrid formation and also diploidization after polyploidization are not rare in angiosperms. The fact that approximately 70% of angiosperms are thought to have experienced one or more episodes of polyploidy at some point in the past (Masterson 1994; The French–Italian Public Consortium for Grapevine Genome Characterization 2007) makes the situation even more complicated. However, the same basic number of chromosomes (2x = 16) in both *A. jesdianum* subsp. *angustitepalum* and *A. winklerianum* questions the hypothesis of polyploidization and re-diploidization resulting in exactly the same basic number of chromosomes. Nonetheless, further work and sequencing of additional genomic regions are necessary for safe conclusions about homoploid hybrid status for *A. jesdianum* subsp. *angustitepalum* and *A. winklerianum*.

Although nuclear and chloroplast loci used in this study do not support the hypothesis of frequent gene flow among species, in several cases (*A. nevskianum*, *A. hissaricum*, and *A. rosenorum*) single individuals clearly possess "captured" chloroplast types (Chapter 3). Thus, information from both marker systems shows that natural interspecies hybrids occur rarely in subgenus *Melanocrommyum*.

6.5 Advantages of network-based methods over bifurcating phylogenies

The outcome from chloroplast *trn*L-*trn*F sequence analysis is generally congruent with nuclear data and supports the divisions of all large taxonomic sections. However, several characteristics of the data violate the main assumptions of traditional tree building analysis methods and consequently could cause problems in species-level phylogenetic studies.

These characteristics are: (i) sharing of haplotypes among several morphologically well separable species, which also are differentiated based on the ITS sequences, (ii) presence of several closely related haplotypes within single species on the other hand, and (iii) persistence of ancestral sequences and their co-existence with several descendant alleles, thus, creating multifurcating structure. Unfortunately, all these phenomena are largely overlooked in species-level studies when multiple individuals per species are not sampled. The findings of this study are consistent with previous studies which showed how incomplete lineage sorting, and survival of ancestral alleles through several speciation events (Dobeš et al. 2004; Jakob and Blattner 2006) shaped the present distribution of chloroplast haplotypes, and resulted in several shared alleles among closely related (Comes and Abbott 2001) or relatively distantly related (Jakob and Blattner 2006) plant species. Thus, the present extensive survey of chloroplast *trn*L-*trn*F region in *Melanocrommyum* is in agreement with the above mentioned studies, and revealed a chloroplast haplotype distribution which is consistent with incomplete lineage sorting processes. The present study advocates including multiple individuals per species in phylogenetic studies, and shows the advantages of networks over tree building methods for analyzing the sequences from chloroplast loci in species-level phylogenies.

6.6 Genome size and Phylogeny

Nearly two-fold variation of genome size without doubling of the chromosome number across the species of subgenus *Melanocrommyum* (e. g., *A. tashkenticum* 30.51 pg DNA and *A. rosenbachianum* subsp. *kwakense* 47.70 pg DNA, both species 2n = 2x = 16) can not be explained by any adaptive theory of genome size variation, as life form and even habitat are nearly identical for the species under study. Combining phylogenetic and genome size data in subgenus *Melanocrommyum* does also not support a speciation model with most genome size changes occurring immediately after speciation (punctuated model), as species-rich groups do not display higher variation than species-poor groups. Intraspecies genome size differences coupled with genetic divergence show that genome size difference can indeed be indicative of micro-evolutionary differentiation. Another outcome of this study, the occurrence of different cytotypes within *A. akaka*, together with morphological and molecular evidence confirms the hybridization hypothesis within a closely related species group containing also *A. akaka*.

Abstract

The subgenus *Melanocrommyum* is one of the largest in *Allium*, and comprises about 140 arid adapted perennial species from Eurasia, with the highest diversity in mountainous Central and Southwest Asia. Several taxonomic classifications of the subgenus proposed during the last century differ considerably in the number of accepted taxa and infra-subgeneric groups. Previous studies hypothesized that a high level of hybridization among *Melanocrommyum* species occurs. Within the frame of this thesis the phylogenetic relationships within subgenus *Melanocrommyum* were evaluated using nuclear rDNA ITS and chloroplast *trn*L-*trn*F sequences. Both markers were sequenced in multiple individuals of about 110 species, covering most sections and subsections from the entire geographical distribution of the subgenus. Phylogenetic groups resulting from the analyses of the ITS region were mostly in conflict with traditional taxonomy, i.e. most sections proved either para- or polyphyletic. As the phylogenetic analysis of maternally inherited chloroplast sequences resulted in poorly resolved trees, a genealogical network approach was used to analyze chloroplast relationships within the subgenus. The chloroplast haplotype network showed that ancient chloroplast types co-occur with their descendants, and are often involved in multifurcating relationships, as they gave rise to multiple chloroplast lineages. The taxon groups inferred from *trn*L-*trn*F sequences turned out to be mainly congruent with the nuclear phylogeny, as genealogical lineages are consistent with the groups inferred from ITS sequences. Thus, both bi- and uniparentally inherited data-sets strongly contradict the morphology-based taxonomic classification. As shape and position of the septal nectaries and their excretory canals differ remarkably among several taxonomic groups of *Allium,* the nectary anatomy in subgenus *Melanocrommyum* was investigated to see if groups defined by ITS data are supported by this character. The split of the large poly- or paraphyletic sections into smaller groups could also be confirmed by the anatomy of flower nectaries.

ITS data separated two distinct groups within each of the following three species *A. stipitatum*, *A. rosenorum* and *A. darwasicum*. In two cases (*A. stipitatum* and *A. rosenorum*) these lineages were close relatives and also geographically separated. In *A. darwasicum*, however, they fall in quite different clades of the phylogenetic tree and show no relation to geographical distribution. Random amplified polymorphic DNA (RAPD) fingerprinting markers resulted in clearly different banding patterns between these groups of *A. darwasicum*. Thus, RAPD analyses proved that the split is not caused by introgression of rDNA (ITS capture), but genetic differentiation involves large parts of the genome. The result was further confirmed by differences between genome sizes of the inferred cryptic lineages. The present study reports an example of sympatric non-sister cryptic taxa, where high genetic distances between the two lineages of *A. darwasicum* imply long-term reproductive isolation.

Nuclear genome size evolution in subgenus *Melanocrommyum* was analyzed in a phylogenetic framework. For that purpose, DNA contents were measured in about 70 taxa by flow cytometry. The observed 2C genome size variation is quite high varying from 26 to

78 pg among the members of the subgenus. The smallest genome sizes occur in basal taxa, while the largest genome sizes were detected in a group of closely related Mediterranean species. However, estimation of ancestral genome sizes using Phylogenetic Generalized Least Squares (PGLS) revealed lineages with increasing as well as decreasing DNA content. The analyses showed that significant differences exist in DNA content among phylogenetic clades, while closely related taxa show highly similar genome size values. Intraspecies genome size differences found in species, which also show genetic differentiation at the ITS locus, could reflect incipient speciation/diversification, while two-fold differences of DNA content in other taxa indicate the occurrence of diploid and tetraploid cytotypes in these species.

The biogeography of the genus *Allium*, its origin, and the causes of North American – Eurasian disjunction, were studied using relaxed molecular clock approaches. Asparagales *rbc*L sequences and calibration points from studies on monocots were used to infer the age of Alliaceae s. str. The resulting age was applied to an *Allium* ITS phylogeny to obtain age estimates for lineages in *Allium*. Precise date estimates for diversifications could not be obtained due to several reasons (absence of *Allium* fossils and therefore the lack of safe calibration points, and not perfectly suitable molecular marker). However, the following conclusions can be drawn from available results: the genus has a Eurasian origin, the present North American – Eurasian disjunct distribution of the subgenus *Amerallium* is probably the result of vicariance after opening of the North Atlantic. However, later arrival of *Amerallium* in North America via the North Atlantic land bridge cannot be completely ruled out. The inferred initial radiation of *Melanocrommyum* is likely connected to progressing Miocene/Pliocene aridification in Central and Southwest Asia, as well as in the Mediterranean, whereas the later diversifications falling within the last one million years could be promoted by changing conditions due to the Quaternary climate oscillations.

To summarize, the taxonomy of subgenus *Melanocrommyum* proved to be artificial by phylogenetic analyses of nuclear and chloroplast markers, rendering most sections polyphyletic, although used markers could not fully resolve the relationships among the clades. The results presented in this thesis (i) infer two major diversification events, and (ii) show that hybridization and polyploidy did not play a major role in this radiations. These outcomes and the absence of pronounced adaptive phenotypic differentiation in *Melanocrommyum* species, support the hypothesis of non-adaptive radiations, promoted by newly emerging arid habitats during the extensive aridification in the Central and Southwest Asia in the upper Tertiary, and migrations and subdivisions of populations during Quaternary climatic changes.

Zusammenfassung

Die Untergattung *Melanocrommyum* ist eine der größten innerhalb der Gattung *Allium*. Sie umfaßt ca. 140 an Trockenklimate angepasste, ausdauernde eurasische Arten. Die größte Diversität dieser Untergattung findet sich in den Bergregionen Zentral- und Südwestasiens. Während des letzten Jahrhunderts wurden mehrere, teils widersprüchliche taxonomische Klassifikationen der Untergattung *Melanocrommyum* aufgestellt, die sich nicht nur in der akzeptierten Artenzahl, sondern auch hinsichtlich der infrasubgenerischen Gruppierung stark voneinander unterscheiden. Frühere Studien der Untergattung postulierten auch, dass Hybridisierung eine wichtige Rolle in der Gruppe spielt.

Im Rahmen dieser Doktorarbeit wurden die Verwandtschaftsbeziehungen innerhalb der Untergattung *Melanocrommyum* mit molekularen Methoden untersucht. Als Marker wurden Sequenzen der Spacer nuklärer rDNA (ITS) sowie der *trn*L-F Region des Chloroplastenchromosoms untersucht. Beider Regionen wurden für jeweils mehrere Individuen von 110 Arten analysiert. Diese Arbeit deckt damit das gesamte geographische Verbreitungsgebiet der Untergattung, sowie nahezu alle Sektionen und Untersektionen ab.

Die auf ITS-Sequenzen basierende molekulare Phylogenie stand im Widerspruch zu der bisher gängigen taxonomischen Klassifikation der Untergattung, da die meisten Sektionen para- oder polyphyletische Gruppierungen bildeten.

Ein netzwerkbasierender Analyseansatz wurde für die Chloroplastendaten verwendet, da baumbasierende Methoden die Datenstruktur (gleichzeitiges Vorkommen von Vorläufer- und Nachfahreallelen sowie multiple Verzweigungen) nur unzulänglich wider geben konnten. Das Chloroplasten-Haplotypennetzwerk zeigte, dass ältere Vorläuferhaplotypen oft mit mehreren ihrer Nachfolger in den analysierten Individuen koexistieren. Die in der Chloroplastenanalyse gefundenen Gruppen sind jedoch größtenteils in Übereinstimmung mit den Gruppierungen des ITS-Stammbaums. Die durch die molekulare Analyse definierten neuen Gruppierungen in *Melanocrommyum* werden durch anatomische Strukturen der Septalnektarien unterstützt, deren unterschiedliche Form und Position für die Mehrheit dieser Gruppen typisch ist.

Die ITS-Analyse zeigte innerhalb von drei Arten (*A. stipitatum*, *A. rosenorum* und *A. darwasicum*) zwei jeweils klar getrennte innerartliche Gruppen. In zwei Fällen (*A. stipitatum* and *A. rosenorum*) waren diese Gruppen zwar geographisch klar getrennt, aber phylogenetisch nahe verwandt. Im Falle von *A. darwasicum* dagegen fielen die beiden molekularen Gruppen in weit voneinander entfernte Clades des Stammbaums. Geographisch sind sie jedoch nicht getrennt, sonder treten im Gegenteil beide innerhalb von Populationen auf. Mit RAPD-*fingerprinting* konnte gezeigt werden, dass die beiden innerartlichen Gruppen klar unterschiedliche Bandenmuster besaßen. Dies bestätigte genetische Unterschiede auch in weiteren Teilen des Genoms. Die Analyse konnte somit nachweisen, dass die gefundenen Unterschiede im ITS nicht auf Introgression und biparentale Homogenisierung der rDNA-Region beruhen, sondern andere, bisher unbekannte Ursachen hat. Die Trennung zwischen den beiden *A. darwasicum*-Typen spiegelt sich auch in unterschiedlichen Genomgrößen wider. Diese beiden Taxa stellen

somit ein seltenes Beispiel für kryptische Arten dar, die nicht in einem Schwestergruppenverhältnis stehen.

Im Rahmen dieser Doktorarbeit wurde ebenfalls die Evolution von Genomgrößen in der Untergattung *Melanocrommyum* in einem phylogenetischen Rahmen analysiert. Dafür wurden für 70 Arten die Genomgrößen mit Durchflusszytometrie bestimmt. Es wurden 2C-Werte von 26 bis 78 pg gemessen. Die kleinsten Genome kommen in den basalen Gruppen der Untergattung vor, während die größten Werte in einer Gruppe eng verwandter Arten aus dem Mittelmeergebiet gefunden wurden. Mittels *Generalized Least Squares* wurden Genomgrößen an den Knoten des phylogenetischen Baumes abgeschätzt. Diese Abschätzung von Vorläufer-Genomgrößen ergab, dass in der Untergattung sowohl Linien mit ab- als auch mit zunehmenden 2C-Werten vorkommen. Darüber hinaus weisen eng verwandte Arten ähnliche Genomgrößen auf, während zwischen unterschiedlichen Verwandtschaftsgruppen deutlich größere Differenzen in den 2C-Werten auftraten. Zudem zeigten die Messungen, dass in einigen Arten tetraploide Zytotypen vorkommen.

Um die Entstehung der Gattung *Allium*, ihre Biogeographie, ihre disjunkte Verbreitung in Eurasien und Nordamerika sowie das Alter von Radiationen innerhalb der Gattung zeitlich einordnen zu können wurden molekulare Datierungsmethoden (*molecular clock*) verwendet. Sequenzen des Chloroplastengens *rbc*L aus den Asparagales sowie Kalibrierungspunkte aus den Monokotylen wurden verwendet um das Alter der Alliaceae s. str. zu bestimmen. Diese Datierung wurde dann auf einen ITS-Stammbaum der Gattung übertragen, um das Alter von Untergattungen und Sektionen ermitteln zu können. Präzise Datierungen waren nicht möglich, da Fossilien in der Gattung und Familie fehlen und Kalibrierungspunkte daher aus weiter entfernten Gruppen der Monokotylen stammen. Weiterhin zeigte der *rbc*L-Locus nur sehr geringe Sequenzvariation innerhalb der Alliaceae, während die ITS-Sequenzen innerhalb Gattung extrem divers waren, was diese Marker problematisch zur präziesen Datierung macht. Die Ergebnisse der Alterdatierung ließen jedoch folgende Schlussfolgerungen zu: (1) *Allium* enstand vor ca. 64 Millionen Jahren (Ma) in Eurasien. (2) Die einzige auch in Amerika vorkommende Untergattung *Amerallium* entstand an der Grenze Eozän/Oligozän (ca. 35 Ma) und besiedelte Nordamerika entweder bereits bevor sich der Nordatlantik öffnete oder erreichte Nordamerika über eine nordatlantische Landbrücke. (3) Eine erste Radiation innerhalb von *Melancrommyum* wurde wahrscheinlich durch die neogene Aridifizierung Zentral- und Südwestasiens ausgelöst (Miocän/Pliocän, Beginn vor ca. 15 Ma). Jüngere schnelle Radiationen in der Untergattung fallen dagegen zeitlich in das Pleistozän und stehen vermutlich mit den zyklischen Klimawechseln dieser Perioden im Zusammenhang.

