Pl. Syst. Evol. 204: 109–123 (1997)

Plant Systematics and Evolution

Phylogenetic analysis of *Iridaceae* with parsimony and distance methods using the plastid gene *rps*4

Tatiana T. Souza-Chies, Gabriel Bittar, Sophie Nadot, Leigh Carter, Evelyne Besin, and Bernard Lejeune

Received January 4, 1996; in revised version March 27, 1996

Key words: *Iridaceae*. – *rps*4, phylogenetic tree, molecular systematic, chloroplast evolution.

Abstract: A molecular phylogeny of the family *Iridaceae* based on the plastid gene *rps*4 was obtained using both parsimony and distance methods. Thirty-four species were examined together with eight outgroup species. Results show that the *Iridaceae* are monophyletic, and that *Isophysis* is likely to be the earliest emerging genus. Subfamily *Ixioideae* plus the genera *Aristea* and *Nivenia* form a strongly supported clade. Within subfam. *Iridoideae*, the tribe *Irideae* includes the genus *Bobartia* (of disputed position), and the tribe *Mariceae* includes *Cypella*. The division of *Iridoideae* into tribes is consistent with their geographical distribution.

The chloroplast genome has been extensively used in plant systematics for phylogenetic reconstructions. It is well suited for this because of its relatively small size, its conservative mode of evolution (i.e. its slow rate of nucleotide substitution) and its relative abundance as a component of plant total DNA. It has therefore provided much basic information to support comparative evolutionary research (Lamppa & Bendich 1979, Palmer & Zamir 1982, Palmer & Stein 1986, Zurawski & Clegg 1987, Soltis & al. 1989, Lavin & al. 1990, Wolfe & al. 1991, Clegg & Zurawski 1992, Clegg & al. 1994).

The plastid gene most commonly used for phylogenetic analyses in plants is *rbc*L, which encodes the large subunit of ribulose biphosphate carboxylase/oxygenase. More than 500 complete *rbc*L sequences are now available for land plants (Chase & al. 1993). Other chloroplast genes have also been used, such as *mat*K (Johnson & Soltis 1994, Steele & Vilgalys 1994, Johnson & Soltis 1995), *atp*B (Hoot & al. 1995), *ndh*F (Olmstead & Reeves 1995) and *rps* 4 (Nadot & al. 1994, 1995). Some non-coding regions of chloroplast DNA, such as the *atp*B-*rbc*L intergenic region (Manen & al. 1994), the *trn*L intron and the *trn*L-*trn*F spacer (Taberlet & al. 1991, Gielly & Taberlet 1994, Ham & al. 1994, Mes & Hart 1994) have been successfully used for evolutionary studies of closely related taxa. For the present work on *Iridaceae*, we chose to use the *rps* 4 gene (encoding protein 4 of the small chloroplastic ribosomal subunit) for its relatively small size (generally

600 bp), its good sequence variation and its previously successful use in a phylogenetic reconstruction in the *Poaceae* (Nadot & al. 1994).

The *Iridaceae* are a medium-sized family of plants including 77 genera and 1750 extant species (Goldblatt, pers. comm.), mostly found in the Southern Hemisphere. Africa is the centre of diversity of the family, and the majority of species are concentrated in the temperate and Mediterranean regions in the Southern part of the continent (Goldblatt & al. 1995). The tropical and subtropical America is also a centre of *Iridaceae*. Some genera may have a markedly disjunct distribution in Oceania. Many genera of *Iridaceae* are economically important because of their ornamental value. Dahlgren & al. (1985) classified *Iridaceae* into 5 subfamilies (*Isophysoideae, Aristeoideae, Sisyrinchioideae, Iridoideae,* and *Ixioideae*). The recent classification of *Iridaceae* by Goldblatt (1990) divides the family into only four different subfamilies (*Isophysidoideae, Nivenioideae, Iridoideae,* and *Ixioideae*). These classifications differ mainly in the position of *Patersonia* (subfam. *Nivenioideae* according to Goldblatt) and subfam. *Sisyrinchioideae* sensu Dahlgren & al. (1985), and in the inclusion or not of the genus *Geosiris* in the family *Iridaceae*.

