Evolution of genome size across some cultivated *Allium* species

A. Ricroch, R. Yockteng, S.C. Brown, and S. Nadot

Abstract: *Allium* L. (Alliaceae), a genus of major economic importance, exhibits a great diversity in various morphological characters and particularly in life form, with bulbs and rhizomes. *Allium* species show variation in several cytogenetic characters such as basic chromosome number, ploidy level, and genome size. The purpose of the present investigation was to study the evolution of nuclear DNA amount, GC content, and life form. A phylogenetic approach was used on a sample of 30 *Allium* species, including major vegetable crops and their wild allies, belonging to the 3 major subgenera *Allium*, *Amerallium,* and *Rhizirideum* and 14 sections. A phylogeny was constructed using internal transcribed spacer (ITS) sequences of 43 accessions representing 30 species, and the nuclear DNA amount and the GC content of 24 *Allium* species were investigated by flow cytometry. For the first time, the nuclear DNA content of *Allium cyaneum* and *Allium vavilovii* was measured, and the GC content of 16 species was measured. We addressed the following questions: (*i*) Is the variation in nuclear DNA amount and GC content linked to the evolutionary history of these edible *Allium* species and their wild relatives? (*ii*) How did life form (rhizome or bulb) evolve in edible *Allium*? Our results revealed significant interspecific variation in the nuclear DNA amount as well as in the GC content. No correlation was found between the GC content and the nuclear DNA amount. The reconstruction of nuclear DNA amount on the phylogeny showed a tendency towards a decrease in genome size within the genus. The reconstruction of life form history showed that rhizomes evolved in the subgenus *Rhizirideum* from an ancestral bulbous life form and were subsequently lost at least twice independently in this subgenus.

Key words: Allium, nuclear DNA amount, GC content, flow cytometry, internal transcribed spacer (ITS), phylogeny, life form.

Résumé : Le genre *Allium* L. (Alliaceae), d'une importance économique majeure, présente une grande diversité dans divers caractères morphologiques et particulièrement dans la forme de vie avec des bulbes et des rhizomes. Les espèces d'*Allium* montrent une variation dans plusieurs caractères cytogénétiques tels que le nombre de chromosome de base, le niveau de ploïdie ou la taille du génome. Le but de cette présente investigation était d'étudier l'évolution de la quantité en ADN nucléaire, du contenu en GC et la forme de vie en utilisant une approche phylogénétique sur un échantillon de 43 accessions représentant 30 espèces d'*Allium*, incluant les espèces cultivées légumières majeures et leurs aïeux sauvages, appartenant aux 3 sous-genres *Allium*, *Amerallium* et *Rhizirideum* et 14 sections. Une phylogénie a été construite à l'aide des séquences ITS de 43 accessions représentant 30 espèces, et la quantité en ADN nucléaire et le contenu en GC de 24 espèces d'*Allium* ont été examinés par cytométrie en flux. Pour la première fois, le contenu en ADN nucléaire de *Allium cyaneum* and *Allium vavilovii* a été mesuré et le contenu en GC de 16 espèces. Les auteurs ont posé les questions suivantes : (*i*) La variation en quantité en ADN nucléaire et le contenu en GC est-elle liée à l'histoire évolutive de ces *Allium* comestibles et leurs parents sauvages? (*ii*) Comment la forme de vie a-t-elle évolué chez les *Allium* comestibles? Nos résultats révèlent une variation interspécifique significative aussi bien de la quantité en ADN nucléaire que du contenu en GC. Aucune corrélation n'est trouvée entre le contenu en GC et la quantité en ADN nucléaire. La reconstruction de la quantité en ADN nucléaire sur la phylogénie montre une tendance vers la diminution de la taille du génome à l'intérieur du genre. La reconstruction de l'histoire de la forme de vie montre que les rhizomes ont évolué dans le sous-genre *Rhizirideum* à partir d'une forme de vie ancestrale bulbeuse et ont été par conséquent perdus au moins 2 fois indépendamment dans ce sous-genre.

Mots clés : Allium, quantité en ADN nucléaire, contenu en GC, cytométrie en flux, espace transcrit interne, phylogénie, forme de vie.

Received 13 September 2004. Accepted 8 February 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 10 June 2005.

Corresponding Editor: P.B. Moens.

A. Ricroch,[1,2] R. Yockteng, and S. Nadot. Laboratoire d'Ecologie, Systématique et Evolution, Université Paris-Sud, CNRS UMR 8079, Bâtiment 360, 91405 Orsay, France.
S.C. Brown. Laboratoire de Cytométrie, Institut des Sciences du Végétal, CNRS UPR 2355, 91198 Gif-sur-Yvette, France.

[1]Corresponding author (e-mail: agnes.ricroch@ese.u-psud.fr).
[2]Present address: Chaire de Génétique évolutive et d'amélioration des plantes, Institut National Agronomique Paris-Grignon, 16, rue Claude Bernard, 75231 Paris CEDEX 05, France.

 doi: 10.1139/G05-017

Introduction

The genus *Allium* L. (Alliaceae) exhibits a great diversity in various morphological characters, particularly in life form (bulb or rhizome) and ecological habitat. It is of major economic importance as vegetable crops, herbal crops, and ornamental plants. This genus consists mostly of perennial and bulbous plants (Stearn 1992); and it is widely distributed over Holarctic regions from the dry subtropics to the boreal zone. *Allium schoenoprasum* even occurs in the subarctic belt. The major centre of species diversity is a region stretching from the Mediterranean basin to Pakistan and central Asia. The next most important centre is located in western North America.

A multidisciplinary approach, including morphological, anatomical, and karyological investigations; studies of life cycles, distribution, and ecology; and systematic studies using biochemical and molecular markers, led to an infrageneric classification recognizing 6 subgenera (*Allium*, *Amerallium*, *Bromatorrhiza*, *Caloscordum*, *Melanocrommyum*, and *Rhizirideum*) and 43 sections (Hanelt et al. 1992; Hanelt and Fritsch 1994; Khassanov and Fritsch 1994; for review see Klaas 1998).

The phylogenetic relationships within *Allium* have been investigated by several authors using various molecular markers such as RAPDs, RFLPs, and amplified fragment length polymorphisms (AFLPs) (see for review Klaas 1998; Fritsch and Friesen 2002; Klaas and Friesen 2002); sequences from the nuclear ribosomal ITS (nrITS) region in *Allium* subg. *Melanocrommyum* (Mes et al. 1999); or sequences from the chloroplast trnD(GUC)-trnT(GGU) region in 14 Himalayan species of *Allium* (Friesen et al. 2000). Comparison of the molecular phylogenies with the classification led Fritsch and Friesen (2002) to propose a phylogenetic classification of the genus, recognizing 67 sections and 14 subgenera.