Zusammenfassend erweist sich die Taxonomie der Untergattung *Melanocrommyum* als künstlich, verglichen mit den phylogenetischen Analysen der Kern- und Chloroplasten-Marker, die die meisten Sektionen in polyphyletischer Anordnung aufzeigten. Die Verwandtschaftsbeziehungen zwischen den Clades konnten jedoch mit den verwendeten Marker nicht vollständig aufgelöst werden. Die Ergebnisse dieser Arbeit lassen (i) auf zwei Haupt-Diversifizierungsereignisse schließen und zeigen, (ii) dass Hybridisierung und

Polyploidie keine wesentliche Rolle bei diesen Radiationen spielten. Diese Resultate, sowie das Fehlen ausgeprägter adaptiver phenotypischer Differenzierung in *Melanocrommyum* Arten, unterstützen die Hypothese von nicht-adaptiver Artbildung in dieser Untergattung. Die nicht-adaptiver Artbildung wurden weiterhin gefördert durch neu entstandene Trockenhabitate, die sich während der ausgeprägten Aridifizierung im Oberen Tertiär in Zentral- und Südwestasien bildeten, sowie durch Migration und Teilung von Populationen während der Klimaschwankungen des Quartär.

Literature cited

Achigan-Dako E. G., Fuchs J., Ahanchede A., Blattner F. R. 2008. Flow cytometric analysis in *Lagenaria siceraria* (Cucurbitaceae) indicates correlation of genome size with usage types and growing elevation. Plant syst. Evol. Doi: 10.1007/s00606-008-0075-2.

Albach D. A., Greilhuber J. 2004. Genome size variation and evolution in *Veronica*. Ann. Bot. 94: 897–911.

Albach D. C., Martínez-Ortega M. M., Chase M. W. 2004. *Veronica*: parallel morphological evolution and phylogeography in the Mediterranean. Plant Syst. Evol. 246: 177–194.

Alfaro M. E., Zoller S., Lutzoni F. 2003. Bayes or Bootstrap? A simulation study comparing the performance of Bayesian Markov Chain Monte Carlo sampling and bootstrapping in assessing hylogenetic relationships. Mol. Biol. Evol. 20: 255–266.

Álvarez I., Wendel J. F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29: 417–434.

An Z., Kutzbach J. E., Prell W. L., Porter S. C. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature 411: 62–66.

Anderson C. L. 2007. Dating Divergence Times in Phylogenies. Doctoral thesis, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology, Acta Universitatis Upsaliensis.

Anderson C. L., Janssen T. (in press) Monocots. In: Hedges B., Kumar S. (eds) Timetree of life.

Anderson C. L., Bremer K., Friis E. M. 2005. Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil constraints. Amer. J. Bot. 92: 1737–1748.

Arbogast B. S., Edwards S. V., Wakeley J., Beerli P., Slowinski J. B. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu. Rev. Ecol. Syst. 33: 707–740.

Archibald J. K., Mort E. M., Crawford D. J. 2003. Bayesian inference of phylogeny: a non-technical primer. Taxon 52: 187–191.

Arnheim N., Krystal M., Schmickel R., Wilson G., Ryder O., Zimmer E. 1980. Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc. Natl. Acad. Sci. USA 77: 7323–7327.

Axelrod D. I. 1979. Age and origin of Sonoran Desert vegetation. Occas. Paper. Calif. Acad. Sci. 132. 74 pp.

Bachmann K. 1997. Nuclear DNA markers in plant biosystematic research. Op. Bot. 132: 137–148.

Baldwin B. G. 2007. Adaptive radiation of shrubby tarweeds (*Deinandra*) in the California Islands parallels diversification of the Hawaiian silversword alliance (Compositae–Madiinae). Amer. J. Bot. 94: 237–248.

Baldwin B. G., Sanderson M. J. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc. Natl. Acad. Sci. USA 95: 9402–9406.

Baldwin B. G., Sanderson M. J., Wojciechowski J. M., Campbell C. S., Donoghue M. J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Missouri Bot. Gard. 82: 247–277.

Bänfer G., Moog U., Fiala B., Mohamed M., Weising K., Blattner F. R. 2006. A chloroplast genealogy of myrmecophytic *Macaranga* species (Euphorbiaceae) in Southeast Asia reveals hybridization, vicariance and long-distance dispersals. Mol. Ecol. 15: 4409–4425.

Barraclough T. G., Reeves G. 2005. The causes of speciation in flowering plant lineages: species-level DNA trees in the African genus *Protea*. In: Bakker F. T., Chartou L.W., Gravendeel B.,

Pelser P. B. (eds) Plant species-level systematics: new perspectives on pattern and process. A.R.G Gantner Verlag, Liechtenstein. pp. 31–46.

Barraclough T. G., Savolainen V. 2001. Evolutionary rates and species diversity in flowering plants. Evolution 55: 677–683.

Beardsley P. M., Olmstead R. G. 2002. Redefining Phrymaceae: The placement of *Mimulus*, tribe Mimuleae and *Phryma*. Amer. J. Bot. 89: 1093–1102.

Beaulieu J. M., Moles T. A., Leitch I. J., Bennett M. D., Dickie J. B., Knight C. A. 2007. Correlated evolution of genome size and seed mass. New Phytol. 173: 422–437.

Bennett M. D. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. Proc. Roy. Soc., B. Biol. Sci. 181: 109–135.

Bennett M. D. 1987. Variation in genomic form in plants and its ecological implications. New Phytol. 106: 177–200.

Bennett M. D., Leitch I. J. 1995. Nuclear DNA amounts in angiosperms. Ann. Bot. 76: 113–176.

Bennetzen J. L., Kellogg E. A. 1997. Do plants have a one-way ticket to genomic obesity? Plant Cell 9: 1509–1514.

Bennetzen J. L., Ma J. X., Devos K. M. 2005. Mechanisms of recent genome size variation in flowering plants. Ann. Bot. 95: 127–132.

Bergsten J. 2005. A review of long-branch attraction. Cladistics 21: 163–193.

Berry P. E., Hahn W. J., Sytsma K. J., Hall J. C., Mast A. 2004. Phylogenetic relationships and biogeography of *Fuchsia* (Onagraceae) based on noncoding nuclear and chloroplast DNA data. Amer. J. Bot. 91: 601–614.

Birky C. W. Jr. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms and models. Annu. Rev. Genet. 35: 125–148.

Blattner F. R. 1999. Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. Biotechniques 29: 1180–1186.

Blattner F. R. 2004. Phylogenetic analysis of Hordeum (Poaceae) as inferred by nuclear rDNA ITS sequences. Mol. Phylogenet. Evol. 33: 289–299.

Blattner F. R., Kadereit J. W. 1995. Three intercontinental disjunctions in Papaveraceae subf. Chelidonioideae: evidence from chloroplast DNA. – In: Jensen U., Kadereit J.W. (eds) Systematics and Evolution of the Ranunculiflorae. Pl. Syst. Evol. Suppl. 9: 147–157

Blattner F. R., Weising K., Bänfer G., Maschwitz U., Fiala B. 2001. Molecular analyses of phylogenetic relationships among myrmecophytic *Macaranga* Species (Euphorbiaceae). Mol. Phylogenet. Evol. 19: 331–344.

Boore J. L. 2006. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol. Evol. 21: 439–446.

Borsch T., Hilu K. W., Quandt D., Wilde V., Neinhuis C., Barthlott W. 2003. Noncoding plastid *trn*T-*trn*F sequences reveal a well resolved phylogeny of basal angiosperms. J. Evol. Biol. 16: 558–576.

Bowen G. J. 2007. When the world turned cold. Nature 445: 607–608.

Bremer K. 2000. Early Cretaceous lineages of monocot flowering plants. Proc. Natl. Acad. Sci. USA 97: 4707–4711.

Britten R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science 231: 1393–1398.

Cannon C. H., Manos P. S. 2003. Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). J. Biogeogr. 30: 211–226.

Chase M. W. 2004. Monocot relationships: An overview. Amer. J. Bot. 91: 1645–1655.

Chase M. W., Hanson L., Albert V. A., Whitten W. M., Williams N. H. 2005. Life history evolution and genome size in subtribe Oncidiinae (Orchidaceae). Ann. Bot. 95: 191–199.

Chase M. W., Soltis D. E., Olmstead R. G., Morgan D., Les D. H., Mishler B. D., Duval M. R., Price R. A., Hills H. G., Qiu Y-L., Kron K. A., Rettig J. H., Conti E., Palmer J. D., Manhart J. R., Sytsma K. J., Michael H. J., Kress W. J., Karol K. G., Clark W. D., Hedren M., Gaut B. S., Jansen R. K., Kim K-J., Wimpee C. F., Smith J. F., Furnier G. R., Strauss S. H., Xiang Q-Y., Plunkett G. M., Soltis P. S., Swensen S. M., Williams S. E., Gadek P. A., Quinn C. J., Eguiarte L. E., Golenberg E., Learn G. H. Jr., Graham S. W., Barrett S. C. H., Dayanandan S., Albert V. A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Miss. Bot. Gard. 80: 528–580.

Chiu W. L., Stubbe W., Sears B. B. 1988. Plastid inheritance in *Oenothera*: organelle genome modifies the extent of biparental transmission. Curr. Genet. 13: 181–189.

Clauss M., Koch M. A. 2006. *Arabidopsis* and its poorly known relatives. Trends Plant Sci. 11: 449–459.

Clément C., Pacini E. 2001. Anther plastids in angiosperms. Bot. Rev. 67: 54–73.

Clement M., Posada D., Crandall K. A. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657–60.

Collevatti R. G., Grattapaglia D., Hay J. D. 2003. Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Brazilian cerrado based on the analysis of chloroplast DNA sequences and microsatellite haplotype variation. Mol. Ecol. 12: 105–115.

Comes H. P., Kadereit J. W. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3: 432–438.

Comes H. P., Abbott R. J. 2001. Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* section *Senecio* (Asteraceae). Evolution 55: 1943–1962.

Cooper A., Penny D. 1997. Mass survival of birds across the Cretaceous – Tertiary boundary: molecular evidence. Science 275: 1109–1113.

Cox A. V., Abdelnour G. J., Bennett M. D., Leitch J. 1998. Genome size and karyotype evolution in the slipper orchids (Cypripedioideae: Orchidaceae). Amer. J. Bot. 85: 681–687.

Crandall K. A., Templeton A. R. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134: 959–69.

Dally A. M., Second G. 1990. Chloroplast DNA diversity in wild and cultivated species of rice (Genus *Oryza*, section *Oryza*). Cladistic-mutation and genetic-distance analysis. Theor. Appl. Genet. 80: 209–222.

Daly D. C., Cameron K. M., Stevenson D. W. 2001. Plant systematics in the age of genomics. Pl. Physiol. 127: 1328–1333.

Davis C. C., Bell C. D., Mathews S., Donoghue M. J. 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA 99: 6833–6837.

Devos K., Brown J., Bennetzen J. 2002. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. Genome Res. 12: 1075–1079.

Dobeš C., Mitchell-Olds T., Koch M. 2004. Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii, A.* × *divaricarpa,* and *A. holboellii* (Brassicaceae). Mol. Ecol. 13: 349–370.

Dodd M. E., Silvertown J., Chase M. W. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. Evolution 53: 732–744.

Don G. 1832. A monograph of the genus *Allium*. Memoirs of the Wernerian Natural History Society. Vol. 6. Adam Black - North Bridge, Edinburgh. pp. 1–102.

Downie S. R., Palmer J. D. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis P. S., Soltis D. E., Doyle J. J. (eds) Molecular systematics of plants. Chapman and Hall, New York. pp. 14–35.

Doyle J. J., Luckow M. S. 2003. The rest of the iceberg: legume diversity and evolution in a phylogenetic context. Plant Physiol. 131: 900–910.

Doyle J. J., Gaut B. 2000. Evolution of genes and taxa: a primer. Plant Mol. Biol. 42: 1–23.

Doyle J. J., Davis J. I., Soreng R. J., Gravin D., Anderson M. J. 1992. Chloroplast inversions and the origin of the grass family (Poaceae). Proc. Natl. Acad. Sci. USA 89: 7722–7726.

Drábkova L., Kirschner J., Vlček Č., Paček V. 2004. *Trn*L-*trn*F intrgenic spacer and *trn*L intron define major clades within *Luzula* and *Juncus* (Juncaceae): Importance of structural mutations. J. Mol. Evol. 59: 1–10.

Dubouzet J. G., Shinoda K., 1998. Phylogeny of *Allium* L. subgenus *Melanocrommyum* (Webb et Berth.) Rouy based on DNA sequence analyses of the internal transcribed spacer region of nrDNA. Theor. Appl. Genet. 97: 541–549.

Dubouzet J. G., Shinoda K. 1999. Relationships among Old and New World *Allium* according to ITS DNA sequence analysis. Theor. Appl. Genet. 98: 422–433.

Dubouzet J. G., Shinoda K., Murata N. 1997. Phylogeny of *Allium* L. subgenus *Rhizirideum* (G. Don ex Koch.) Wendelbo according to dot blot hybridization with randomly amplified DNA probes. Theor. Appl. Genet. 95: 1223–1228.

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

Dupont-Nivet G., Krijgsman W., Langereis C. G., Abels H. A., Dai S., Fang X. 2007. Tibetan plateau aridification linked to global cooling at the Eocene–Oligocene transition. Nature 445: 635–638.

Efron B., Tibshirani R. J. 1993. An introduction to the bootstrap. Chapman and Hall, New York. 436 pp.

Fay M. F., Chase M. W. 1996. Resurrection of Themidaceae for the *Brodiaea* alliance, and recircumscription of Alliaceae, Amaryllidaceae and Agapanthoideae. Taxon 45: 441–451.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27: 401–410.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Felsenstein J. 2004. Inferring phylogenies. Sinauer Associates, Sunderland, MA. 664 pp.

Fernandez C. C., Shevock J. R., Glazer A. N., Thompson J. N. 2006. Cryptic species within the cosmopolitan desiccation-tolerant moss *Grimmia laevigata*. Proc. Natl. Acad. Sci. USA 103: 637–642.

Friesen N., Blattner F. R. 2000. RAPD analysis reveals geographic differentiations within *Allium schoenoprasum* L. (Alliaceae). Plant Biol. 2: 297–305.