Some aspects of the phylogeny of *Iridaceae* are still disputed, such as the position of the genera *Isophysis* (Dahlgren & al. 1985, Goldblatt 1990, Rudall 1994, Chase & al. 1995), *Geosiris* (Cronquist 1981, Dahlgren & al. 1985, Goldblatt 1990) and *Bobartia* (Goldblatt & Rudall 1992), the delimitation of the genera within the *Iridoideae* tribes *Mariceae, Tigridieae*, and *Irideae* (Goldblatt 1991) and the position of the subfam. *Ixioideae* within the *Iridaceae* (Goldblatt 1990, Rudall 1994). Our goal was to contribute to the solution of some of these problems using the molecular information provided by the *rps* 4 sequences. The results are presented in this paper. In order to add information to the reconstructions obtained by a parsimony program (PAUP), we also used the new distance method Anataxis (Bittar 1995), that we have already applied to a phylogeny of monocots (Nadot & al. 1995). In the discussion, the phylogenetic relationships in *Iridaceae* provided by this molecular approach are compared to other phylogenetic and biogeographic data on this family.

Material and methods

Plant material. Fresh leaves were collected from 34 species of *Iridaceae* listed in Table 1. In order to select the best set of outgroup species to reconstruct phylogenies, comparative tests were carried out using plants belonging to the *Liliales* sensu Dahlgren (in which *Iridaceae* are included) and to the *Asparagales,* based on Rudall's (1994) suggestion that *Iridaceae* are closer to *Asparagales* than to *Liliales*. Representatives from all *Liliales* families except *Iridaceae* were included in this study. *Asparagales* encompass 33 families; samples from 12 of them were assayed. Comparative trials showed that much better resolution and robustness of trees was obtained with *Asparagales* than with *Liliales,* thus confirming suggestion by Rudall (1994). Using this information, eight *Asparagales* species, listed in Table 1, were finally retained to constitute the outgroup species.

DNA isolation, amplification and sequencing. Total DNA was extracted by the CTAB method modified by Doyle & Doyle (1987). A fragment of approximately 800 bp including *rps* 4, a non-coding region and the *trn*S gene was amplified by PCR (Fig. 1). PCR primers (see below) were selected at the 5′ end of the *rps* 4 and the at 5′ end of *trn*S conserved regions. The selection of the primer *trn*S was based on the comparison of the

Table 1. Sequence accession numbers (EMBL Databank), names, geographic distribution and vouchers of the species analysed. The samples came from the "Muséum National d'Histoire Naturelle" – MNHN (Paris, France; *EB* Ecole Botanique, *S* Serre), Botanical Garden of Porto Alegre (Brazil), Botanical Garden of Geneva (Switzerland), Royal Botanical Garden of Kew (England), Missouri Botanical Garden (USA) and Departement of Botany of Campinas (Brazil)

EMBL no.	Species (geographic distribution)	Origin/voucher specimens
Z68231	*Alophia veracruzana* GOLDBLATT (America)	Missouri (GOLDBLATT & HOWARD) 9070 MO
Z68232	*Aristea platycaulis* BAKER (Africa, Madagascar)	MNHN (S 69-1198)
Z68234	*Babiana stricta* KER-GAWL (Africa, Socotra)	MNHN (S 94-578)
Z68235	*Belamcanda chinensis* ADANS. (Asia)	MNHN (JA 84-1389)
Z68268	*Bobartia gladiata* KER-GAWL (Africa)	Kew (G 9490 MO 16)
Z68236	*Crocosmia* sp. PLANCH. (Africa)	Porto Alegre (034)
Z68237	*Crocus nudiflorus* HOHEN. (Europe, Asia, Africa)	MNHN (EB 66-2483bis)
Z68254	*Cypella* sp. HERB. (America)	Porto Alegre (027)
Z68238	*Dietes grandiflora* BROWN (Africa)	MNHN (S 70-531)
Z68239	*Dietes robinsoniana* KLATT (Africa)	MNHN (S 88-82)
Z68240	*Freesia* sp. KLATT (Africa)	MNHN (EB)
Z68241	*Gladiolus communis* TOURN. (Africa, Madagascar, Eurasia)	MNHN (EB)
Z68255	*Gladiolus murielae* HOCHST. (Africa, Madagascar, Eurasia)	MNHN (EB)
Z68256	*Gladiolus papilio* HOOK. (Africa, Madagascar, Eurasia)	Missouri (GOLDBLATT & MANNIG) 9841 MO
Z68257	*Iris ensata* THUNB. (Europe, Asia, Africa, North America)	Missouri (SOLOMON 19400 MO)
Z68242	*Iris lutescens* GUSS. (Europe, Asia, Africa, North America)	MNHN (EB)
EmZ29254	*Iris pallida* LAM. (Europe, Asia, Africa, North America)	MNHN (EB)
Z68243	*Isophysis tasmanica* MOORE (Tasmania)	Kew (J. BRUHL 2)
Z68244	*Lapeirousia neglecta* POURR. (Africa)	Kew (GOLDBLATT & MANNIG) 9489 MO
Z68245	*Libertia formosa* GRAH. (Australasia, America)	Genève (Chile)
Z68246	*Micranthus juncus* ECKL. (Africa)	Kew (CHASE I-156)
Z68258	*Moraea spathulata* KLATT (Africa)	MNHN (EB-90-1494)
Z68247	*Neomarica* sp. SPRAGUE (America)	Porto Alegre (010)
Z68266	*Nivenia corymbosa* BAKER (Africa)	Kew (GOLDBLATT s.n. MO-MWC 418)
Z68248	*Patersonia fragilis* ASCHERS & GRAEBN. (Oceania)	MNHN (S 92-20)
Z68259	*Patersonia* sp. BENTH. & HOOK. (Oceania)	Kew (UNSW 21494-MWC 418)
Z68260	*Pillansia templemannii* BOLUS (Africa)	Kew (BEAN s.n. MO)
Z68261	*Romulea revelieri* JORD. & FOURR. (Africa)	MNHN (MORET 93-88)
Z68262	*Sisyrinchium* sp. BENTH. & HOOK. (America)	Campinas (Herbarium)