The genus *Allium* displays a high diversity in ploidy level, varying from 2*x* to 16*x* (De Sarker et al. 1997; Klaas 1998; Bennett et al. 2000). Basic chromosome numbers of $x = 7$, 8, and 9 have been reported. In a study of 25 *Allium* species, Jones and Rees (1968) found considerable differences among 2C-values measured by Feulgen densitometry. Ohri et al. (1998) confirmed this in a survey of 86 *Allium* species (representing all 6 subgenera), measured in 4C nuclei by Feulgen densitometry. The same conclusions were drawn from the study of genome size in 28 *Allium* species (Baranyi and Greilhuber 1999).

Base composition, expressed for instance as GC content, varies considerably across angiosperms, but is generally quite stable within a genus. However, variation at the generic level occurs in *Allium* (Kirk et al. 1970; Ricroch and Brown 1997). In animals, there is a tendency of larger genomes to have a higher GC frequency (Vinogradov 1998). The positive correlation between GC frequency and genome size may be explained by the greater physical and chemical stability in large genomes containing a relatively high GC frequency (Vinogradov 1998). A similar relation seems to exist in higher plants (Vinogradov 1994). Thus, Barow and Meister (2001) showed no correlation in higher plants (54 species including 3 *Allium* species) as the fluorescence of base-specific dyes is influenced by the nonrandom distribution of bases in the DNA molecule. No correlation could be found between nuclear DNA amount and GC content among the 20 *Allium* species analysed by Kirk et al. (1970). In this present study, we checked any correlation among 16 other *Allium* species. The variation in cytogenetic characters such as nuclear DNA amount and chromosome numbers, when compared with the classification of the genus *Allium*, does not reveal clear discontinuities between the taxonomic groups (Ohri et al. 1998); but the variation mirrors the great diversity observed in morphology, life form, ecology, and breeding systems (Hanelt et al. 1992). However, the biological significance of this variation is little understood.

Until now, no study reported on the evolution of characters such as the nuclear composition and life form among edible *Allium*. The purpose of the present investigation was to use a phylogenetic approach for studying the evolution of characters such as nuclear DNA amount, GC content, and life form (bulb or rhizome), and to examine possible correlations among these characters in crop species of *Allium*. Our sampling focuses on major vegetable crops of economic importance and their wild allies, such as the bulb onion (*Allium cepa* L.), chive (*A. schoenoprasum* L.), Japanese bunching onion (*Allium fistulosum* L.), leek (*Allium ampeloprasum* L. syn. *A. porrum* G. Don), and nodding onion (*Allium cernuum* Roth). The species studied here represent the 3 major subgenera *Allium*, *Amerallium*, and *Rhizirideum* and 14 sections. We also included species belonging to different sections of these subgenera to obtain a panel covering diploid and tetraploid species with a different basic number of chromosomes ($x = 7$, $x = 8$, and $x = 9$) when they are available in germplasms. The sampling was not meant to represent the whole genus *Allium*, but rather the sampling was meant to allow the studying of the evolution of characters previously mentioned in crop species, replacing them in a phylogenetic context. A phylogeny including 43 accessions representing 30 species was constructed using internal transcribed spacer (ITS) sequences. We addressed the following questions: (*i*) Is the variation in nuclear DNA amount and GC content linked to the evolutionary history of these *Allium* species? (*ii*) How did life form (rhizome or bulb) evolve in edible *Allium*? Nuclear DNA amounts (C-values) and genome size are important biodiversity characters of fundamental significance, which have many uses (Bennett et al. 2000).

Material and methods

Plant material

The total number of species examined in this study was 30. The material of 24 species was obtained from *Allium* collections maintained in Gatersleben (Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Germany), Wageningen (Centre for Genetic Resources, the Netherlands (CGN)), and Wellesbourne (Horticulture Research International (HRI), UK), and the information about the other 6 species was obtained from the literature. The sampling included representatives of the major vegetable crops and their wild relatives. Three subgenera and 14 sections of the classification (Hanelt et al. 1992) were represented. Species considered as crops include bulb onion (*A. cepa*), chive (*A. schoenoprasum*), Japanese bunching onion (*A. fistulosum,* subgenus *Rhizirideum)*,

leek (*A. porrum,* subgenus *Allium*), and nodding onion (*A. cernuum,* subgenus *Amerallium*). Table 1 lists the species with their position in the classification and the accession numbers.

Sixteen species were represented by 1 population. Six species (*Allium altaicum*, *A. ampeloprasum*, *A. fistulosum, Allium galanthum*, *Allium heldreichii*, and *Allium senescens*) were represented by several populations from different geographic locations, and each species was maintained in distinct germplasms. Each population was represented by 10 individuals grown from a set of seeds in a frost-free greenhouse at Université Paris-Sud (France).

DNA extraction and amplification

DNA was extracted from fresh plant material taken from plants grown in a greenhouse at Université Paris-Sud (Table 1), using the Plant DNeasy extraction kit (Qiagen, Valencia, California) following the manufacturer's instructions. The ITS sequences of 17 accessions representing 13 *Allium* species and 3 outgroups were taken from GenBank. The outgroups (*Tulbaghia* and *Nothoscordum*) were chosen following Friesen et al. (2000).

DNA amplifications were performed using the universal primers ITS-1 (5′-GGAAGTAGAAGTCGTAACAAG-3′) and ITS-2 (5′-TCCTCCGCTTATTGATATGC-3′). Amplifications were carried out with 2 µL DNA, 0.3 µmol dNTP/L, 240 µmol of each primer/L, 2 mMol of $MgCl_2$/L, and 1U of *Taq* DNA polymerase (Promega, Madison, Wisconsin) in 50 µL reaction volume. We used a PTC-100 thermal cycler (MJ-Researcher, Inc., Boston, Massachusetts) programmed as follows: 3 min at 94 °C; a cycle of 40 s at 94 °C, 45 s at 50 °C, and 90 s at 72 °C repeated 35 times; and a final step of 5 min at 72 °C.

Concentration and quality of the PCR products were checked on 1.5% agarose gels. PCR products (150 ng) were purified with a polyethylene glycol (PEG) precipitation (26.2% PEG 8000, 6.6 mMol $MgCl_2$/L, 0.6 mol NaOAc/L) following the method described by Rosenthal et al. (1993) and sequenced on both strands using the same primers as for PCR amplification. Sequencing products were run on an ABI capillary sequencer (GMI, Inc., Minnesota, USA) Sequences were deposited in the GenBank database.

Phylogenetic analysis

ITS sequences were aligned using Clustal X (Thompson et al. 1997) with final corrections made manually. The 5.8S coding region was included in the alignment.