Friesen N., Fritsch R. M., Bachmann K. 1997. Hybrid origin of some ornamentals of *Allium* subgenus *Melanocrommyum* verified with GISH and RAPD. Theor. Appl. Genet. 95: 1229–1238.

Friesen N., Fritsch R. M., Blattner F. R. 2006. Phylogeny and new intrageneric classification of *Allium* L. (Alliaceae) based on nuclear ribosomal DNA ITS sequences. Aliso 22: 372–395.

Friesen N., Fritsch R.M., Pollner S., Blattner F. R. 2000. Molecular and morphological evidence for an origin of the aberrant genus *Milula* within Himalayan species of *Allium* (Alliaceae). Mol. Phylogenet. Evol. 17: 209–218.

Fritsch R. M., 1992a. Infra-subgeneric grouping in subgenus *Melanocrommyum* (Webb et Berth) Rouy. In: Hanelt P., Hammer K., Knüpffer H. (eds) The Genus *Allium* - Taxonomic Problems and Genetic Resources (Proc. Int. Symp. Gatersleben, June 11-13, 1991). Gatersleben. pp. 67–75.

Fritsch, R.M., 1992b. Septal nectaries in the genus *Allium* – shape, position and excretory canals. In: Hanelt P., Hammer K., Knüpffer H. (eds) The Genus *Allium* - Taxonomic Problems and Genetic Resources (Proc. Int. Symp. Gatersleben, June 11-13, 1991). Gatersleben. pp. 77–85.

Fritsch R. 1993. Taxonomic and nomenclatural remarks on *Allium* L. subgen. *Melanocrommyum* (Webb & Berth.) Rouy section *Megaloprason* Wendelbo. Candollea 48: 417–430.

Fritsch R. M. 2000. Taxonomic revision of *Allium* L. section *Regeloprason* Wendelbo in Middle Asia. In: Ashurmetov O., Khassanov F., Salieva Y. (eds) Plant life in South-West and Central Asia (5th International Symposium, Tashkent '98). Chinor ENK, Tashkent. pp. 62–74.

Fritsch R. M., Astanova S. B. 1998. Uniform karyotypes in different sections of *Allium* L. subgen. *Melanocrommyum* (Webb & Berth.) Rouy from Central Asia. Feddes Repert. 109: 539–549.

Fritsch R. M., Friesen N. 2002. Evolution, domestication, and taxonomy. In: Rabinowitch H. D., Currah L. (eds) *Allium* Crop Science: Recent Advances, CABI Publishing, Wallingford, U.K. pp. 5–30.

Fritsch R. M., Keusgen M. 2006. Occurrence and taxonomic significance of cysteine sulphoxides in the genus *Allium* L. (Alliaceae). Phytochem. 67: 1127–1135.

Fritsch R. M., Khassanov F. O., Friesen N. W. 1998. New taxa, new combinations, and taxonomic remarks on *Allium* L. from Fergan depression, Middle Asia. Linzer Biol. Beitr. 30: 281–292 + 843.

Fritsch R. M., Khassanov F. O., Matin F. 2002. New *Allium* taxa from Middle Asia and Iran. Stapfia 80: 381–393.

Galbraith D. W., Harkins R. K., Maddox M. J., Ayres M. N., Sharma P. D., Firoozabady E. 1983. Rapid flow cytometry analysis of the cell cycle in intact plant tissues. Science 220: 1049–1051.

Gillespie J. H. 1991. The causes of molecular evolution. Oxford University Press, New York.

Gilley L. P. Taberlet 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. Mol. Biol. Evol. 11: 769–777.

Givnish T. J., Sytsma K. J. 1997. Homoplasy in molecular vs. morphological data: the likelihood of correct phylogenetic inference. In: Givnish T. J., Sytsma K. J. (eds) Molecular Evolution and Adaptive Radiation. Cambridge University Press. pp. 55–103.

Gold J. R., Amemiya C. T. 1987. Genome size variation in North American minnows (Cyprinidae). II. Variation among 20 species. Genome 29: 481–489.

Goldblatt P., Savolainen V., Porteous O., Sostaric I., Powell M., Reeves G., Manning J. C., Barraclough T. G., Chase M. W. 2002. Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. Mol. Phylogenet. Evol. 25: 341–360.

Govindaraju D. R., Dancik B. R., Wagner D. B. 1989. Novel chloroplast DNA polymorphism in a sympatric region of two pines. J. Evol. Biol. 2: 49–59.

Graham M. J., Nickell C. D., Rayburn A. L. 1994. Relationship between genome size and maturity group in soybean. Theor. Appl. Genet. 88: 429–432.

Graham S. W., Reeves P. A., Burns C. E., Olmstead R. G. 2000. Microstructural changes in noncoding DNA: Interpretation, evolution and utility of indels and inversions in basal angiosperm phylogenetic inference. Int. J. Plant Sci. 161: 83–96.

Graur D., Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet. 20: 80–86.

Gregory T. R. 2004. Macroevolution, hierarchy theory, and the C-value enigma. Paleobiology 30: 179–202.

Greilhuber J. 1998. Intraspecific variation in genome size: a critical reassessment. Ann. Bot. 82: 27–35.

Greilhuber J., Speta F. 1985. Geographical variation of genome size at low taxonomic levels in the *Scilla bifolia* alliance (Hyacinthaceae). Flora 176: 431–439.

Greilhuber J., Doležel J., Lysák M. A., Bennett M. D. 2005. The origin, evolution and proposed stabilization of the terms 'Genome Size'and 'C-Value' to describe nuclear DNA contents. Ann. Bot. 95: 255–260.

Greilhuber J., Borsch T., Müller K., Worberg A., Porembski S., Barthlott W. 2006. Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size. 8: 770–777.

Grundt H. H., Kjølner S., Borgen L., Rieseberg L. H., Brochmann C. 2006. High biological species diversity in the arctic flora. Proc. Natl. Acad. Sci. USA 103: 972–975.

Gurushidze M., Mashayekhi S., Blattner F. R., Friesen N., Fritsch R. M. 2007. Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. Plant Syst. Evol. 269: 259–269.

Gurushidze M., Fritsch R. M., Blattner F. R. 2008. Phylogenetic analysis of *Allium* subgenus *Melanocrommyum* infers cryptic species and demands a new sectional classification. Mol. Phylogenet. Evol. 49: 997–1007.

Hall A. E., Fiebig A., Preuss D. 2002. Beyond the *Arabidopsis* genome: opportunities for comparative genomics. Pl. Physiol. 129: 1439–1447.

Hanelt P., Fritsch R., Kruse J., Maaß H., Ohle H., Pistrick H. 1989. *Allium* L. section *Porphyroprason* Ekberg – Merkmale und systematische Stellung. Flora 182: 69–86.

Hanelt P., Schultze-Motel J., Fritsch R., Kruse J., Maaß H. I., Ohle H., Pistrick K. 1992. Infrageneric grouping of *Allium* - the Gatersleben approach. In: Hanelt P., Hammer K., Knüpffer H. (eds) The Genus *Allium* - Taxonomic Problems and Genetic Resources (Proc. Int. Symp. Gatersleben, June 11-13, 1991). Gatersleben. pp. 107–123.

Hansen A. K., Escobar L. K., Gilbert L. E., Jansen R. K. 2007. Paternal, maternal, and biparental inheritance of the chloroplast genome in *Passiflora* (Passifloraceae): implications for phylogenetic studies. Amer. J. Bot. 94: 42–46.

Harpke D., Peterson A. 2007. Quantitative PCR revealed a minority of ITS copies to be functional in *Mammillaria* (Cactaceae). Int. J. Pl. Sci. 168: 1157–1160.

Harris S. A., Ingram R. 1991. Chloroplast DNA and biosystematics: the effects of intraspecific diversity and plastid transmission. Taxon 40: 393–412.

Hastings W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57: 97–109.

Havey M. J. 1992. Restriction enzyme analysis of the chloroplast and nuclear 45S ribosomal DNA of *Allium* sections *Cepa* and *Phyllodolon* (Alliaceae). Plant Syst. Evol. 183: 17–31.

Hendy M. D., Penny D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38: 297–309.

Hennig W. 1953. Kritische Bemerkungen zum phylogenetischen System der Insekten. Beitr. Entomol. 3 (Sonderh.). 1-85.

Hennig W. 1966. Phylogenetic Systematics. Univ. Illinois Press, Urbana. 263 pp.

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

Hewitt G. M. 2001. Speciation, hybrid zones and phylogeography—or seeing genes in space and time. Mol. Ecol. 10: 537–549.

Hill T., Lewicki P. 2006. Statistics: methods and applications. A comprehensive reference for Science, Industry, and Data Mining. 1st ed. Tulsa: StatSoft.

Hillis D. M., Bull J. J. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42: 182–192.

Hipkens V. D., Marshall K. A., Neale D. B., Rottmann W. H., Strauss S. H. 1995. A mutation hotspot in the chloroplast genome of a conifer (Douglas fir: *Pseudotsuga*) is caused by variability in the number of direct repeats from a partially duplicated tRNA gene. Curr. Genet. 27: 527–579.

Hörandl E., Paun O., Johansson J. T., Lehnebach C., Armstrong T., Chen L., Lockhart P. 2005. Phylogenetic relationships and evolutionary traits in *Ranunculus* s. l. (Ranunculaceae) inferred from ITS sequence analysis. Mol. Phylog. Evol. 36: 305–327.

Hudson R. R. 1990. Gene genealogies and the coalescent process. In: Havey P. H., Partridge L. (eds) Oxford surveys in evolutionary biology, vol 7. New York, Oxford University Press. pp 1–44.

Huelsenbeck J. P., Ronquist F., Nielsen R., Bollback J. P. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310–2314.

Huelsenbeck J. P., Larget B., Miller R. E., Ronquist F. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51: 673–688.

Ingvarsson P. K., Ribstein S., Taylor D. R. 2003. Molecular evolution of insertions and deletions in the chloroplast genome of *Silene*. Mol. Biol. Evol. 20: 1737–1740.

Jakob S. S., Blattner F. R. 2006. A chloroplast genealogy of *Hordeum* (Poaceae): Long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. Mol. Biol. Evol. 23: 1602–1612.

Jakob S. S., Meister A., Blattner F. R. 2004. The considerable genome size variation in *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. Mol. Biol. Evol. 21: 860–869.

Jakob S. S., Ihlow A., Blattner F. R. 2007. Combined ecological niche modeling and molecular phylogeography revealed the evolutionary history of *Hordeum marinum* (Poaceae) – niche differentiation, loss of genetic diversity, and speciation in Mediterranean Quaternary refugia. Mol. Ecol. 16: 1713–1727.

Janssen T., Bremer K. 2004. The age of major monocot groups inferred from 800+ *rbcL* sequences. Bot. J. Linn. Soc. 146: 385–398.

Johnson L. B., Palmer J. D. 1989. Heteroplasmy of chloroplast DNA in *Medicago*. Plant Mol. Biol. 12: 3–11.

Johnston J. S., Pepper A. E., Hall A. E., Chen Z. J., Hodnett G., Drabek J., Lopez R., Price H. J. 2005. Evolution of Genome Size in Brassicaceae. Ann. Bot. 95: 229–235.

Kadereit J. W., Comes H. P. 2005. The temporal course of alpine plant diversification in the Quaternary. In: Bakker F. T., Chartou L.W., Gravendeel B., Pelser P. B. (eds) Plant species-level systematics: new perspectives on pattern and process. A.R.G Gantner Verlag, Liechtenstein. pp. 117–130.

Kalendar R., Tanskanen J., Immonen S., Nevo E., Schulman A. H. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc. Natl. Acad. Sci. USA 97: 6603–6607.

Kamelin R. V. 1973. Florogeneticheskij Analiz Estestvennoj Flory Gornoj Srednej Azii. Nauka, Leningrad. 237 pp.

Kamelin R. V., Seisums A. G. 1996. Tri novykh vidov roda *Allium* L. (Alliaceae) iz Jugo-zapadnoj Azii. Tres species novae generis *Allium* L. (Alliaceae) ex Asia austro-occidentalis. Novosti sist. vyssh. rast. 30: 29–33.

Kanno A., Hirai A. 1993. A transcription map of the chloroplast genome from rice (*Oryza sativa*). Curr. Genet. 23: 166–174.

Kay K. M., Reeves P. A., Olmstead R. G., Schemske D. W. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. Amer. J. Bot. 92: 1899–1910.

Kay K. M., Whittall J. B., Hodges S. A. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. BMC Evol. Biol. 6: 36 doi:10.1186/1471-2148-6-36.

Khassanov F. O. 1992. A revision of the genus *Allium* L. in the flora of Uzbekistan. In: Hanelt P., Hammer K., Knüpffer H. (eds) The Genus *Allium* - Taxonomic Problems and Genetic Resources (Proc. Int. Symp. Gatersleben, June 11-13, 1991). Gatersleben. pp. 153–159.

Khassanov F. O., Fritsch R. M. 1994. New taxa in *Allium* L. subgenus *Melanocrommyum* (Webb & Berth.) Rouy from Central Asia. Linzer Biol. Beitr. 26: 965–990.

Kimura M. 1968. Evolutionary rate at the molecular level. Nature 68: 624–626.

Kimura M., Ohta T. 1971. On the rate of molecular evolution. J. Mol. Evol. 1: 1–17.

Klak C., Reeves G., Hedderson T. 2004. Unmatched tempo of evolution in Southern African semi-desert ice plants. Nature 427: 63–65

Knight C. A., Ackerly D. D. 2002. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. Ecol. Lett. 5: 66–76.

Knight C. A., Molinari N. A., Petrov D. A. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. Ann. Bot. 95: 177-190.

Koch C. 1849. Beiträge zu einer Flora des Orientes. Linnaea 22: 177–336.

Koch M. A., Haubold B., Mitchell-Olds T. 2001. Molecular systematics of the Brassicaceae: evidence from coding plastid *mat*K and nuclear *Chs* sequences. Am J. Bot. 88: 534–544.

Koch M., Dobeš C., Matschinger M., Bleeker W., Vogel J., Kiefer M., Mitchell-Olds T. 2005. Evolution of the plastidic *trn*F(GAA) gene in *Arabidopsis* relatives and the Brassicaceae family: Monophyletic origin and subsequent diversification of a plastidic pseudogene. Mol. Biol. Evol. 22: 1032–1043.

Koch M., Dobeš C., Kiefer M., Schmickl R., Klimes L., Lysak M. A. 2007. Supernetwork identifies multiple events of plastidic *trn*F(GAA) pseudogene evolution in the Brassicaceae. Mol. Biol. Evol. 24: 63–73.

Leitch I. J., Bennett M. D. 1997. Polyploidy in angiosperms. Trends Plant Sci. 2: 470–476.

Leitch I. J., Chase M. W., Bennett M. D. 1998. Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. Ann. Bot. 82: 85–94.

Leitch I. J., Soltis D. E., Soltis P. S., Bennett M. D. 2005. Evolution of DNA amounts across land plants (Embryophyta). Ann. Bot. 95: 207–217.

Lewis P. 1998. Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data. In: Soltis D. E., Soltis P. S., Doyle J. J. (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer, Dordrecht. pp. 132–187

Li W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.

Li W.-H., Zharkikh A. 1994. What is the bootstrap technique? Syst. Biol. 43: 424–430.

Lidholm J., Szmidt A., Gustafsson P. 1991. Duplication of the *psb*A gene in the chloroplast genome of two *Pinus* species. Mol. Gen. Genet. 226: 345–352.

Lilly J. W., Havey M. J. 2001. Sequence analysis of a chloroplast intergenic spacer for phylogenetic estimates in *Allium* section *Cepa* and a PCR–based polymorphism detecting mixtures of male–fertile and male–sterile cytoplasmic onion. Theor. Appl. Genet. 102: 78–82.

Linder C. R., Rieseberg L. H. 2004. Reconstructing patterns of reticulate evolution in plants. Amer. J. Bot. 91: 1700–1708.

Linder H. P., Hardy C. R. 2005. Species richness in the Cape flora: a macroevolutionary and macroecological perspective. In: Bakker F. T., Chartou L. W., Gravendeel B., Pelser P. B. (eds) Plant species-level systematics: new perspectives on pattern and process. A.R.G Gantner Verlag, Liechtenstein. pp. 47–73.

Linder H. P., Hardy,C. R., Rutschmann F. 2005. Taxon sampling effects in molecular dating: an example from the African Restionaceae. Mol. Phylogenet. Evol. 35: 569–582.

Linné C. von 1753. Species Plantarum. Holmiae. pp. 294–302.

Linne von Berg G., Samoylov A., Klaas M., Hanelt P. 1996. Chloroplast DNA restriction analysis and the infrageneric grouping of *Allium* (Alliaceae). Plant Syst. Evol. 200: 253–261.

Löhne C., Borsch T. 2005. Molecular evolution and phylogenetic utility of the *pet*D group II intron: A case study in basal angiosperms. Mol. Biol. Evol. 22: 1–16.

Ma J., Devos K., Bennetzen J. 2004. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. Genome Res. 14: 860–869.

Maaß H. I., Klaas M. 1995. Infraspecific differentiation of garlic (*Allium sativum* L.) by isozyme and RAPD markers. Theor. Appl. Genet. 91: 89–97.

Maddison W. P. 1997. Gene trees in species trees. Syst. Biol. 46: 523–536.

Magallon S., Sanderson M. J. 2001. Absolute diversification rates in angiosperm clades. Evolution 55: 1762–1780.

Mahelka V., Suda J., Jarolímová V., Trávnícek P., Krahulec F. 2005. Genome size discriminates between closely related taxa *Elytrigia repens* and *E. intermdia* (Poaceae: Triticeae) and their hybrid. Folia Geobot. 40: 367–384.

Martin W., Stoebe B., Goremykin V., Hansmann S., Hasegawa M., Kowallik K. V. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393: 161–165.

Martínes-Ortega M. M., Delgado L., Albach D. C., Elena-Rossello J. A., Rico E. 2004. Species boundaries and phylogeographic patterns in cryptic taxa inferred from AFLP markers: *Veronica* subgen. *Pentasepalae* (Scrophulariaceae) in the Western Mediterranean. Syst. Bot. 29: 965–986.

Martins E. P. 2004. COMPARE, version 4.6b. Computer program for the statistical analysis of comparative data. http://compare.bio.indiana.edu/. Department of Biology, Indiana University, Bloomington, IN.

Martins E. P., Hansen T. F. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Amer. Naturalist 149: 646–667.

Mason-Gamer R. J., Holsinger K. E., Jansen R. K. 1995. Chloroplast DNA haplotype variation within and among populations of *Coreopsis grandiflora* (Asteraceae). Mol. Biol. Evol. 12: 371–381.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264: 421–424.

Mathews S., Donoghue M. J. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science 286: 947–950.

McKinnon G. E., Vaillancourt R. E., Jackson H. D., Potts B. M. 2001. Chloroplast sharing in the Tasmanian eucalypts. Evolution 55: 703–711.

McNeal D. W. 1992. Taxonomy of North American species of *Allium*. In: Hanelt P., Hammer K., Knüpffer H. (Eds.) The Genus *Allium* – Taxonomic Problems and Genetic Resources (Proc. Int. Symp. Gatersleben, June 11–13, 1991). IPK, Gatersleben, pp. 107–123.

Mes T. H. M., Friesen N., Fritsch R. M., Klaas M., Bachmann K. 1997. Criteria for sampling in *Allium* based on chloroplast DNA PCR-RFLP's. Syst. Bot. 22: 701–712.

Mes T. H. M., Fritsch R. M., Pollner S., Bachmann K. 1999. Evolution of the chloroplast genome and polymorphic ITS regions in *Allium* subgenus *Melanocrommyum*. Genome 42: 237–247.

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E. 1953. Equations of state calculations by fast computing machines. J. Chem. Phys. 21: 1087–1092.

Metzlaff M., Boerner T., Hagemann R. 1981. Variations of chloroplast DNA in the genus *Pelargonium* and their biparental inheritance. Theor. Appl. Genet. 60: 37–41.

Milne R. I., Abbott R. J. 2002. The origin and evolution of Tertiary relict floras. Adv. Bot. Res. 38: 281–314.

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

Mitchell-Olds T., Al-Shehbaz I. A., Koch M. A., Sharbel T. 2005. Crucifer evolution in the post-genomic era. In: Henry R. J. (ed) Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants. CAB International.

Moore M. J., Jansen R. K. 2006. Molecular evidence for the age, origin, and evolutionary history of the American desert plant genus *Tiquilia* (Boraginaceae). Mol. Phylogenet. Evol. 39: 668–687.

Morgan M. T. 2001. Transposable element number in mixed mating populations. Genetical Res. 77: 261–275.

Murphy W. J., Eizirik E., O'Brien S. J., Madsen O., Scally M., Douady C. J., Teeling E., Ryder O. A., Stanhope M. J., de Jong W. W., Springer M.S. 2001. Resolution of the early placental mammal radiation using Bayesian Phylogenetics. Science 294: 2348–2351.

Murray B. G. 2005. When does intraspecific C-value variation become taxonomically significant? Ann. Bot. 95: 119–125.

Nei M., Kumar S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, Oxford

Newton A. C., Allnutt T. R., Gillies A. C. M., Lowe A. J., Ennos R. A. 1999. Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends Ecol. Evol. 14: 140–145.

Nguyen N. H., Driscoll H. E., Specht C. D. 2008. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. Mol. Phylogenet. Evol. 45: 1157–1172.

Nie Z.–L., Sun H., Beardsley P. M., Olmstead R. G., Wen J. 2006. Evolution of biogeographic disjunction between eastern Asia and eastern North America in *Phyrma* (Phyrmaceae). Amer. J. Bot. 93: 1343–1356.

Obermayer R., Greilhuber J. 2005. Does genome size in *Dasypyrum villosum* vary with fruit colour? Heredity 95: 91–95.

Ohri D., Fritsch R. M., Hanelt P. 1998. Evolution of genome size in *Allium* (Alliaceae). Plant Syst. Evol. 210: 57–86.

Olmstead R. G., Palmer J. D. 1994. Chloroplast DNA systematics: A review of methods and data analysis. Amer. J. Bot. 81: 1205–1224.

Özhatay N. 1986. *Allium* species grown in Northern Anatolia and their chromosome number. Doga, TU Bio. D. 10: 452–458.

Özkan H., Levy A., Feldman M. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. Plant Cell 13: 1735–1747.

Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature 401: 877–884.

Palme A. E., Su Q., Palsson S., Lascoux M. 2004. Extensive sharing of chloroplast haplotypes among European birches indicates hybridization among *Betula pendula*, *B. pubescens* and *B. nana*. Mol. Ecol. 13: 167–178.

Palmer J. D. 1991. Plastid chromosomes: structure and evolution. In: Bogorad L., Vasil I. K. (eds) Cell culture and somatic cell genetics in plants, Vol. 7. The molecular biology of plastids. Academic Press, San Diego. pp. 5–53.

Palmer J. D., Nugent J. M., Herbon L. A. 1987. Unusual structure of geranium chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. Proc. Natl. Acad. Sci. USA 84: 769–773.

Pan J., Zhang D., Sang T. 2007. Molecular phylogenetic evidence for the origin of a diploid hybrid of *Paeonia* (Paeoniaceae). Amer. J. Bot. 94: 400–408.

Persson K., Wendelbo P. 1979. The taxonomic position of *Nectaroscordum koelzii* (Liliaceae). Bot. Notiser 132: 191–196.

Petrov D. A. 2001. Evolution of genome size: new approaches to an old problem. Trends Genet. 17: 23–28.

Petrov D. A. 2002. DNA loss and evolution of genome size in Drosophila. Genetica 115: 81–91.

Pirie M. D., Vargas M. P. B., Botermans M., Bakker F. T., Chatrou L. W. 2007. Ancient paralogy in the cpDNA *trn*L-F region in Annonaceae: implications for plant molecular systematics. Amer. J. Bot. 94: 1003–1016.

Posada D., Crandall K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817–818.

Posada D., Crandall K. A. 2001. Intraspecific gene genealogies: Trees grafting into networks. Trends Ecol. Evol. 16: 37–45.

Price H. J., Chambers K. L., Bachmann K. 1981. Genome size variation in diploid *Microseris bigelovii* (Asteraceae). Bot. Gaz. 142: 156–159.

Price H. J., Dillon S. L., Hodnett G., Rooney W. L., Ross L., Johnston J. S. 2005. Genome evolution in the genus *Sorghum* (Poaceae). Ann. Bot. 95: 219–227.

Provan J., Soranzo N., Wilson N. J., Goldstein B. D., Powell W. 1999. A low mutation rate for chloroplast microsatellites. Genetics 153: 943–947.

Quigley F., Weil J. H. 1985. Organization and sequence of five tRNA genes and of an unidentified reading frame in the wheat chloroplast genome: Evidence for gene rearrangements during evolution of chloroplast genomes. Curr. Genet. 9: 495–503.

Raven P. H., Axelrod D. I. 1978. Origin and relationships of the California flora. University of California Publications in Botany. 72: 1–134.

Rayburn A. L., Auger J. A., Benzinger E. S., Hepburn A. G. 1989. Detection of intraspecific DNA content variation in *Zea mays* L. by flow cytometry. J. Exp. Bot. 40: 1179–1183.

Reboud X., Zeyl C. 1994. Organelle inheritance in plants. Heredity 72: 132–140.

Reeves G., Francis D., Davies M. S., Rogers H. J., Hodkinson T. 1998. Genome size is negatively correlated with altitude in natural populations of *Dactylis glomerata*. Ann. Bot. 82: 99–105.

Regel E. 1875. Alliorum adhuc cognitorum monographia. Acta Horti Petropolitani 3: 1–266.

Renner S. S. 2005. Relaxed molecular clocks for dating historical plant dispersal events**.** Trends in Plant Sci. 10: 550–558.

Richardson J. E., Pennington R. T., Pennington T. D., Hollingsworth P. M. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293: 2242–2245.

Riede C. R., Fairbanks D. J., Andersen W. R., Kehrer R. L., Robison L. R. 1994. Enhancement of RAPD analysis by restriction-endonuclease digestion of template DNA in wheat. Plant Breed. 113: 254–257.

Rieseberg L. H. 1996. Homology among RAPD fragments in interspecific comparisons. Mol. Ecol. 5: 99–105.

Rieseberg L. H. 1998. Genetic mapping as a tool for studying speciation. In: Soltis D. E., Soltis P. S., Doyle J. J. (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer, Dordrecht. pp. 459–487.

Rieseberg L. H., Soltis D. E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5: 65–83.

Rieseberg L. H., van Vossen C., Desrochers A. 1995. Genomic reorganization accompanies hybrid speciation in wild sunflowers. Nature 375: 313–316.

Rieseberg L. H., Whitton J., Linder C. R. 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. Acta Bot. Neerl. 45: 243–262.

Rodman J., Price R. A., Karol K., Conti E., Sytsma K. J., Palmer J. D. 1993. Nucleotide sequences of the *rbc*L gene indicate monophyly of mustard oil plants. Ann. Missouri Bot. Gard. 80: 686–699.

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

Saitou N., Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Sanderson M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14: 1218–1231.

Sanderson M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19: 101–109.

Sanderson M. J. 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19: 301–302.

Savolainen V., Chase M. W., Morton C. M., Hoot S. B., Soltis D. E., Bayer C., Fay M. F., de Bruijn A., Sullivan S., Qiu Y-L. 2000. Phylogenetics of flowering plants based upon a combined analysis of plastid *atp*B and *rbc*L gene sequences. Syst. Biol. 49: 306–362.

Schmuths H., Meister A., Horres R., Bachmann K. 2004. Genome size variation among accessions of *Arabidopsis thaliana*. Ann. Bot. 93: 317–321.

Schneider H., Schuettpelz E., Pryer K. M., Cranfill R., Magallón L. R. 2004. Ferns diversified in the shadow of angiosperms. Nature 428: 553–557.

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

Schwarzbach A. E., Rieseberg L. H. 2002. Likely multiple origins of a diploid hybrid sunflower species. Mol. Ecol. 11: 1703–1717.

Seisums A. G. 1994. Podrod *Melanocrommyum* (Webb et Berth.) Rouy roda *Allium* L. (Monograficheskij obzor). Avtoreferat dissertatsii doktora biol. nauk. Riga. 28 pp.

Shaul S., Graur D. 2002. Playing chicken (*Gallus gallus*): Methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. Gene 300: 59–61.

Shopova M. 1972. Citoloshki istrazhuvajna na rodot *Allium* od florata na Makedonijha. Godishen Zbornik Biol., Skopje 24: 83–103.

Shore J. S., McQueen K. L., Little S. H. 1994. Inheritance of plastid DNA in the *Turnera ulmifolia* complex (Turneraceae). Amer. J. Bot. 81: 1636–1639.

Smarda P., Bures P. 2006. Intraspecific DNA content variability in *Festuca pallens* on different geographical scales and ploidy levels. Ann. Bot. 98: 665–678.

Smith S. E. 1989. Influence of parental genotype on plastid inheritance in *Medicago sativa*. J. Hered. 80: 214–217.

Sneath P. H. A., Sokal R. R. 1973. Numerical Taxonomy: the principles and practice of numerical classification. Freeman, San Francisco. 573 pp.

Soltis D. E., Soltis P. S. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis D. E., Soltis P. S., Doyle J. J. (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer, Dordrecht. pp. 1–42.

Soltis D. E., Soltis P. S. 2000. Contributions of plant molecular systematics to studies of molecular evolution. Plant Mol. Biol. 42: 45–75.

Soltis D. E., Soltis P. S. 2003a. The role of phylogenetics in comparative genetics. Pl. Physiol. 132: 1790–1800.