Phylogenetic analyses were performed with the computer program PAUP* 4.0*b*10*a* (Swofford 2002). Indel regions were scored as additional characters using GapCoder (Young and Healy 2003). Maximum parsimony (MP) analysis was performed using the heuristic search algorithm, with TBR branch swapping, and 100 replicates of random addition taxa. Bootstrapping (1000 replicates) was completed with TBR branch swapping and 10 replicates of random addition taxa. A distance analysis was conducted with the neighbor-joining method, using GTR pairwise distances. The evolution of nuclear DNA content and life form (bulb with or without rhizome) was reconstructed using Mesquite software (Maddison and Maddison 2003). Although rhizomes in *Allium* can take diverse shapes and growth forms, to simplify we have chosen to consider the rhizome as a single character state.

Nuclear DNA amount, GC content, and chromosome number determinations

Fresh leaves were collected from plants grown in a greenhouse at Université Paris-Sud. Nuclei were isolated according to Galbraith et al. (1983). Nuclei of *A. fistulosum* (23.3 pg DNA per 2C; GC = 39.8%; Ricroch and Brown 1997) were added to the sample as an internal reference by including a smaller piece of leaf before chopping. The concentrations of ethidium bromide (50 µg/mL, Sigma) and mithramycin A (50 µg/mL, Serva) for assessing nuclear DNA amount and GC content, respectively, have been determined previously (Ricroch and Brown 1997). Flow cytometry analysis of 2000–5000 nuclei was performed on an Epics V cytometer (Beckman-Coulter, Roissy, France) fitted with Spectra-Physics lasers using the following excitation and emission configurations: for ethidium bromide (514 nm; emission >590 nm) and for mithramycin A (458 nm; emission >550–590 nm). The intensities of fluorescence emitted by the nuclei were recorded and analysed on a Coulter MDADS graphics display computer. For the 24 *Allium* species, 3 replicate plants were analysed 4 times, so that each estimate is based on 12 measurements. No statistically significant differences were found between replicates. The 2C DNA value was divided by the ploidy index to obtain the genome size (1C DNA). The DNA amount was converted into genome size expressed in megabase pairs (Mbp) of nucleotides according to Bennett et al. (2000), namely 1 pg = 980 Mbp.

Chromosome counts of the studied species were established according to the procedure previously described by Barthes and Ricroch (2001).

Results

Phylogenetic analysis

Within *Allium*, lengths of the ITS sequences ranged from 450 bp in *Allium lineare* to 488 bp in *Allium triquetrum*. *Tulbaghia fragrans, Nothoscordum bivalve*, and *Nothoscordum gracile* have ITS regions of 476, 499, and 498 bp, respectively. The resulting alignment totaled 751 positions including the *5.8S* gene (183 positions), with 446 parsimony informative sites.

The mean sequence divergence between pairs of *Allium* species ranged from 0.003 (*Allium saxatile* versus *Allium schoenoprasum*) to 0.3327 (*Allium ursinum* versus *Allium flavum*).

Heuristic searches generated 149 most-parsimonious trees of 1585 steps, with CI = 0.587, RC = 0.466, and RI = 0.794. The strict consensus of 149 trees was relatively well resolved, and several nodes were supported by high bootstrap values (Fig. 1). The neighbor-joining tree (Fig. 2) was congruent with the most-parsimonious tree for most nodes. Differences appeared relative to the positions of *A. altaicum*, *A. fistulosum*, and *A. galanthum* as well as the position of *Allium hymenorrhizum*. However, none of the nodes involving these species was supported by the bootstrap analysis in MP.

Table 1. Taxonomy, genetic composition, and life form of 32 accessions of 30 *Allium* species.

Species	Subgenus[a]	Section[a]	Collection number[b]	ITS[c]	Rhizome[e]	*x*	Ploidy level	Genome size[f] (pg)	Genome size (Mbp)	2C DNA (pg) ± SE	GC (%) ± SE
A. altaicum Pall.	Rhizirideum	Cepa	All 233	none	+[i]	8	2	11.7	11 466	23.3±0.6	39.9[g]±0.5
—	—	—	CGN 14769	AY427527	—	—	—	—	—	—	—
A. altyncolicum Friesen	Rhizirideum	Schoenoprasum	TAX 0042	AY427528	+[i]	8	4	7.0	6 860	27.9±0.6	40[g]±0.2
A. ampeloprasum L.	*Allium*	Allium	CGN 16394	AY427529	–[j]	8	2	16.7	16 366	33.4±0.7	38.5[g]±0.5
—	—	—	HRI 4549	AY427530	—	—	—	—	—	—	—
—	—	—	HRI 4550	AY427531	—	—	—	—	—	—	—
—	—	—	HRI 5971	AY427553	—	—	—	—	—	—	—
A. angulosum L.	Rhizirideum	Rhizirideum	TAX 0534	AY427532	+[k]	8	2	13.5	13 230	27.0±0.3	40±0.3
A. ***cepa*** L.	*Rhizirideum*	Cepa	CGN 96030	AJ411944[d]	–[k]	8	2	16.2	15 876	32.4±0.2	38.7±0.3
A. ***cepa var. aggregatum*** G.Don	—	—	—	AJ411906[d]	—	—	—	—	—	—	—
A. ***cernuum*** Roth	*Amerallium*	Lophioprason	TAX 0497	AY427533	+[h]	7	2	23.2	22 736	46.5±0.9	41.2[g]±0.5
A. cyaneum Regel	Rhizirideum	Reticulatobulbosa	HRI 1265	AJ411880[d]	+[h]	8	4	11.6[q]	11 368[q]	46.5±1[q]	40.5[g]±0.1
A. drummondii Regel	Amerallium	Amerallium	HRI 1270	AY427534	–[h]	7	2	18.9	18 522	37.9±0	40.2[g]±0.01
A. ***fistulosum*** L.	*Rhizirideum*	Cepa	CGN 14764	AY427535	+[k]	8	2	11.7	11 466	23.5±0.3	39.8±0.3
—	—	—	HRI 2464	AY427536	—	—	—	—	—	—	—
A. flavum L.	Allium	Codonoprasum	HRI 1275	AJ411926[d]	+[h]	8	2	15.4	15 092	30.8±0.15	40.5[g]±0.4
A. galanthum Kar. et Kir	*Rhizirideum*	Cepa	All 256	AY427537	+[h]	8	2	14.3	14 014	28.7±0.4	38.9±0.1
A. heldreichii Boiss	Allium	Allium	NVRS 012507	AY427539	–[j]	8	4	15.8	15 484	63.3±1.3	38.7[g]±0.1
—	—	—	HRI 1283	AY427538	—	—	—	—	—	—	—
A. hymenorrhizum Ldb.	*Rhizirideum*	Oreiprason	TAX 0157	AY427540[d]	+[k]	8	2	8.9	8 722	17.9±0.2	39.6[g]±0.6
A. ledebourianum Roem. et Schult.	Rhizirideum	*Schoenoprasum*	HRI 1291	AJ411925[d]	+[i]	8	2	7.8	7 644	15.6±0.3	39.9[g]±0.1
A. leucocephalum Turcz	Rhizirideum	Reticulatobulbosa	—	AJ412757[d]	+[i]	8	2	13.2[m]	12 936[m]	—	—
A. lineare L.	Rhizirideum	Reticulatobulbosa	—	AJ411951[d]	+[i]	8	2	13.1[m]	12 838[m]	—	—
A. moly L.	Amerallium	Molium	—	AF055108[d]	–[k]	7	2	25.0[m]	24 500[m]	—	—
—	—	—	—	AJ412703[d]	—	—	—	—	—	—	—
A. nutans L.	Rhizirideum	*Rhizirideum*	TAX 1499	AJ411924[d]	+[i]	8	4	10.5	10 290	41.9±0.8	40.2±0.4
—	—	—	—	AJ411909[d]	—	—	—	—	—	—	—
A. obliquum L.	Rhizirideum	Petroprason	TAX 0589	AY427542	– or +[k,i]	8	2	11.6	11 668	23.3±0.2	39.4[g]±0.3
A. ochroleucum W. et K. (syn. *A. ericetorum* Thore)	Rhizirideum	*Oreiprason*	—	AJ412755[d]	–[h]	8	2	16.0[p]	15 680[p]	—	—
A. ***paradoxum*** M. Bieb.	Amerallium	Briseis	—	AJ412741[d]	–[k]	8	2	26.7[m]	26 166[m]	—	—
A. ***porrum*** L.	*Allium*	Allium	NVRS 014549	AY427543	–[j]	8	4	12.7	12 446	50.7±0.7	39.4±0.3
A. roylei Stearn	*Rhizirideum*	*Oreiprason*	CGN 20520	AY427544	+[h]	8	2	15.3	14 994	30.7±0.3	39.4[g]±0.2
A. saxatile M. Bieb.	*Rhizirideum*	*Oreiprason*	TAX 1251	AY427545	+[k]	8	2	9.9	9 702	19.8±0.3	39.3[g]±0.1
A. ***schoenoprasum*** L.	Rhizirideum	*Schoenoprasum*	HRI 1340	AY427546	+[k]	8	2	7.7	7 546	15.4±0.2	39.9±0.2
—	—	—	HRI 1341	AY427547	—	—	—	—	—	—	—
A. senescens L.	Rhizirideum	*Rhizirideum*	CGN 15758	AY427549	+[i]	8	4	10.9	10 682	43.6±0.6	40.2±0.2
—	—	—	HRI 1256	AY427548	—	—	—	—	—	—	—
—	—	—	HRI 2498	AY427550	—	—	—	—	—	—	—