Soltis P. S., Soltis D. E. 2003b. The bootstrap in phylogeny reconstruction. Statist. Sci. 18: 256–267.

Soltis D. E., Soltis P. S., Bennett M. D., Leitch I. J. 2003. Evolution of genome size in the angiosperms. Amer. J. Bot. 90: 1596–1603.

Soltis D. E., Soltis P. S., Chase M. E., Mort M. E., Albach D. C., Zanis M., Savolainen V., Hahn W. H., Hoot S. B., Fay M. F., Axtell M., Swensen S. M., Nixon K. C., Farris J. S. 2000. Angiosperm phylogeny inferred from a combined data set of 18S rDNA, *rbcL,* and *atp*B sequences. Bot. J. Linn. Soc. 133: 381–461.

Soltis D.E., Albert V. A., Savolainen V., Hilu K., Qiu Y., Chase M. W., Farris J. S., Stefanović S., Rice D. W., Palmer J. D., Soltis P. S. 2004. Genome–scale data, angiosperm relationships, and 'ending incongruence': a cautionary tale in phylogenetics. Trends Plant Sci. 9: 477–483.

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.

Song K., Lu P., Tang K., Osborn T. C. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA 92: 7719–7723.

SPSS for Windows, Rel. 12.0.0. 2003. Chicago: SPSS Inc.

Stahlhut R., Park G., Petersen R., Ma W., Hylands P. 1999. The occurrence of the anti-cancer diterpene taxol in *Podocarpus gracilior* Pilger (Podocarpaceae). Biochem. Sys. Ecol. 27: 613–622.

Strauss S. H., Palmer J. D., Howe G. T., Doerksen A. H. 1988. Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. Proc. Natl. Acad. Sci. USA 85: 3898–3902.

Susuki Y., Glazko G. V., Nei M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA 99: 16138–16143.

Swofford D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, MA.

Swofford D. L., Olsen G. J., Waddell P. J., Hillis D. M. 1996. Phylogenetic Inference. In: Hiillis D. M., Moritz D., Mable B. K. (eds) Molecular Systematics, Sinauer Associates, Sunderland, Massachusetts. pp. 407-514.

Sytsma K. J. 1990. DNA and morphology: inference of plant phylogeny. Trends Ecol. Evol. 5: 104–110.

Taberlet P., Gielly L., Pautou G., Bouvet J. 1991. Universal primers for amplification of three noncoding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105–1109.

Tamura K., Dudley J., Nei M., Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596–1599.

Templeton A. R., Crandall K. A., Sing C. F. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619–633.

Terry R. G., Nowak R. S., Tausch R. J. 2000. Genetic variation in chloroplast and nuclear ribosomal DNA in Utah juniper (*Juniperus osteosperma*, Cupressaceae): evidence for interspecific gene flow. Amer. J. Bot. 87: 250–258.

The Angiosperm Phylogeny Group 1998. An ordinal classification for the families of flowering plants. Ann. Missouri Bot. Gard. 85: 531–553.

The Angiosperm Phylogeny Group 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141: 399–436.

The French–Italian Public Consortium for Grapevine Genome Characterization. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 463–467.

Tierney L. 1994. Markov chains for exploring posterior distributions. Ann. Stat. 4: 1701–1728.

Tiffney B. H. 1985a. Perspectives on the origin of the floristic similarity between Eastern Asia and Eastern North America. J. Arnold. Arbor. 66: 73–94.

Tiffney B. H. 1985b. The Eocene North Atlantic land bridge: Its importance in Tertiary and modern phylogeography of the Northern Hemisphere. J. Arnold. Arbor. 66: 243–273.

Traub H. P. 1968. The subgenera, sections and subsections of *Allium* L. Plant Life 24: 147–163.

Tscholokaschvili N. 1975. Ad cognitionem systematis generis *Allium*. Zam. Sist. Geogr. Rast. Tbilisi. 31: 36–54.

Tzagolova V. G. 1977. K sistematike sektsii *Molium* G. Don Kazakhstanskikh lukov. Bot. Mat. Gerb. Inst. Bot. Acad. Nauk Kaz. SSR 10: 13–14.

Tzagolova V. G. 1983. K voprosu o sistematicheskom polozhenii endemichnykh lukov Kazakhstana i ikh ekologo-geograficheskoj kharakteristike. - Bot. Mat. Gerb. Inst. Bot. Acad. Nauk Kaz. SSR 13: 23–27.

Vakhtina L. I. (1969) Sravnitel'no-kariologicheskoe issledovanie nekotorykh vidov luka sekcii *Molium* Don. Bot. Zh. 54: 143–153.

van Raamsdonk L. W. D., Ensink W., van Heusden A. W., Vrielink–van Ginkel M., Kik C. 2003. Biodiversity assessment based on cpDNA and crossability analysis in selected species of *Allium* subgenus *Rhizirideum*. Theor. Appl. Genet. 107: 1048–1058.

Vargas P., Morton C. M., Jury S. L. 1999. Biogeographic patterns in Mediterranean and Macaronesian species of *Saxifraga* (Saxifragaceae) inferred from phylogenetic analyses of ITS sequences. Amer. J. Bot. 86: 724–734.

Vijverberg K., Bachmann K. 1999. Molecular evolution of a tandemly repeated *trn*F(GAA) gene in the chloroplast genome of *Microseris* (Asteraceae) and the use of structural mutations in phylogenetic analysis. Mol. Biol. Evol. 16: 1329–1340.

von Bothmer R., Jacobsen N., Baden C., Jørgensen R. B., Linde-Laursen I. 1995. An Ecogeographical Study of the Genus *Hordeum* (second ed.), IPGRI, Rome.

Vvedensky A. I. 1935. Rod Luk - *Allium* L. In: Komarov V. L. (ed) Flora SSSR; 4. Izd. Akad. Nauk SSSR, Leningrad. pp. 112–280.

Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T., Sigiura M. 1994. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii.* Proc. Natl. Acad. Sci. USA 91: 9794–9798.

Webb P., Berthelot S. 1848. Histoire naturelle des Isles Canaries. – vol 3, Phytographia Canariensis sectio III, pp 137–252.

Webb T., Bartlein P. J. 1992. Global changes during the last three million years: climatic controls and biotic responses. Annu. Rev. Ecol. Syst. 23: 141–173.

Weiss-Schneeweiss H., Greilhuber J., G. M. Schneeweiss. 2005. Genome size evolution in holoparasitic *Orobanche* (Orobanchaceae) and related genera Amer. J. Bot. 93: 148–156.

Welsh J., McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucl. Acid Res. 18: 7213–7218.

Wen J. 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu. Rev. Ecol. Syst. 30: 421–455.

Wen J. 2001. Evolution of eastern Asian–Eastern North American biogeographic disjunctions: a few additional issues. Int. J. Plant Sci. 162: 117–122.

Wendel J. F., Doyle J. J. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis D. E., Soltis P. S., Doyle J. J. (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer, Dordrecht. pp. 265–296.

Wendel J. F., Schnabel A., Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA92: 280–284.

Wendel J. F., Cronn R. C., Johnston S., Price H. J. 2002. Feast and famine in plant genomes. Genetica 115: 37–47.

Wendelbo P. 1966. New taxa and synonyms in *Allium* and *Nectaroscordum* of SW Asia. Acta Horti Gotoburg 28: 15–55.

Wendelbo P. 1969. New subgenera, sections and species of *Allium*. Bot. Notiser 122: 25–37.

Wendelbo P. 1971. Alliaceae. In: Rechinger K. H. Flora Iranica. No. 76. Graz, Akademische Druck- und Verlagsanstalt. 100 pp.

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

Whitfield J. B., Lockhart P. J. 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22: 258–265.

Whittall J. B., Hellquist C. B., Schneider E. L., Hodges S. A. 2004. Cryptic species in an endangered pondweed community (*Potamogeton*, Potamogetonaceae) revealed by AFLP markers. Amer. J. Bot. 91: 2022–2029.

Williams J. G., Kubelik A. R., Livak K. J., Rafalski J. A., Tingey S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531–6535.

Wissemann V. 2003. Hybridization and the evolution of the nrITS spacer region. In: Sharma A. K., Sharma A. (eds) Plant Genome, Biodiversity and Evolution, vol. 1, part A. Science Publishers, Enfield, New Hampshire. pp. 57–71.

Wittzell H. 1999. Chloroplast DNA variation and reticulate evolution in sexual and apomictic sections of dandelions. Mol. Ecol. 8: 2023–2035.

Wolfe A. D., Liston A. 1998. Contributions of PCR-based methods to plant systematics and evolutionary Biology. In: Soltis D. E., Soltis P. S., Doyle J. J. (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer, Dordrecht. pp. 43–86.

Won H., Renner S. S. 2006. Dating dispersal and radiation in the gymnosperm *Gnetum* (Gnetales)—clock calibration when outgroup relationships are uncertain. Syst. Biol. 55: 610–622.

Wyman J., Laliberte S., Tremblay M. F. 1997. Nuclear DNAcontent variation in seeds from 22 half-sib families of jack pine (*Pinus banksiana*, Pinaceae). Amer. J. Bot. 84: 1351–1361.

Xiang Q.–Y., Soltis D. E., Soltis P. S., Manchester R., Crawford D. J. 2000. Timing the eastern Asian–eastern North American floristic disjunction: Molecular clock corroborates paleontological data. Mol. Phylogenet. Evol. 15: 462–472.

Xiang Q.–Y., Zhang W. H., Ricklefs R. E., Qian H., Cheng Z. D., Wen J., Li J. H. 2004. Regional differences in rates of plant speciation and molecular evolution: A comparison between eastern Asia and eastern North America. Evolution 58: 2175–2184.

Zachos J., Pagani M., Sloan L., Thomas E., Billups K. 2001. Trends, rhythms and aberrations in global climate 65 Ma to present. Science 292: 686–693.

Zhang L.–B., Comes H. P., Kadereit J. W. 2001. Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. Amer. J. Bot. 88: 2331–2345.

Zhang L.–B., Comes H. P., Kadereit J. W. 2004. The temporal course of Quaternary speciation in the European high mountain endemic *Primula* sect. *Auricula* (Primulaceae). Int. J. Pl. Sci. 165: 191–207.

Zharkikh A., Li W.-H. 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. II. Four taxa without a molecular clock. Mol. Biol. Evol. 9: 1119–1147.

Zhou S., Renner S. S., Wen J. 2006. Molecular phylogeny and inter– and intracontinental biogeography of Calycanthaceae. Mol. Phylogenet. Evol. 39: 1–15.

Zuckerkandl E., Pauling L. 1962. Molecular disease, evolution, and genic heterogeneity. In: Kasha M., Pullman B. (eds) Horizons in Biochemistry. Academic Press, New York. pp 189–225.

Abbreviations, Figure legends and Table captions

Abbreviations

ANOVA	analysis of variance
APG	angiosperm phylogeny group
BI	Bayesian inference
bs	bootstrap support
cp	chloroplast
DNA	Deoxyribonucleic acid
EECO	early Eocene climatic optimum
GTR	general time reversible
HT	haplotype
IR	inverted repeat
ITS	internal transcribed spacer
LBA	long branch attraction
MCMC	Markov chain Monte Carlo
ML	maximum likelihood
MMCO	middle Miocene climatic optimum
MP	maximum parsimony
Myr	million years
NHG	northern hemisphere glaciation
NJ	neighbor joining
NPRS	nonparametric rate smoothing
PCR	polymerase chain reaction
pg	picogram
PGLS	phylogenetic generalized least squares
PL	penalized likelihood
pp	posterior probability
RAPD	random amplified polymorphic DNA
*rbc*L	ribulose-1,5-bisphosphate carboxylase/oxygenase
RFLP	restriction fragment length polymorphism
SD	standard deviation
SE	standard error
TBE	Tris/Borate/EDTA (ethylene-diamine-tetra-acetic acid)
TBR	tree bisection reconnection
*trn*F	transfer RNA for Phenylalanine
*trn*L	transfer RNA for Leucine

Figure Legends

Fig. 1. Distribution map of *Allium* subgenus *Melanocrommyum.*

Fig. 2. Schematic representation of the cross-section of an ovary; α – septal nectary, β – excretory canal, γ – locule. In all remaining nectary types in subgenus *Melanocrommyum* excretory canals are not shown (Fig. 3), because they start in the lower part of the nectaries and are not visible on the cross sections of the ovaries.

Fig. 3(a – c). Phylogenetic tree resulting from a Bayesian analysis of the ITS sequences from species of subgenus *Melanocrommyum* and two outgroup species. The sequences of hitherto unnamed new taxa (see discussion) are abbreviated with Un. The numbers along the branches depict Bayesian posterior probabilities and bootstrap support values (%). One asterisk indicates posterior probabilities ≥ 0.99 and MP bootstrap values between 80 and 90%, while two asterisks indicate posterior probabilities of 1.0, and bootstrap values ≥ 90%. Branches present in the Bayesian tree and not present in the strict consensus tree of MP are indicated by gray double-lines. Schemes of the cross-sections of ovaries indicate molecular groups where taxa possess the same type of nectaries. Species for which I was not able to study nectary anatomy are indicated by "#". The current taxonomic (sectional) affiliations of species are shown on the right side, where shaded bars correspond to names of the respective sections.

Fig. 4. RAPD reaction profile (Ethidium-Bromide-stained agarose gel) of primer C 07. The dashed line separates the individuals belonging to ITS clusters 4 and 7. *Allium darwasicum* individuals from both ITS clusters have clearly different RAPD profiles, indicating that not only ITS sequences but large parts of the genomes are different.

Fig. 5. A part of the Neighbor-Joining tree of *trn*L-*trn*F sequences of 440 *Melanocrommyum* individuals calculated using K81uf+I distances. Zero-length branches at the base of the nodes of this tree correspond to the sequences (haplotypes) which are inferred as ancestral haplotypes (occupying the interior positions within the lineages) by TCS network.

Fig. 6. Bayesian tree of 73 chloroplast haplotypes and two outgroup species. The black bars show the nodes, which also exist in the strict consensus tree of MP. The numbers above branches depict the posterior probabilities from BI and bootstrap support from MP. The lineage numbering corresponds to the lineages in TCS network represented in Fig. 7.

Fig. 7. Statistical parsimony network of 80 chloroplast haplotypes found in subgenus *Melanocrommyum*. Circle sizes correspond to the number of taxa possessing the haplotype. Taxon names are abbreviated by the first three letters. Asterisks depict possible connections to the outgroup taxa (root), where the larger asterisk represents the more probable connection. Inferred haplotypes (not present in the data set) are depicted as black dots. Dashed grey lines represent less probable connections according to criteria suggested by Crandall and Templeton (1993). The haplotypes of the backbone network (1-73) are arbitrarily numbered. Extensions with lower case letters (a to c) indicate haplotype division resulting from variation at two mononucleotide repeats and one highly variable indel.

Fig. 8. A schematic representation of ITS phylogenetic tree from the chapter 2. The color codes correspond to the colors of respective chloroplast lineages from Fig. 7, and thus, show the distribution of chloroplast lineages among the "ITS groups". The taxa for which the *trn*L-*trn*F region could not be amplified are given in black.