Table 1. Taxonomy, genetic composition, and life form of 32 accessions of 30 *Allium* species.

Species	Subgenus[a]	Section[a]	Collection number[b]	ITS[c]	Rhizome[e]	x	Ploidy level	Genome size[f] (pg)	Genome size (Mbp)	2C DNA (pg) ± SE	GC (%) ± SE
—	—	—	HRI 8027	AY427551	—	—	—	—	—	—	—
A. senescens L. subsp. *montanum* (F.W. Schmidt) Holub	Rhizirideum	*Rhizirideum*	TAX 0363	AY427541	+[k]	8	4	10.8	10 584	43.1±0.4	40.6[g]±0.3
A. triquetrum L.	Amerallium	Briseis	—	AJ412742[d]	–[k]	9	2	18.1[p]	17 738[p]	—	—
A. ursinum L.	Amerallium	Arctoprason	—	AJ412744[d]	–[k]	7	2	31.5[n]	30 870[n]	—	—
A. vavilovii M. Pop. et Vved.	*Rhizirideum*	*Cepa*	CGN 18744	AJ411839[d]	+[h]	8	2	16.4[q]	16 072[q]	32.8±0.4[q]	39.3[g]±0.7
A. zebdanense Boiss. et Noe	Amerallium	Molium	HRI 1371	AY427552	–[h]	9	2	19.2	18 816	38.4±0.2	40.2[g]±0.2

Note: Cultivated species (crops) are in boldface. SE, standard error.

[a]Classification after Hanelt et al. (1992).

[b]Germplasm origin: CGN, Centre for Genetic Resources (Wageningen, the Netherlands); HRI, Horticulture Research Institute (Wellesbourne, UK); TAX and All, Department of Taxonomy, Institut für Pflanzengenetik und Kulturpflanzenforschung (Gatersleben, Germany); NVRS, Novosibirsk.

[c]GenBank accession numbers. For the outgroups: *Tulbaghia fragrans* AJ250300; *Nothoscordum bivalve* AJ250301; and *Nothoscordum gracile* AJ412716.

[d]Sequences retrieved from GenBank.

[e]+, presence of rhizome; –, absence of rhizome.[h]Personal observation. Data from the literature: [i]Malyschev, L.I., and Peschkova, G.A. (*Editors*). 2001. Flora of Siberia. Vol. 4. [j]Mathew, B. (*Editor*). 1996. A review of *Allium* section *Allium*. Royal Botanical Garden, Kew, U.K. [k]Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M., and Webb, D.A. (*Editors*). 1980. Flora Europea. Vol. 5. Cambridge University Press, U.K.).

[f]Genome size (1x equivalent corrected for ploidy level). The conversion from pg to Mb was made as follows: 1 pg = 980 Mb (Bennett et al. 2000). Estimation method from Feulgen densitometry. [m](Ohri et al. 1998); [n](Labani and Elkington 1987); [p](Jones and Rees 1968). [q]New estimates resulting from this study.

[g]New estimates of content resulting from this study.

All species from the subgenus *Amerallium* form a strongly supported clade, which is a sister group to the rest of the genus. The subgenus *Allium* appears nonmonophyletic, owing to the position of *A. flavum* (section *Codonoprasum*, subgenus *Allium*), accession HRI 1275, nested within the subgenus *Rhizirideum*.

Nuclear DNA amount and GC content

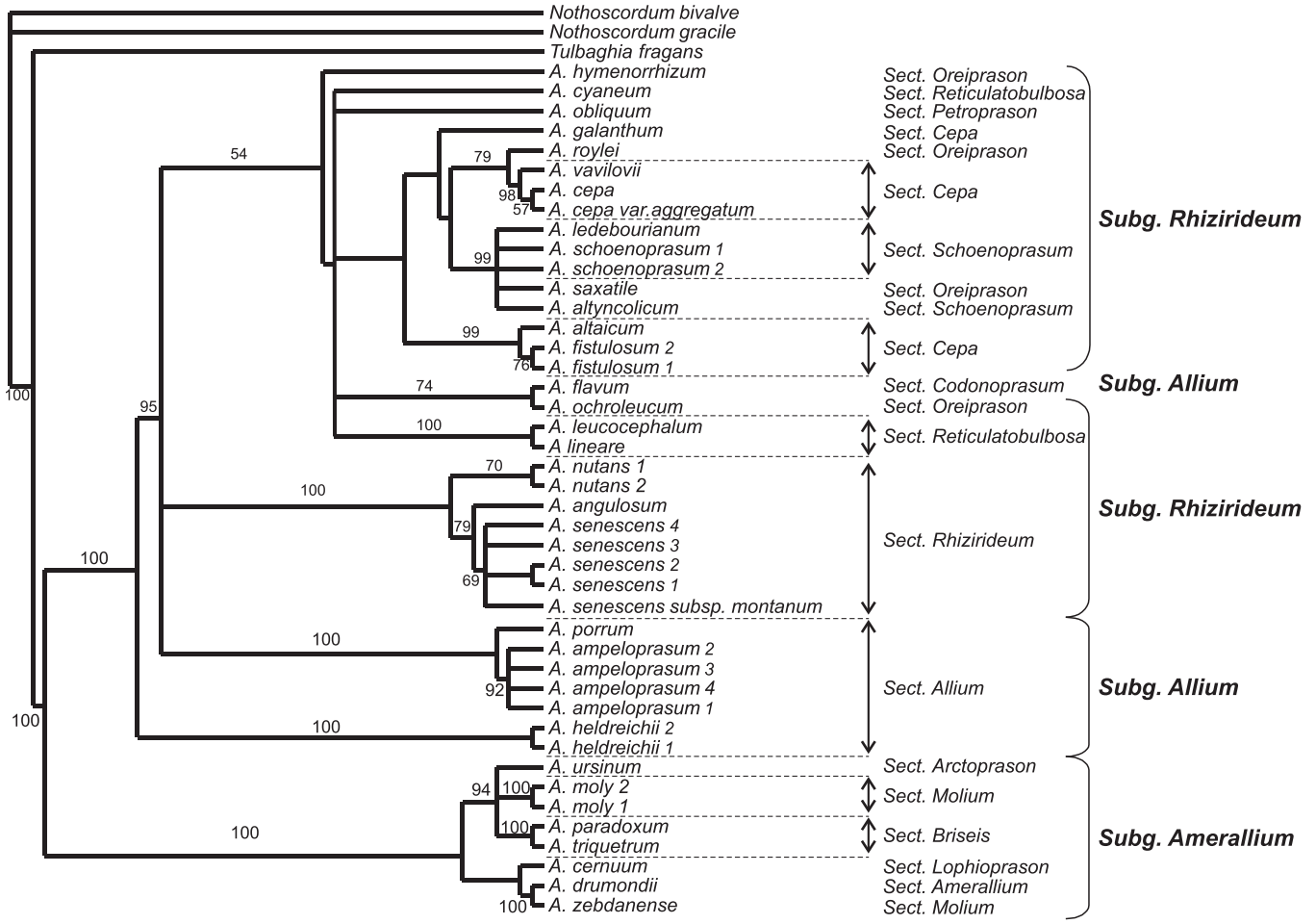
The nuclear DNA amount (2C DNA) was measured for 31 accessions representing 30 species (Table 1) and compared with the data obtained from the Angiosperms C-values database (http://www.rbgkew.org.uk) for 29 of the 30 species. For the first time, the nuclear DNA amount of *Allium cyaneum* and *Allium vavilovii* was measured. Data were taken from the Angiosperms C-values database for the following species: *Allium leucocephalum*, *A. lineare*, *Allium moly*, *Allium ochroleucum*, *Allium paradoxum*, *A. triquetrum*, and *A. ursinum*. Overall, genome size (1C DNA) values varied from 7 pg (*Allium altyncolicum*, $2n = 4x = 32$) to 31.5 pg (*A. ursinum*, $2n = 2x = 14$) showing 3.3-fold variation and a coefficient of variation of 27.4%. The analysis revealed a significant variation among species in the genome size (1C) as indicated by the ANOVA ($P < 0.001$***, Table 2). The GC content, evaluated in 24 accessions representing 23 *Allium* species (Table 1), displayed values ranging from 38.5% (*A. ampeloprasum*, $2n = 2x = 16$) to 41.2% (*A. cernuum*, $2n = 2x = 14$), with a median value of 39.9% (SD = 0.66%). The shift of 2.7 percentiles corresponded to a coefficient of variation of 1.68%, and the ANOVA revealed that the differences in GC content among species were significant ($P < 0.05$*, Table 2). Among the 23 *Allium* species examined, the GC content of 16 accessions representing 15 species was measured for the first time (Table 1). Our cytometric values for GC content of the other 8 species are in accordance with those previously obtained (Kirk et al. 1970; Ricroch and Brown 1997 for *A. cepa* and *A. fistulosum*).

Character evolution

The evolution of genome size and life form (Fig. 2) and GC content (not shown) was reconstructed with Mesquite (Maddison and Maddison 2003) using the phylogeny based on ITS sequences. The neighbor-joining tree was preferred for reconstructing character evolution, because its topology was fully resolved and in agreement with previously published molecular phylogenies (Mes et al. 1999). The tree presented in Fig. 2 does not include the outgroups, since values for the characters examined were not always available for the outgroups. The rooting was completed by comparison with the MP analysis presented in Fig. 1. A comparative analysis conducted using Mesquite (Maddison and Maddison 2003) revealed the existence of a correlation between genome size and ploidy level (Fig. 2, $P = 0.031$): tetraploidy is correlated with a small genome size.

No clear evolutionary pattern was found for the GC content of 24 accessions representing the 23 species examined in this study, and no correlation was found between nuclear DNA amount and GC content, as previously found in 20 species by Kirk et al. (1970) whose 10 species were also examined in this study.

Fig. 1. Consensus of the 149 most-parsimonious trees, resulting from a heuristic search, based on the alignment of ITS sequences from 43 accessions of *Allium*. Bootstrap percentages resulting from 1000 replicates are indicated above or below each branch. Only values above 50% are indicated. The systematic position (section and subgenus) according to Hanelt et al. (1992) follows species names.



Discussion

Phylogeny and systematics of *Allium*

Although based on a limited data set of ITS sequences, the phylogeny presented here is in overall agreement with previous phylogenetic analyses of the genus *Allium*, including some species examined in our study (Mes et al. 1999; Friesen et al. 1999, Friesen et al. 2000; Fritsch and Friesen 2002). The subgenus *Amerallium*, characterized by variation in basic chromosome number (x = 7, 8, or 9), is a sister group to the rest of the genus as found previously with various molecular data (review in Fritsch and Friesen 2002 and in Klaas and Friesen 2002). Phylogenetic relationships among species representing subgenera *Allium* and *Rhizirideum* are similar to the results obtained by Klaas and Friesen (2002) using chloroplast data. There is a problem concerning the position of *A. flavum* in our tree. *Allium flavum* is classified in the subgenus *Allium* (section *Codonoprasum*) and is described as having a true bulb. However, the plants grown from the seeds provided as accession HRI 1275 produced a visible rhizome, indicating misdetermination of the accession in the germplasm. The phylogenetic position of accession HRI 1275 observed in Fig. 1 suggests that this accession belongs to a species, which is actually part of the *Allium ericetorum* alliance, such as *A. ochroleucum*.