Fig. 9. Phylogeny of *Allium* based on *rbc*L sequences and Bayesian analysis. The numbers above the branches depict posterior probabilities from BI, below the branches bootstrap supports from parsimony analyses are given. The taxonomic affiliations (at subgeneric level) of the groups are shown on the right side.

Fig. 10. A schematic representation of Asparagales *rbcL* tree including *Allium* sequences from this analyses. The names of the taxa are replaced by the family names, calibration points for the node ages, as well as node numbers are taken from the work of Anderson and Janssen (in press). The only difference with their study is the node 45-A, which does not have node number in their paper, and is referred as crown node age for Alliaceae-Amaryllidaceae clade (Anderson 2007). The resulting age from the present study for Alliaceae s. str. is shown on the scheme in Myr.

Fig. 11. Chronogram of *Allium* inferred by penalized likelihood (PL) based on the ITS tree from Friesen et al. (2006), which was calibrated with the root node using 76 (Pl) – 79 (NPRS) Myr as minimum and maximum ages (see text for details). The species names are replaced by the names of the subgenera according to Friesen et al. (2006).

Fig. 12. Chronogram of subgenus *Melanocrommyum* based on ITS data, calculated using penalized likelihood (PL). The age of the crown node of subgenus *Melanocrommyum* was taken from the dated *Allium* ITS tree (chronogram, Fig. 11) as a calibration point. Scale below is in million years.

Fig. 13. Flow-cytometric genome size measurements in subgenus *Melanocrommyum*. Examples of histograms of relative fluorescence intensity obtained after analysis of propidium-iodide stained nuclei of the species: *A. taeniopetalum*, *A. robustum*, *A. bakhtiaricum*, *A. pseudobodeanum* together with the internal standard *Vicia faba*.

Fig. 14. Phylogenetic tree resulting from a Bayesian analysis of ITS sequences of subgenus *Melanocrommyum* species and two outgroups. Gray double-lines indicate branches not present in the strict consensus (MP) tree. The numbers next to the species names depict 2C DNA values and standard deviation (SD), and the numbers along the branches show estimated ancestral genome sizes with standard errors (SE). Species for which genome sizes are not available, are depicted by **#,** the well-supported groups (according to Chapter 2) are indicated by Latin uppercase letters on the right side.

Fig. 15. Box plot displaying the distribution of 2C DNA values (Y – axis) across the defined phylogenetic clades (X – axis), resulting from ITS analyses. ■ represent the median, whiskers the standard deviation.

Table captions

Table 1. Comparison of methods of phylogeny reconstruction (modified from Soltis and Soltis 2003).

Table 2. Haplotype distribution among *Melanocrommyum* species.

Table 3. PCR and sequencing primers for *rbc*L.

Table 4. List of the accessions, ploidy level (wherever available) and genome size of the *Melanocrommyum* taxa used in this study.

Table 5. Genome size amount in phylogenetic clades.

Acknowledgments

It is a pleasure to thank the many people who provided various help during working on this thesis.

A very special thank you goes to my supervisor Frank Blattner for his enthusiasm, inspiration, patience, and great efforts to explain things clearly. Throughout my thesis-writing, he provided encouragement, sound advice and lots of good ideas.

I am deeply indebted to Dr. Reinhard Fritsch for his invaluable help in *Allium* taxonomy and productive discussions and criticism. Without his knowledge of species and collecting activities of *Alliums*, particularly the most beautiful ones (of course subgenus *Melanocrommyum*) this work could not be accomplished.

I am very grateful to Prof. Martin Röser at the University of Halle for agreeing to supervise my thesis.

I thank Dr. Nikolai Friesen not only for providing DNA samples for North American *Allium* species, but also for fruitful and friendly discussions.

Dr. Klaus Pistrick was always helpful and supportive in many *Allium* related topics. Besides, he and his family were organizers of various cultural activities in and around Gatersleben, which spiced up the life in a small village.

I would like to thank Dr. Peter Hanelt, who initiated the *Allium* work in Gatersleben. His interest in the situation in my home country (which is never "boring and banal") was often the reason of our discussions during his visits in the Vavilov-house of the IPK.

I can't help thanking all the members of taxonomy section for relaxed and interesting working atmosphere. My colleagues Sabine Jakob, Christina Baier, Nicolai Nürk, Enoch Achigan Dako, Thekla Pleines, Matthias Höffken, Jana Ekhvaia, Karl Schmid, Solomon Benor, Firouzeh Bordbar, Saeideh Mashayekhi, and Violetta Kotseruba were good companions at many scientific or nonscientific occasions. Sabine Jakob, Nicolai Nürk, and Christina Baier kindly helped me in translating the abstract of the thesis into German. I am especially thankful to Matthias Höffken for his various and patient assistance with command line programs, and explaining to me many kinds of "distributions" and statistical tests.

Petra Oswald, Birgit Wohlbier, Ina Faustman, Barbara Sperling, Christina Koch, Brunhilde Wedemeier provided expert technical assistance, and kept a nice working atmosphere in the lab.

Furthermore, my thanks go to Dr. Jörg Fuchs for his excellent cooperation in genome size measurements and constructive criticism of the earlier version of the "Genome size chapter".

I am indebted to Prof. Maia Akhalkatsi and Prof. Gia Nakhutsrishvili for giving me possibility to work at the Tbilisi Institute of Botany and Botanical Garden, in the collaborative project with the IPK-Gatersleben and the University of Marburg.

Financial support from the IPK Gatersleben and VolkswagenStiftung (Hannover, Germany) as a part of "PharmAll" project is gratefully acknowledged. I am thankful to be a participant of the first Kazbegi workshop on "Botany, taxonomy and phytochemistry of wild *Allium* L. species of the Caucasus and Central Asia" supported by VolkswagenStiftung.

Many thanks go to the PhD Student Board of the IPK for organizing scientific, educational, cultural, and leisure events, which made my life far from close friends easier and more interesting.

Last but not least I would like to express my gratitude to my family for the support they provided me through my entire life, and especially to my son for his patience and encouragements during my thesis-writing time.

Curriculum Vitae

Personal Information

Name:	Maia Gurushidze
Date of Birth:	11.09.1976
Place of Birth:	Kvareli, Georgia
Address:	Finkenweg 1 D-06466 Gatersleben
e-mail:	maiagur1@yahoo.com

Education:

1986-1993	Shilda Secondary School (with distinction), Kvareli district, Georgia
1993-1995	Telavi I. Gogebashvili Pedagogical Institute (Georgia), Faculty of Biology and Geography
1995-1999	Tbilisi I. Javakhishvili State University (Georgia), Faculty of Medicine and Biology Bachelor degree in Biology (with honors), Qualification of teacher of Biology and Chemistry
1999-2001	Tbilisi I. Javakhishvili State University Master degree in Biology
2001-2003	Teaching assistant at the Tbilisi I. Javakhishvili State University
2003-2004	Research assistant at the N. Ketskhoveli Institute of Botany, (*current name:* Tbilisi Institute of Botany and Botanical Garden), Tbilisi, Georgia
Sept.-Dec. 2004	Research internship at the Institute of Plant Genetics and Crop Research (IPK Gatersleben, Germany)
March 2005-till present	PhD (doctoral) student at the IPK Gatersleben. The title of the Thesis: Phylogenetic relationships and diversification processes in *Allium* subgenus *Melanocrommyum*. Supervisor: Dr. Frank Blattner.

Publications (peer reviewed)

Gurushidze M., Fritsch R. M., Blattner F. R. 2008. Phylogenetic analysis of *Allium* subgenus *Melanocrommyum* infers cryptic species and demands a new sectional classification. Mol. Phylogenet. Evol. 49: 997–1007.

Gurushidze M., Mashayekhi S., Blattner F. R., Friesen N., Fritsch R. M. 2007. Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. Plant Syst. Evol. 269: 259–269.

Fritsch R.M. **Gurushidze M.,** Jedelská J., Keusgen M. 2006. More than a pretty face - ornamental "drumstick onions" of *Allium* subg. *Melanocrommyum* are also potential medicinal plants. Herbertia 60: 26–59

Gurushidze M., Abramidze T., Klimiashvili T., Dzidziguri D. 2004. Comparative study of hepatocyte transcriptional activity at the early stage of reparative growth in phylogenetically distant organisms. Proc. Georg. Acad. Sci. Tbilisi.

Modebadze I., Shavlakadze T., **Gurushidze M.,** Kokrashvili Z., Dziziguri D. 2001. Growth controlling gene expression in the rat hepatocytes under conditions of hormonal disbalance. Bull. Georg. Acad. Sci. Tbilisi. 164: 524-536.

Other publications

Gurushidze M. 2008. Phylogenetic Analysis Reveals Multiple Independent Origins of Dithiodipyrrole Containing Species of *Allium* Subgenus *Melanocrommyum.* In: Keusgen M., Fritsch R. M. (eds) Proceedings of Wokshop "Botany, taxonomy and Phytochemistry of wild Allim L. species of the Caucasus and Central Asia". pp. 85-91.

Fritsch R. M., **Gurushidze M.,** Jedelská J., Keusgen M. 2007. Ornamental "Drumstick Onions" of *Allium* Subgenus *Melanocrommyum* Used as Medicinal Plants in Southwest and Central Asia. Planta Med. 73: P. 138.

Lectures

Gurushidze M.: Evolution, Diversification and Biogeographic Patterns in *Allium*. Systematics 2008, April 07-11, 2008. Göttingen, Germany.

Gurushidze M.: Phylogeny of *Allium* subgenus *Melanocrommyum* – evidence from molecular data. International Symposium "Plant life of South West Asia" June 25-29, 2007. Eskisehir, Turkey.

Gurushidze M.: Phylogeny of *Allium* subgenus *Melanocrommyum* and taxonomic position of red dye containing species. First Kazbegi workshop on „Botany, taxonomy and Phytochemistry of wild *Allim* L. species of the Caucasus and Central Asia" June 4-8, 2007. Kazbegi, Georgia.

Gurushidze M., Blattner F. R., Fritsch R. M.: Congruent molecular data-sets contradicting morphology – phylogenetic pattern or uniparental inheritance? IPK Student Conference, May 30-June 1, 2006. Gatersleben, Germany.

Gurushidze M.: Why are nuclear and chloroplast phylogenies incompatible in *Allium* subg. *Melanocrommyum*?" IPK Student Conference, June 22-25, 2005. Gatersleben, Germany.

Posters

Gurushidze M., Blattner F. R.: Molecular markers infer cryptic species within *Allium* subg. *Melanocrommyum*. Botany 2008 – annual meeting of Botanical Society of America, July 26–30 2008. University of British Columbia, Vancouver, Canada.

Gurushidze M., Blattner F. R., Fritsch R. M.: Congruent nuclear and chloroplast phylogenies strongly contradict morphology in *Allium* subgenus *Melanocrommyum*. Botany 2006 – annual meeting of Botanical Society of America, July 28-August 2, 2006. Chico, USA.

Gurushidze M., Fritsch R. M., Blattner F. R.: Congruent molecular data sets *vs.* contradicting morphology in *Allium*: phylogenetic pattern or hybridization and uniparental inheritance? 17th International Symposium-Biodiversity and Evolutionary Biology of the German Botanical Society, September 24-28, 2006. Bonn, Germany.

Awards and fellowships:

2000	Second prize for the best scientific contribution at the Student and Postgraduate Student Conference of Tbilisi State University.
2001	Second prize at the International Soros Science Education Program.
2006-2007	*Volksvagenstiftung* Scholarship, research project on Pharmaceutical values of onions and related species of Middle Asia and the Caucasus.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Arbeit bisher weder der Naturwissenschaftlichen Fakultät der Martin-Luther-Universität Hale Wittenberg noch irgendeiner anderen wissenschaftlichen Einrichtung zum Zweck der Promotion vorgelegt wurde.

Ich erkläre, dass ich mich bisher noch nie um einen Doktorgrad beworben habe.

Ferner erkläre ich, dass ich nie die vorliegende arbeit selbständig and ohne fremde Hilfe verfasst habe. Es wurden keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, und die den benutzten Werken wörtlich oder inhaltlich entnommen Stellen wurden als solche kenntlich gemacht.

Halle/Saale, den 28. 11. 2008

Maia Gurushidze

Appendices

Table 1. Species, accession numbers (Gatersleben collection), taxonomic affiliations, and the origin of the plant material used in the ITS analyses of the subgenus *Melanocrommyum* (Chapter 2)

species	subspecies	Taxonomy (section)	accession number	origin or/and source
Allium aflatunense B. Fedtsch.		Megaloprason	5562	Kyrgyzstan, Olomouc collection
Allium aflatunense B. Fedtsch.		Megaloprason	5632	Kyrgyzstan
Allium aflatunense B. Fedtsch.		Megaloprason	5694	Kyrgyzstan
Allium akaka S.G. Gmelin ex Schult. et Schult. f.		Acanthoprason	1138I	Iran
Allium akaka S.G. Gmelin ex Schult. et Schult. f.		Acanthoprason	5967	Turkey
Allium akaka S.G. Gmelin ex Schult. et Schult. f.		Acanthoprason	6397	Iran
Allium akaka S.G. Gmelin ex Schult. et Schult. f.		Acanthoprason	6398	Iran
Allium akaka S.G. Gmelin ex Schult. et Schult. f.		Acanthoprason	5980	Turkey
Allium alaicum Vved.		Acmopetala	5676	Kyrgyzstan
Allium alexeianum Regel		Kaloprason	2557	Tajikistan
Allium alexeianum Regel		Kaloprason	6303	Tajikistan
Allium alexeianum Regel		Kaloprason	6528	Uzbekistan
Allium altissimum Regel		Megaloprason	2976	Kazakhstan
Allium altissimum Regel		Megaloprason	6357	Uzbekistan
Allium arkitense R.M. Fritsch		Acmopetala	5692	Kyrgyzstan
Allium aroides Popov & Vved.		Aroidea	6189	Uzbekistan
Allium aroides Popov & Vved.		Aroidea	2517	Uzbekistan, Tashkent Botanical Garden
Allium aroides Popov & Vved.		Aroidea	3703	Uzbekistan
Allium aschersonianum W. Barbey in C. & W. Barbey		Melanocrommyum	5531	Israel
Allium assadii Seisums		Acmopetala	3927	Iran
Allium atropurpureum Waldst. et Kit.		Melanocrommyum	0799	Jena Botanical Garden
Allium atropurpureum Waldst. et Kit.		Melanocrommyum	1017	private collection Dr. Wandelt, Quedlinburg
Allium atropurpureum Waldst. et Kit.		Melanocrommyum	2194	Bulgaria
Allium backhousianum Regel		Acmopetala	2680	Kyrgyzstan
Allium backhousianum Regel		Acmopetala	2681	Kyrgyzstan