The good support obtained for most nodes in our analysis entitled us to use our analysis as a basis for our purpose, that is, studying the evolution of nuclear DNA amount and life form.

Evolution of nuclear DNA amount and GC content

With the exception of the nuclear DNA amount of *A. cyaneum* and *A. vavilovii*, which were measured for the first time, our cytometric values for nuclear DNA amount are in accordance with those previously obtained (Jones and Rees 1968; Labani and Elkington 1987; Ohri et al. 1998). In several instances, nuclear DNA values allowed us to confirm groupings between species suggested by classifications and molecular phylogenies. For example, this study and the phylogeny based on molecular markers (AFLP) in 20 species of the subgenus *Rhizirideum* (Van Raamsdonk et al. 2000;) show that *A. vavilovii* and *A. altaicum* are immediate presumed ancestors of *A. cepa* and *A. fistulosum*, respectively. Friesen et al. (1999) used RAPDs and noncoding chloroplast DNA to reveal only a sister group relationship, but could not

Fig. 2. Comparative analysis of genome size values versus ploidy level in cultivated *Allium* species and their wild relatives. Left tree: reconstruction of genome size evolution. Values are given in pg; underlined values correspond to tetraploid species. The legend for the character states of genome size is in the framed box on the left side of the figure. Right tree: reconstruction of life form evolution. Symbols for character states are as follows: ●, presence of a rhizome; ○, no rhizome detectable outside of the bulbs. The species names have been framed according to their systematic position (subgenus). *Allium montanum* is used as a short name for *A. senescens* subspecies *montanum*.

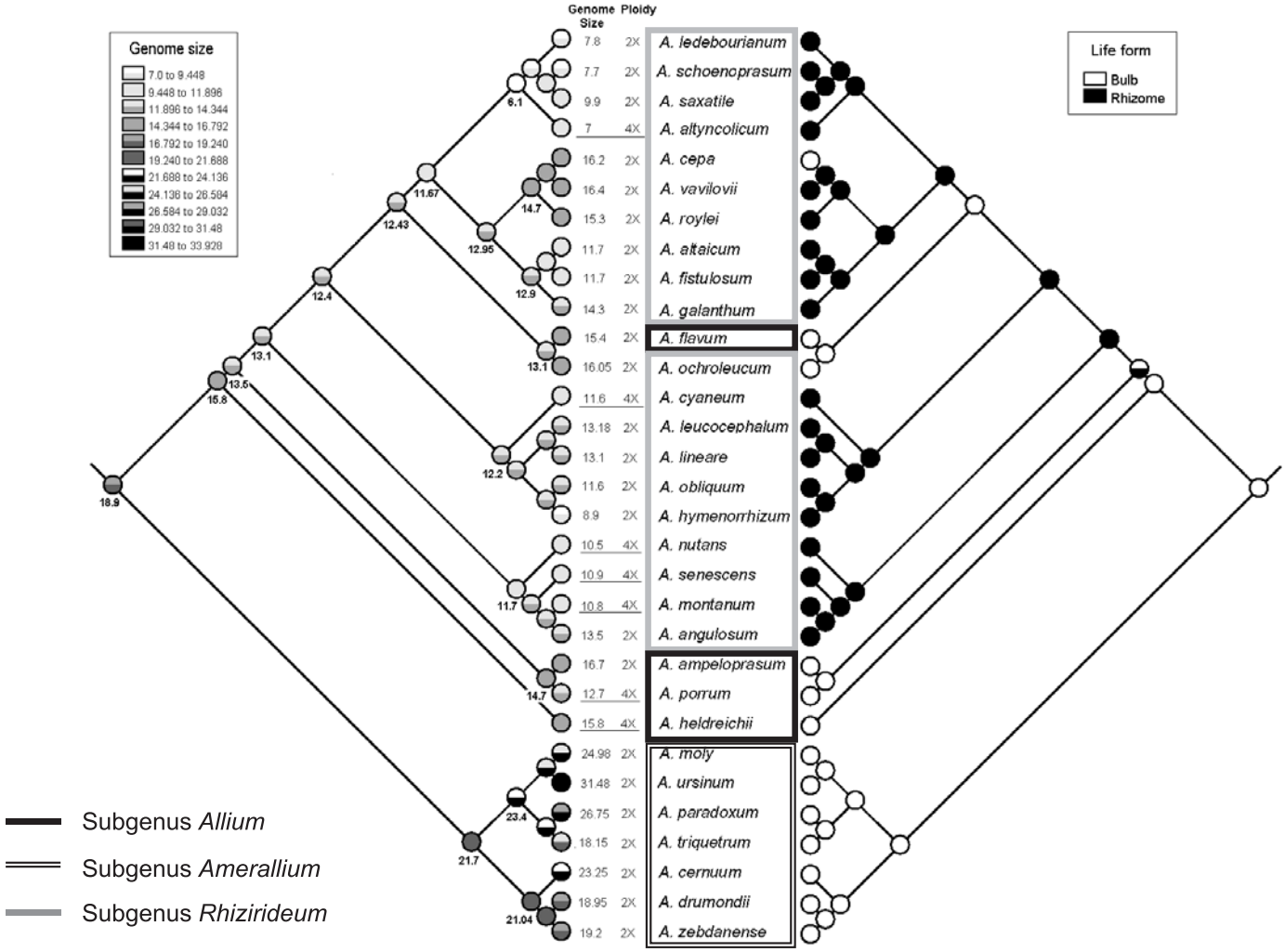


Table 2. ANOVA on 2C DNA (in pg), genome size (1C DNA in pg), and GC content (%) measured by flow cytometry in 24 *Allium* accessions.

	Degrees of freedom	MS	*F* value
2C DNA			
$P - 1$	23	571.161	1867.677***
$P(n - 1)$	48	0.306	—
Genome size (1C DNA)			
$P - 1$	23	66.482	1506.415***
$P(n - 1)$	48	0.044	—
GC			
$P - 1$	23	0.237	2.087*
$P(n - 1)$	48	0.114	—

Note: *, $P < 0.05$; ***, $P < 0.001$.

proof the progenitor. The comparison of 2C DNA values shows that there are no significant differences between these crops and their presumed wild relatives. Similarly, 2C DNA values support the close relationship between *Allium roylei* (sect. *Oreiprason*) and *A. cepa* (sect. *Cepa*) suggested by crossing experiments (Van Raamsdonk et al. 1992), molecular cytogenetics studies (Pich et al. 1996), and the phylogeny presented here. Tetraploid species examined in this study were found to have significantly smaller genome sizes than diploid species. Verma and Rees (1974) explained the diminution of nuclear DNA in polyploid genomes on the basis of higher chromatin condensation in polyploids than in diploids. The mean DNA amount per genome (genome size) tended to decrease with increasing ploidy level in monocots, eudicots, and individual families (e.g., Poaceae and Fabaceae) (Bennett et al. 2000). Transposable elements, genes, and other sequences are gradually rearranged and removed by illegitimate recombination, unequal crossing over,

and, possibly, other mechanisms in polyploid species (Devos et al. 2002).

Although previous studies had shown no clear discontinuities in base composition of nuclear DNA among taxonomic groups in *Allium* (Kirk et al. 1970), the reconstruction of 1C DNA amount (genome size) evolution onto the phylogeny (Fig. 2) brings evidence that this character has evolved towards a reduction of the genome size within *Amerallium* species examined, with reconstructed ancestral values ranging from 21.7 pg at the base of the tree to 6.1 pg in one of the most derived groups of the phylogeny (including *A. altyncolicum*, *A. ledebourianum*, *A. saxatile*, and *A. schoenoprasum*). Thus, our investigation of GC content, completed for 24 species, in relation to their genome size (1C DNA), and phylogenetic position shows no correlation between genome size and GC content without either loss or gain of GC bases. Our results support previous data on 54 species including 3 *Allium* species, which were investigated in our study (Barrow and Meister 2001), and 8 other *Allium* species (Kirk et al. 1970).

The highest genome size values are found in the monophyletic subgenus *Amerallium* at the base of the genus *Allium*, with values ranging from 18.1 pg in *A. triquetrum* to 31.5 pg in *A. ursinum* and a reconstructed ancestral value of 17.9 pg. Values then decrease across the phylogeny. The smallest values are found in the section *Schoenoprasum*, with a reconstructed ancestral value of 6.1 pg. In terms of ecology, this group is hygrophytic, inhabiting moist and marshy places (Pistrick 1992); and the wild species classified in this section are geographically restricted. The 2 wild species from the section *Schoenoprasum* included in our study are endemic to western Siberia for *A. altyncolicum* and to western, central southern Altay for *A. ledebourianum* (Friesen 2001). Thus, reduction in genome size in these species has occurred along with ecological specialization and restricted geographic distribution. The exception is the cultivated *A. schoenoprasum* (chive), which is the most widespread species in the genus, with distribution up to the arctic belt (Hanelt et al. 1992). *Allium schoenoprasum* has a small genome size in agreement with the results obtained for the other species of the section.

Evolution of life form

It has been suggested that bulbous forms have evolved as an adaptation to dry summer periods, leading to 1-shoot forms of plants with shortened vegetative periods (Cheremushkina 1992; Kruse 1992). Indeed, most bulbous species of *Allium* grow in arid areas. All species other than *A. cernuum* that represent the subgenus *Amerallium* in the phylogeny produce true bulbs. *Allium moly, A. paradoxum*, *A. triquetrum*, and *A. ursinum* are distributed in Europe and parts of Asia. *Allium drummondii* occurs mainly in dry conditions in mountains of western North America (Hanelt et al. 1992). The subgenus *Allium* occurs throughout Eurasia and is ecologically restricted to dry, open habitats with sparse vegetation (Hanelt et al. 1992). Species from this subgenus produce true bulbs. They possess the growth form and rhythm as well as the morphological and anatomical traits that provide adaptability to arid conditions (Hanelt et al. 1992; Kamenetsky 1992; Kruse 1992; Pistrick 1992). The reconstruction of life form in 30 species representing the 3 major subgenera *Allium*, *Amerallium*, and *Rhizirideum* and 14 sections, using the ITS-based phylogeny (Fig. 2) shows that the presence of rhizomes is a derived condition in the genus, as suggested by several molecular studies (Friesen et al. 1997; Mes et al. 1997; Fritsch and Friesen 2002). Bulb as storage organ is, therefore, ancestral within the genus, although former hypotheses considered rhizomes as an indication of primitive or ancestral origin, irrespective of the existing morphological diversity (Cheremushkina 1992). We interpret the absence of rhizomes in *A. ochroleucum* and in the cultivated *A. cepa* as 2 independent reversions towards the ancestral condition.

It has been hypothesized that temperate angiosperm species beginning their growth in early spring do so by sudden expansion of cells formed during the previous summer and have large genomes (2C DNA), whereas those growing later in the year do so by quick cell divisions and have small genomes (Grime and Mowforth 1982; Grime 1983). The comparison of genome size evolution and life form evolution (Fig. 2) shows that the appearance of rhizomes goes along with a reduction in 1C DNA values. However, the fact that all rhizomatous species in our sampling share a common ancestor prevents us from trying to correlate both characters. As bulbous species in *Allium* generally begin their growth in early spring, they are expected to have a large genome size.

In spite of the relatively small number of species included in the phylogenetic reconstruction, the sampling was sufficient to reveal tendencies in the evolution of the characters studied in edible *Allium* species and their wild relatives. The reconstruction of genome size evolution showed that genome size varies among the different subgroups of the genus and indicated a tendency towards a decrease in genome size within the genus. Although our data revealed significant interspecific variation in the GC content, no evolutionary pattern was found for this character. The reconstruction of life form evolution confirmed that rhizomatous species are derived within the genus. The absence of rhizomes in *A. ochroleucum* and in the cultivated *A. cepa* are 2 independent reversions towards the ancestral bulbous condition. This study shows that variation in genome size has evolutionary implications.

Acknowledgements

We thank M.T. Crosnier, D. de Nay, J.M. Bureau, and O. Catrice for their technical assistance in cytometry (Institut des Sciences du Végétal, Gif-sur-Yvette, France), G. Nicolo for assistance in DNA analysis (Université Paris-Sud, France), Dr. I. Boukema (Centre for Genetic Resources, Wageningen, the Netherlands) and Dr. J. Keller (Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatesleben, Germany) for providing plant material, and Prof. P.H. Gouyon for scientific support (Université Paris-Sud, France). Experiments presented comply with the current laws of France.

References

Baranyi, M., and Greilhuber, J. 1999. Genome size in *Allium*: in quest of reproducible data. Ann. Bot. **83**: 687–695.

Barow, M., and Meister, A. 2001. Lack of correlation between AT frequency and genome size in higher plants and the effect of nonrandomness of base sequences on dye binding. Cytometry, **47**: 1–7.

Barthes, L., and Ricroch, A. 2001. Interspecific chromosomal rearrangements in monosomic addition lines of *Allium*. Genome, **44**: 929–935.

Bennett, M.D., and Smith, J.B. 1976. Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond. Ser. B. Biol. Sci. **274**: 227–274.

Bennett, M.D., Bhandol, P., and Leitch, I.J. 2000. Nuclear DNA amounts in angiosperms and their modern uses — 807 new estimates. Ann. Bot. **86**: 859–909.

Cheremushkina, V.A. 1992. Evolution of life forms of species in subgenus *Rhizirideum* (Koch) Wendelbo genus *Allium* L. *In* The genus *Allium* — taxonomic problems and genetic resources. Proceedings of an international symposium held at Gatersleben, Germany, 11–13 June 1991. *Edited by* P. Hanelt, K. Hammer, and H. Knupffer. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany. pp. 27–34.