Allium backhousianum Regel		Acmopetala	5680	Kyrgyzstan
Allium bakhtiaricum Regel		Megaloprason	3924	Iran
Allium brachyscapum Vved.		Acanthoprason	5922	Turkmenistan
Allium breviscapum Stapf		Acanthoprason	6478	Iran
Allium bucharicum Regel		Kaloprason	6126	Tajikistan
Allium bucharicum Regel		Kaloprason	6114	Tajikistan
Allium cardiostemon Fischer et Meyer		Pseudoprason	3947	Iran
Allium cardiostemon Fischer et Meyer		Pseudoprason	5978	Turkey
Allium cardiostemon Fischer et Meyer		Pseudoprason	6081	Armenia
Allium caspium (Pall.) M. Bieb.	*baissunense* (Lipsky) Khassanov et R.M. Fritsch	Kaloprason	5873	Uzbekistan
Allium caspium (Pall.) M. Bieb.	*baissunense* (Lipsky) Khassanov et R.M. Fritsch	Kaloprason	6187	Uzbekistan
Allium caspium (Pall.) M. Bieb.	*caspium*	Kaloprason	5927	Turkmenistan
Allium caspium (Pall.) M. Bieb.	*caspium*	Kaloprason	5936	Turkmenistan
Allium chelotum Wendelbo		Melanocrommyum	6236	Iran
Allium chelotum Wendelbo		Melanocrommyum	I323	Iran
Allium chelotum Wendelbo		Melanocrommyum	I324	Iran
Allium chitralicum Wang et Tang		Brevicaule	6153	Tajikistan
Allium chodsha-bakirganicum Gaffarov et Turakulov		Regeloprason	3787	Kyrgyzstan
Allium chrysantherum Boiss.et Reut.		Melanocrommyum	2E	Iran
Allium costatovaginatum Kamelin et Levichev ex Krassovskaja et Levichev		Acmopetala	5780	Uzbekistan
Allium costatovaginatum Kamelin et Levichev ex Krassovskaja et Levichev		Acmopetala	6364	Uzbekistan
Allium costatovaginatum Kamelin et Levichev ex Krassovskaja et Levichev		Acmopetala	5783	Uzbekistan
Allium cristophii Trautv.		Kaloprason	1920	Manchester Botanic Garden
Allium cristophii Trautv.		Kaloprason	6246	Iran
Allium cristophii Trautv.		Kaloprason	5920	Turkmenistan
Allium cristophii Trautv.		Kaloprason	5262	Turkmenistan
Allium cristophii Trautv. ***s.l.***		Kaloprason	6102	Iran
Allium cupuliferum Regel	*cupuliferum*	Regeloprason	5808	Uzbekistan

Allium cupuliferum Regel	*cupuliferum*	Regeloprason	6196	Uzbekistan
Allium cupuliferum Regel	*cupuliferum*	Regeloprason	6350	Uzbekistan
Allium cupuliferum Regel	*cupuliferum*	Regeloprason	6356	Uzbekistan
Allium cyrilli Tenore		Melanocrommyum	5296	Turkey
Allium cyrilli Tenore		Melanocrommyum	1150	Leipzig Botanical Garden
Allium darwasicum Regel		Regeloprason	1625	Tajikistan
Allium darwasicum Regel		Regeloprason	2537	Tajikistan
Allium darwasicum Regel		Regeloprason	6431	Tajikistan
Allium darwasicum Regel		Regeloprason	6138	Tajikistan
Allium darwasicum Regel		Regeloprason	6134	Tajikistan
Allium darwasicum Regel		Regeloprason	6144	Tajikistan
Allium darwasicum Regel		Regeloprason	6143	Tajikistan
Allium dasyphyllum Vved.		Acmopetala	5027	Kyrgyzstan
Allium decipiens Fischer ex Schult. et Schult. f.		Melanocrommyum	2709	Ukraina
Allium decipiens Fischer ex Schult. et Schult. f.		Melanocrommyum	5837	Ukraina
Allium derderianum Regel		Acanthoprason	6390	Iran
Allium derderianum Regel		Acanthoprason	6516	Iran
Allium derderianum Regel		Acanthoprason	114G	Iran
Allium dodecadontum Vved.		Acmopetala	5695	Kyrgyzstan
Allium elburzense Wendelbo		Acanthoprason	6392	Iran
Allium ellisii J.D.Hooker		Kaloprason	6254	Iran
Allium ellisii J.D.Hooker		Kaloprason	6249	Iran
Allium fetisowii Regel		Acmopetala	3693	Kazakhstan
Allium fetisowii Regel		Acmopetala	5685	Kyrgyzstan
Allium fetisowii Regel		Acmopetala	5688	Kyrgyzstan
Allium fetisowii Regel		Acmopetala	5052	Kyrgyzstan
Allium fetisowii Regel		Acmopetala	2975	Kazakhstan
Allium giganteum Regel		Compactoprason	6188	Uzbekistan
Allium giganteum Regel		Compactoprason	6263	Iran
Allium giganteum Regel		Compactoprason	6446	Tajikistan
Allium gypsaceum Popov et Vved. ex Vved.		Popovia	3661	Uzbekistan
Allium gypsaceum Popov et Vved. ex Vved.		Popovia	6185	Uzbekistan
Allium helicophyllum Vved.		Kaloprason	5272	Turkmenistan

Allium helicophyllum Vved.		Kaloprason	6240	Iran
Allium hexaceras Vved.		Acaule	119G	Uzbekistan
Allium hissaricum Vved.		Regeloprason	2935	Tajikistan
Allium hissaricum Vved.		Regeloprason	2947	Tajikistan
Allium hissaricum Vved.		Regeloprason	3129	Tajikistan
Allium hissaricum Vved.		Regeloprason	6157	Tajikistan
Allium hollandicum R.M.Fritsch		Megaloprason	1631	private collection Dr. Fritsch, Gatersleben
Allium hollandicum R.M.Fritsch		Megaloprason	2615	Botanic Garden Vrije Univ. Amsterdam
Allium hollandicum R.M.Fritsch		Megaloprason	2800	from Bundesgartenschau (German Federal Garden Show) Frankfurt/Main
Allium insufficiens Vved.		Acmopetala	6119	Tajikistan
Allium isakulii R.M. Fritsch et F.O. Khass.	*balkhanicum* R.M. Fritsch et Khassanov	Regeloprason	5264	Turkmenistan
Allium isakulii R.M. Fritsch et F.O. Khass.	*balkhanicum* R.M. Fritsch et Khassanov	Regeloprason	5931	Turkmenistan
Allium isakulii R.M. Fritsch et F.O. Khass.	*subkopetdagense* R.M. Fritsch et Khassanov	Regeloprason	5278	Turkmenistan
Allium isakulii R.M. Fritsch et F.O. Khass.	*subkopetdagense* R.M. Fritsch et Khassanov	Regeloprason	5919	Turkmenistan
Allium isakulii R.M. Fritsch et F.O. Khass.	*nuratense* (Kamelin) R.M. Fritsch et F.O. Khass.	Regeloprason	111G	Uzbekistan
Allium jesdianum Boiss. et Buhse	*angustitepalum* (Wendelbo) Khassanov et R.M. Fritsch	Megaloprason	3666	Uzbekistan
Allium jesdianum Boiss. et Buhse	*angustitepalum* (Wendelbo) Khassanov et R.M. Fritsch	Megaloprason	1083	Afghanistan, Göteborg Botanical Garden
Allium jesdianum Boiss. et Buhse	*angustitepalum* (Wendelbo) Khassanov et R.M. Fritsch	Megaloprason	3671	Uzbekistan
Allium jesdianum Boiss. et Buhse	*jesdianum*	Megaloprason	3951	Iran
Allium jesdianum Boiss. et Buhse	*jesdianum*	Megaloprason	3953	Iran
Allium jesdianum Boiss. et Buhse	*jesdianum*	Megaloprason	6261	Iran
Allium karataviense Regel		Miniprason	3676	Kazakhstan
Allium karataviense Regel		Miniprason	5713	Tajikistan
Allium karataviense Regel		Miniprason	6133	Tajikistan
Allium karataviense Regel		Miniprason	6147	Tajikistan
Allium koelzii (Wendelbo) K.Persson et Wendelbo		Pseudoprason	1122I	Iran

Allium koelzii (Wendelbo) K.Persson et Wendelbo		Pseudoprason	6501	Iran
Allium koelzii (Wendelbo) K.Persson et Wendelbo		Pseudoprason	6509	Iran
Allium komarowii Lipsky		Compactoprason	3144	Uzbekistan
Allium komarowii Lipsky		Compactoprason	6296	Tajikistan
Allium komarowii Lipsky		Compactoprason	6530	Uzbekistan
Allium kuhsorkhense R.M. Fritsch et Joharchi		Acanthoprason	6262	Iran
Allium kujukense Vved.		Vvedenskya	3625	Kazakhstan
Allium lipskyanum Vved.		Regeloprason	1384	Tajikistan, Dushanbe Botanical Garden
Allium lipskyanum Vved.		Regeloprason	3118	Tajikistan
Allium lipskyanum Vved.		Regeloprason	5076	Uzbekistan
Allium macleanii J. G. Baker		Compactoprason	2415	Tajikistan, Chorog Botanical Garden
Allium macleanii J. G. Baker		Compactoprason	6150	Tajikistan
Allium macleanii J. G. Baker		Compactoprason	6445	Tajikistan
Allium majus Vved.		Compactoprason	3362	Uzbekistan
Allium materculae Bordz.		Acanthoprason	3948	Iran
Allium materculae Bordz.		Acanthoprason	6402	Iran
Allium materculae Bordz.		Acanthoprason	6406	Iran
Allium minutiflorum Regel		Acanthoprason	6	Iran (Herbarium Tabriz)
Allium motor Kamelin		Acmopetala	3355	Uzbekistan
Allium motor Kamelin		Acmopetala	6208	Uzbekistan
Allium motor Kamelin		Acmopetala	6369	Uzbekistan
Allium nevskianum Vved. ex Wendelbo		Kaloprason	2262	Tajikistan
Allium nevskianum Vved. ex Wendelbo		Kaloprason	2269	Tajikistan
Allium nevskianum Vved. ex Wendelbo		Kaloprason	5878	Uzbekistan
Allium nigrum L.		Melanocrommyum	0515	Leipzig Botanical Garden
Allium nigrum L.		Melanocrommyum	1650	Leipzig Botanical Garden
Allium nigrum L.		Melanocrommyum	5321	Turkey
Allium nigrum L.		Melanocrommyum	2616	Botanical Garden Vrije Univ. Amsterdam
Allium noëanum Reut ex Regel		Melanocrommyum	6490	Iran
Allium oreophilum C. A. Meyer		Porphyroprason	0348	Botanic Garden Univ. Graz
Allium orientale Boiss.		Melanocrommyum	5352	Turkey
Allium orientale Boiss.		Melanocrommyum	5365	Turkey
Allium orientale Boiss.		Melanocrommyum	5371	Turkey

Allium protensum Wendelbo		Kaloprason	2162	Rostock Botanical Garden
Allium protensum Wendelbo		Kaloprason	3335	Uzbekistan
Allium protensum Wendelbo		Kaloprason	3672	Uzbekistan
Allium pseudobodeanum R.M. Fritsch et Matin		Acanthoprason	6519	Iran
Allium pseudobodeanum R.M.Fritsch et Matin		Kaloprason	3936	Iran
Allium pseudobodeanum R.M.Fritsch et Matin		Kaloprason	3932	Iran
Allium pseudobodeanum R.M.Fritsch et Matin		Kaloprason	6388	Iran
Allium pseudowinklerianum R.M. Fritsch et Khassanov		Regeloprason	5679	Kyrgyzstan
Allium pseudowinklerianum R.M. Fritsch et Khassanov		Regeloprason	112G	Kyrgyzstan
Allium regelii Trautv.		Regeloprason	5917	Turkmenistan
Allium regelii Trautv.		Regeloprason	5937	Turkmenistan
Allium robustum Kar.et Kir.		Melanocrommyum	4I	Kazakhstan
Allium rosenbachianum Regel	*kwakense* R. M. Fritsch	Megaloprason	2541	Tajikistan
Allium rosenbachianum Regel	*kwakense* R. M. Fritsch	Megaloprason	6159	Tajikistan
Allium rosenbachianum Regel		Megaloprason	5614	Tajikistan, Dushanbe Botanical Garden
Allium rosenbachianum Regel	*rosenbachianum*	Megaloprason	6120	Tajikistan
Allium rosenbachianum Regel	*rosenbachianum*	Megaloprason	6132	Tajikistan
Allium rosenbachianum Regel	*kwakense* R. M. Fritsch	Megaloprason	6461	Tajikistan
Allium rosenorum R. M. Fritsch		Megaloprason	2938	Tajikistan
Allium rosenorum R. M. Fritsch		Megaloprason	6158	Tajikistan
Allium rosenorum R. M. Fritsch		Megaloprason	1869	Tajikistan
Allium rosenorum R. M. Fritsch		Megaloprason	5081	Uzbekistan
Allium rosenorum R. M. Fritsch		Megaloprason	6290	Tajikistan
Allium rothii Zucc.		Melanocrommyum	6340	Jordan
Allium rothii Zucc.		Melanocrommyum	6342	Jordan
Allium rothii Zucc.		Melanocrommyum	6345	Jordan
Allium saposhnikovii E. Nikitina		Acmopetala	3655	Kazakhstan
Allium saposhnikovii E. Nikitina		Acmopetala	5686	Kyrgyzstan
Allium sarawschanicum Regel		Megaloprason	1323	Tajikistan
Allium sarawschanicum Regel		Megaloprason	1326	Tajikistan
Allium sarawschanicum Regel		Megaloprason	2946	Tajikistan