De Sarker, D., Johnson, M.A.T., Reynolds, A., and Brandham, P.E. 1997. Cytology of the highly polyploid disjunct species, *Allium dregeanum* (Alliaceae), and of some Eurasian relatives. Bot. J. Linn. Soc. **124**: 361–373.

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

Friesen, N. 2001. The genera *Allium* L. and *Caloscordum* Herbert. *In* Flora of Siberia. Vol. 4. Araceae-Orchidaceae. *Edited by* L. Malyshev and G. Peshkova. Science Publishers, Inc., Enfield, U.S.A. pp. 43–91, 174–192.

Friesen, N., Borisjuk, N., Mes, T.H.M., Klaas, M., and Hanelt, P. 1997. Allotetraploid origin of *Allium altyncolicum* as investigated by karyological and molecular markers. Plant Syst. Evol. **206**: 317–335.

Friesen, N., Pollner, S., Bachmann, K., and Blattner, F.R. 1999. RAPDs and non-coding chloroplast DNA reveal a single origin of the cultivated *Allium fistulosum* from *A. altaicum* (Alliaceae). Am. J. Bot. **86**: 554–562.

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

Fritsch, R.M., and Friesen, N. 2002. Evolution, domestication and taxonomy. *In Allium* crop science: recent advances. *Edited by* H.D. Rabinovitch and L. Currah. CAB international, Wallingford, U.K. pp. 5–27.

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

Grime, J.P. 1983. Prediction of weed and crop response to climate based upon measurements of DNA content. Asp. Appl. Biol. **4**: 87–98.

Grime, J.P., and Mowforth, M.A. 1982. Variation in genome size: an ecological interpretation. Nature (London), **299**: 151–153.

Hanelt, P., and Fritsch, R.M. 1994. Notes on infrageneric taxa in *Allium* L. Kew Bull. **49**: 559–564.

Hanelt, P., Schultze-Motel, J., Fritsch, R.M., Kruse, J., Maass, H.I., Ohle, H., and Pistrick, K. 1992. Infrageneric grouping of *Allium*. The Gatersleben approach. *In* The genus *Allium* — taxonomic problems and genetic resources. Proceedings of an international symposium held at Gatersleben, Germany, 11–13 June 1991. *Edited by* P. Hanelt, K. Hammer, and H. Knupffer. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany. pp. 107–123.

Jones, R.N., and Rees, H. 1968. Nuclear DNA variation in *Allium*. Heredity, **23**: 591–605.

Kamenetsky, I. 1992. Morphological types and root systems as indicators of evolutionary pathways in the genus *Allium*. *In* The genus *Allium* – taxonomic problems and genetic resources. Proceedings of an international symposium held at Gatersleben, Germany, 11–13 June 1991. *Edited by* P. Hanelt, K. Hammer, and H. Knupffer. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany. pp. 129–135.

Khassanov, F.O., and Fritsch, R.M. 1994. New taxa in *Allium* L. subg. *Melanocrommyum* (Webb and Berth) Rouy from central Asia. Linz. Biol. Beitr. **26**: 965–990.

Kirk, J.T.O., Rees, H., and Evans, G. 1970. Base composition of nuclear DNA within the genus *Allium*. Heredity, **25**: 507–512.

Klaas, M. 1998. Applications and impact of molecular markers on evolutionary and diversity studies in the genus *Allium*. Plant Breed. **117**: 297–308.

Klaas, M., and Friesen, N. 2002. Molecular markers in *Allium*. *In Allium* crop science: recent advances. *Edited by* H.D. Rabinovitch and L. Currah. CAB international, Wallingford, U.K. pp. 159–185.

Kruse, J. 1992. Growth form characters and their variation in *Allium* L. *In* The genus *Allium* – taxonomic problems and genetic resources. Proceedings of an international symposium held at Gatersleben, Germany, 11–13 June 1991. *Edited by* P. Hanelt, K. Hammer, and H. Knupffer. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany. pp. 173–179.

Labani, R.M., and Elkington, T.T. 1987. Nuclear DNA variation in the genus *Allium* L. (Liliaceae). Heredity, **59**: 119–128.

Maddison, W.P., and Maddison, D.R. 2003. Mesquite: a modular system for evolutionary analysis. Version 0.996 [computer program]. Available from http://mesquiteproject.org

Mathew, B. 1996. A review of *Allium* section *Allium*. Royal Botanical Garden, Kew, U.K.

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

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

Pich, U., Fritsch, R., and Schubert, I. 1996. Closely related *Allium* species (Alliaceae) share a very similar satellite sequence. Plant Syst. Evol. **202**: 255–264.

Pistrick, K. 1992. Phenological variability in the genus *Allium* L. *In* The genus *Allium* – taxonomic problems and genetic resources. Proceedings of an international symposium held at Gatersleben, Germany, 11–13 June 1991. *Edited by* P. Hanelt, K. Hammer, and H. Knupffer. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany. pp. 243–249.

Ricroch, A., and Brown, S.C. 1997. DNA base composition of *Allium* genomes with different chromosome numbers. Gene, **205**: 255–260.

Rosenthal, A., Coutelle, O., and Craxton, M. 1993. Large scale production of DNA sequencing templates by microtitre format PCR. Nucleic Acids Res. **21**: 173–174.

Stearn, W.T. 1992. How many species of *Allium* are known? Kew Mag. **9**: 180–181.

Swofford, D.L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0*b*10*a*. [computer program]. Sinauer Associates, Sunderland, Mass.

Thompson, J.D., Gibson, T.J., Plewnoak, F., Jeanmougin, F., and Higgins, D.J. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**: 4876–4882.

Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M., and Webb, D.A. (*Editors*). 1980. Flora Europea. Vol. 5. Cambridge University Press, Cambridge, U.K.

Van Raamsdonk, L., Wietsma, W., and De Vries, N.G. 1992. Crossing experiments in *Allium* L. section *Cepa*. Bot. J. Linn. Soc. **109**: 293–303.

Van Raamsdonk, L., Vrielink-van Ginkel, G.R., and Kik, C. 2000. Phylogeny reconstruction and hybrid analysis in *Allium* subgenus *Rhizirideum*. Theor. Appl. Genet. **100**: 1000–1009.

Verma, S.C., and Rees, H. 1974. Nuclear DNA and the evolution of allotetraploid *Brassicae*. Heredity, **33**: 61–68.

Vinogradov, A.E. 1994. Measurement by flow cytometry of genomic AT/GC ratio and genome size. Cytometry, **16**: 34–40.

Vinogradov, A.E. 1998. Genome size and GC-percent in invertebrates as determined by flow cytometry: the triangular relationship. Cytometry, **31**: 100–109.

Young, N.D., and Healy, J. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics, **4**: 6.