Allium sarawschanicum Regel		Megaloprason	6259	Iran
Allium schachimardanicum Vved.		Acmopetala	5066	Uzbekistan
Allium schubertii Zucc.		Kaloprason	3024	private collection Dr. Wiering, Bergen
Allium schugnanicum Vved.		Megaloprason	6151	Tajikistan
Allium severtzovioides R.M. Fritsch		Acmopetala	5043	Kazakhstan
Allium severtzovioides R.M. Fritsch		Acmopetala	5789	Uzbekistan
Allium severtzovioides R.M. Fritsch		Acmopetala	5879	Uzbekistan
Allium sewerzowii Regel		Acmopetala	5047	Kazakhstan
Allium sewerzowii Regel		Acmopetala	5844	Kazakhstan
Allium shelkovnikovii Grossh.		Acanthoprason	6404	Iran
Allium shelkovnikovii Grossh.		Acanthoprason	6405	Iran
Allium sochense R.M. Fritsch et U. Turakulov		Regeloprason	113G	Kyrgyzstan
Allium sp. 1			Un8	Iran (Herbarium Tabriz)
Allium sp. 2			Un9	Iran (Herbarium Tabriz)
Allium sp. 3			Un10	Iran (Herbarium Tabriz)
Allium sp. 4 "intradarvazicum"		Regeloprason	6450	Tajikistan
Allium sp. 4 "intradarvazicum"		Regeloprason	6452	Tajikistan
Allium sp. 4 "intradarvazicum"		Regeloprason	6454	Tajikistan
Allium sp. 5 "hamedanense"		Melanocrommyum	6487	Iran
Allium sp. 6 "ubipetrense"		Acanthoprason	6475	Iran
Allium sp. 6 "ubipetrense"		Acanthoprason	6503	Iran
Allium sp. 6 "ubipetrense"		Acanthoprason	6507	Iran
Allium sp. 7 "chychkanense"		Acmopetala	5057	Kyrgyzstan
Allium sp. 7 "chychkanense"		Acmopetala	5060	Kyrgyzstan
Allium sp. 8 "saralicum"		Melanocrommyum	6506	Iran
Allium sp. 8 "saralicum"		Melanocrommyum	3	Iran (Herbarium Tabriz)
Allium sp. 8 "saralicum"		Melanocrommyum	5	Iran (Herbarium Tabriz)
Allium stipitatum Regel		Megaloprason	6370	Uzbekistan
Allium stipitatum Regel		Megaloprason	1311	Tajikistan
Allium stipitatum Regel		Megaloprason	5263	Turkmenistan
Allium stipitatum Regel		Megaloprason	6502	Iran
Allium stipitatum Regel		Megaloprason	3967	Iran
Allium suworowii Regel		Acmopetala	3121	Tajikistan

Allium suworowii Regel		Acmopetala	5791	Uzbekistan
Allium suworowii Regel		Acmopetala	6141	Tajikistan
Allium suworowii Regel		Acmopetala	6161	Tajikistan
Allium taeniopetalum Popov & Vved	*mogoltavicum* (Vved.) R.M. Fritsch et Khassanov	Acmopetala	5910	Uzbekistan
Allium taeniopetalum Popov & Vved	*mogoltavicum* (Vved.) R.M. Fritsch et Khassanov	Acmopetala	6169	Uzbekistan
Allium taeniopetalum Popov & Vved	*mogoltavicum* (Vved.) R.M. Fritsch et Khassanov	Acmopetala	6372	Uzbekistan
Allium taeniopetalum Popov & Vved	*taeniopetalum*	Acmopetala	5244	Uzbekistan
Allium taeniopetalum Popov & Vved	*turakulovii* R. M. Fritsch et Khassanov	Acmopetala	5068	Kyrgyzstan
Allium tashkenticum Khassanov et R. M. Fritsch		Acmopetala	5788	Uzbekistan
Allium tashkenticum Khassanov et R. M. Fritsch		Acmopetala	5885	Uzbekistan
Allium tashkenticum Khassanov et R. M. Fritsch		Acmopetala	6171	Uzbekistan
Allium tashkenticum Khassanov et R. M. Fritsch		Acmopetala	6205	Uzbekistan
Allium trautvetterianum Regel		Compactoprason	3123	Tajikistan
Allium trautvetterianum Regel		Compactoprason	6130	Tajikistan
Allium tulipifolium Ledeb.		Melanocrommyum	2966	Kazakhstan
Allium tulipifolium Ledeb.		Melanocrommyum	3699	Kazakhstan
Allium tulipifolium Ledeb.		Melanocrommyum	116G	Kazakhstan
Allium verticillatum Regel		Verticillata	2182	Tajikistan
Allium verticillatum Regel		Verticillata	6162	Tajikistan
Allium viridulum Ledeb.		Melanocrommyum	1I	Kazakhstan
Allium viridulum Ledeb.		Melanocrommyum	5I	Kazakhstan
Allium vvedenskyanum Pavlov		Acmopetala	3723	Kazakhstan
Allium winklerianum Regel		Regeloprason	6292	Tajikistan
Allium winklerianum Regel		Regeloprason	6315	Tajikistan
Allium winklerianum Regel		Regeloprason	6434	Tajikistan
Allium zergericum F.O.Khass. & R.M. Fritsch		Acmopetala	2992	Uzbekistan

Table 2. Species, accessions, subgeneric classification and the provenances of plants used for *rbc*L analyses of the genus *Allium* (Chapter 4)

species	subspecies or variety	Taxonomy (subgenus)	Accession No	origin or/and source
Allium altaicum Pall.		Cepa	1422	Russia
Allium aflatunense B. Fedtsch.		Melanocrommyum	2977	Kazakhstan, Dr. Kamenetzkaja, private collection
Allium ampeloprasum L.	var. ampeloprasum	Allium	0907	Botanical Garden Cordoba, Spain
Alliums angulosum L.		Rhizirideum	1746	Botanical Garden Budakalasz, Huhgary
Allium anisopodium Ledeb.		Rhizirideum	2349	Mongolia
Allium aroides Popov & Vved.		Melanocrommyum	6189	Uzbekistan
Allium atrosanguineum Kar. et Kir.		Cepa	OS05-34-0012-20	Kyrgyzstan
Allium backhousianum Regel		Melanocrommyum	2680	Kyrgyzstan
Allium beesianum Smith.		Reticulatobulbosa	2211	Botanical Garden Hauniensis Kopenhagen, Denmark
Allium bidentatum Fisch. ex Prokh.		Rhizirideum	1696	Mongolia
Allium brevidens Vved.		Allium	5037	Uzbekistan
Allium brevistylum Wats.		Amerallium	A125	Herbarium of Vienna W, Austria
Allium bulgaricum (Janka) Prodan		Nectaroscordum	3220	received from R. Dadd, Wokingham, Berkshire, UK
Allium caeruleum Pall.		Allium	1525	Botanical Garden of Academy of Sciences, Moscow, Russia
Allium caesium Schrenk		Allium	5048	Kazakhstan
Allium campanulatum S. Wats.		Amerallium	3061	R. H. S. Garden, Wisley, UK
Allium canadense L.		Amerallium	0913	Botanical Garden Poznan, Poland
Allium cardiostemon Fischer et Meyer		Melanocrommyum	3947	Iran
Allium carolinianum DC.		Polyprason	2570	Tajikistan
Allium cepa L.		Cepa	G2168	Georgia
Allium cernuum Roth		Amerallium	2654	USA
Allium chamaemoly L.	*chamaemoly*	Amerallium	Am-90	Herbarium of Vienna W, Austria
Allium chinense G. Don		Butomissa	0988	Kusawa Research Station, Japan
Allium clathratum Ledeb.		Reticulatobulbosa	3165	Russia

Allium condensatum Turcz.		Cepa	OS04-33-0053-10	Russia
Allium costatovaginatum Kamelin et Levichev ex Krassovskaja et Levichev		Melanocrommyum	6364	Uzbekistan
Allium cratericola Eastw.		Amerallium	Am-004	USA
Allium cristophii Trautv.		Melanocrommyum	6246	Iran
Allium crystallinum Vved.		Allium	3662	Uzbekistan
Allium cupanii Raf.	*hirtovaginatum* (Kunth) Stearn	Allium	5336	Turkey
Allium cyathophorum Bur. et Franch.	var. *farreri* (Stearn) Stearn	Cyathophora	4247	Botanicalal Garden Jena, Germany
Allium cyrilli Tenore		Melanocrommyum	5296	Turkey
Allium daghestanicum Grossh.		Polyprason	1741	Bakuriani collection, Georgia,
Allium darwasicum Regel		Melanocrommyum	6138	Tajikistan
Allium denticulatum (Traub) McNeal		Amerallium	Am-005	USA
Allium drobovii Vved.		Reticulatobulbosa	H 6022	Uzbekistan, herbarium Tashkent (TASH)
Allium eduardii Stearn		Rhizirideum	2761	Russia
Allium eremoprasum Vved.		Allium	5243	Uzbekistan
Allium fasciculatum Rendle		Amerallium	Am-069	China
Allium fasciculatum Rendle		Amerallium	Am-065	Herbarium Frankfurt (FR), Germany
Allium fedschenkoanum Regel		Cepa	H 6022	Herbarium Gatersleben (GAT)
Allium fetisowii Regel		Melanocrommyum	5685	Kyrgyzstan
Allium filidens Regel		Allium	3722	Kazakhstan
Allium fimbriatum Wats.	var. *fimbriatum*	Amerallium	Am-006	USA, California
Allium fistulosum L.		Cepa	0266	R. H. S. Garden, Wisley, UK
Allium flavum L.	var. *flavum*	Allium	2309	Bulgaria
Allium forrestii Diels		Reticulatobulbosa	Am-115	China, SE Tibet
Allium gunibicum Miscz. ex Grossh.		Polyprason	2333	Georgia
Allium gypsaceum Popov et Vved. ex Vved.		Melanocrommyum	5669	Uzbekistan
Allium haneltii Khassanov et R.M. Fritsch		Allium	5796	Uzbekistan
Allium hookeri Thw.		Amerallium	Am-117	Bhutan
Allium howellii Eastw.	var. *howellii*	Amerallium	Am-09	USA
Allium hymenorrhizum Ledeb.		Polyprason	2571	Tajikistan
Allium insubricum Boiss. et Reut.		Amerallium	Am-127	Italy
Allium jesdianum Boiss. et Buhse	*angustitepalum* (Wendelbo)	Melanocrommyum	3783	Hoog & Dix-Export, Heemstede,

	Khassanov et R.M. Fritsch			Netherlands
Allium jodanthum Vved.		Reticulatobulbosa	2869	Tajikistan
Allium kaschianum Regel		Polyprason	2418	Tajikistan
Allium kopetdagense Vved.		Allium	5802	Turkmenistan
Allium kujukense Vved.		Vvedenskya	5046	Kazakhstan
Allium kunthii G.Don		Amerallium	2158	Mexico
Allium litvinovii Drob. ex Vved.		Allium	3350	Kyrgyzstan
Allium macleanii J. G. Baker		Melanocrommyum	6150	Tajikistan
Allium macranthum Bak.		Amerallium	Am-051	Butan
Allium mairei H. Lév.		Cyathophora	5896	Nancy Botanical Garden, France
Allium massaessylum Bat. et Trab.		Amerallium	Am-123	Spain
Allium macrostemon Bge.	var. *macrostemon*	Allium	1959	North Korea
Allium monathum Maxim.		Microscordum	OS0131-20	Russia
Allium mongolicum Regel		Rhizirideum	2372	Mongilia
Allium moschatum L.		Polyprason	6063	Georgia
Allium narcissiflorum Vill.		Amerallium	OS07-33-0001-20	France
Allium neriniflorum (Herb.) Baker		Caloscordum	5392	Botanical Garden of the University of Göttingen, Germany
Allium obliquum L.		Polyprason	O-37	Russia
Allium oliganthum Kar. et Kir.		Cepa	3201	Russia
Allium oreophilum C. A. Meyer		Porphyroprason	Am-105	Kyrgyzstan
Allium oreoprasum Schrenk		Butomissa	5743	Kyrgyzstan
Allium oreoprasoides Vved.		Reticulatobulbosa	H 6027	Uzbekistan, Herbarium Tashkent (TASH)
Allium oschaninii O. Fedt.		Cepa	2177	Tajikistan
Allium ovalifolium Hand.-Mazz.		Anguinum	5092	China
Allium pamiricum Wend.		Allium	H 6057	Pakistan Herbarium Tashkent (TASH)
Allium paniculatum L.	*paniculatum*	Allium	3828	Italy
Allium paradoxum (M. Bieb.) G. Don	var. *paradoxum*	Amerallium	01-38-0007-80	Osnabrück Botanical Garden, Germany
Allium parvulum Vved.		Allium	5055	Kyrgyzstan
Allium pallasii Murr.		Allium	5797	Kyrgyzstan
Allium peninsulare Lemmon		Amerallium	Am-021	USA, California

Allium petraeum Kar. et Kir.		Polyprason	5616	Kazakhstan
Allium platyspathum Schrenk	*platyspathum*	Polyprason	5600	Tajikistan
Allium prattii C. H. Wright		Anguinum	Am-109	China
Allium polyrhizum Turcz. ex Regel		Rhizirideum	1692	Mongolia
Allium przewalskianum Regel		Reticulatobulbosa	3882	Mark Mc Donough, Pepperell, USA
Allium ramosum L.		Butomissa	1695	Mongolia
Allium roylei Stearn		Polyprason/Cepa	5152	Olomouc collection, Czech Republic
Allium rubens Schrad. ex Willd.		Rhizirideum	3193	Russia
Allium schmitzii Coutinho		Cepa	5226	received from Mark Mc Donough, Pepperell, USA
Allium schoenoprasum L.		Cepa	0192	Cluj-Napoca Botanical Garden, Romania
Allium schoenoprasoides Regel		Allium	3610	Kazakhstan
Allium scabriscapum Boiss. ex Kotschy		Reticulatobulbosa	3901	Iran
Allium semenovii Regel		Cepa	OSBU16059	Kyrgyzstan
Allium senescens L.		Rhizirideum	3397	Russia
Allium sergii Vved.		Melanocrommyum	3714	Kazakhstan
Allium spicatum (Prain) Friesen		Cyathophora	Am-114	Nepal
Allium splendens Willd. ex Schult. et Schult.f.		Reticulatobulbosa	3392	Russia
Allium stellatum Ker		Amerallium	Am-054	Herbarium of Vienna (W), Austria
Allium stipitatum Regel		Melanocrommyum	3087	Botanical Garden Univ. Strasbourg, France
Allium strictum Schrader		Reticulatobulbosa	OS02-36-0044-20	Russia
Allium suaveolens Jacq.		Polyprason	1603	Botanical Garden Univ. Padua, Italy
Allium tanguticum Regel		Allium	3779	China
Allium tenuissimum L.		Rhizirideum	3072	Botanical Garden Acad. Pyongyang, North Korea
Allium thunbergii G. Don		Cepa	1963	North Korea
Allium tricoccum Solander in Ait.		Anguinum	2582	Received from J. F. Swenson, Glencoe, USA
Allium triquetrum L.		Amerallium	OS06-18-0030-10	Spain
Allium sphaerocephalon L.	*trachypus* (Boiss. & Spruner) K.Richter	Allium	5315	Turkey

Allium trachyscordum Vved.		Reticulatobulbosa	3615	Kazakhstan
Allium tuberosum Rottl. ex Sprengel		Butomissa	1970	North Korea
Allium tuvinicum (Friesen) Friesen		Rhizirideum	1686	Mongolia
Allium umbilicatum Boiss.		Allium	1060	Gothenburg Botanical Garden, Sweden
Allium ursinum L.	*ursinum*	Amerallium	5831	Denmark
Allium vavilovii Popov et Vved.		Cepa	5238	Turkmenistan
Allium verticillatum Regel		Melanocrommyum	6162	Tajikistan
Allium victorialis L.	*victorialis*	Anguinum	1300	Russia
Allium vodopjanovae Friesen	*vodopjanovae*	Rhizirideum	3192	Russia
Allium wallichii Kunth		Amerallium	Am-066	China
Allium wallichii Kunth		Amerallium	Am-067	China
Allium weschnjakovii Regel		Cepa	OS05-34-0094-20	Kyrgyzstan