Table 1 (continued)

EMBL no.	Species (geographic distribution)	Origin/voucher specimens
Z68263	*Sisyrinchium striatum* L. (America)	Genève (Chile)
Z68249	*Sparaxis* sp. Ker-Gawl (Africa)	MNHN (EB)
Z68250	*Tigridia* sp. Juss. (America)	MNHN (EB)
Z68264	*Trimezia stayermarkii* R. Foster (America)	Missouri (Goldblatt 1345 MO 54)
Z68265	*Watsonia anguta* Ker-Gawl (Africa)	Kew (Goldblatt 6904 MO 4)
X84109	*Agave bracteosa* Wats.	MNHN (S 54-699)
X84115	*Asparagus scaber* Brign	MNHN (EB)
Z68267	*Furcraea gigantea* Vent.	MNHN (S 64-321)
Z68253	*Haemanthus magnificus* L.	MNHN (S 69-18)
Z68251	*Hymenocallis littoralis* Salisb.	MNHN (S 92-303)
X84133	*Narcissus odorus* Tourn.	MNHN (EB)
Z68252	*Nerine bowdenii* Watson.	MNHN (EB)
X84148	*Yucca filamentosa* Wood	MNHN (S)

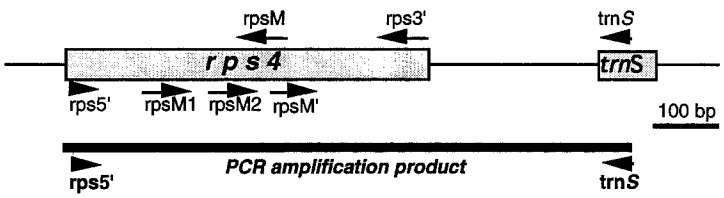

Fig. 1. Positions and directions of the primers used to amplify and to sequence the *rps* 4 gene

Genbank sequences from spinach (accession number AC = M16878), rice (AC = X15901), tobacco (AC = Z00044/S54304) and *Marchantia* (AC = X04465/Y00686). The rps5′ primer was selected by a comparison of *rps* 4 sequences of spinach, tobacco, rice and maize (AC = X01608). PCR was carried out in a volume of 100 µl using a standard technique with Taq polymerase (Appligene, Strasbourg, France). The annealing temperature was 60 °C. PCR products were directly sequenced with the Sequenase PCR Product Sequencing Kit (USB-Amersham) including a treatment of the PCR product with Exonuclease I and Alkaline Phosphatase. The positions of the PCR and sequencing primers of the *rps* 4 gene are shown in Fig. 1. Sequences of primers used in this work are given in Table 2.

Alignment and analysis of sequence data. Nucleotide sequences were aligned using the CLUSTAL W 1.5 program (Thompson & al. 1994). These alignments were used for construction of maximum parsimony (MP) trees using the heuristic and branch-and-bound research modes (TBR option) of program PAUP 3.1.1 (Swofford 1993). The "branch and bound" and heuristic search modes of PAUP were used in several phylogenetic reconstruction attempts. Whether the indels should be taken into account or not was considered: both options were used and the gaps corresponding to the insertions (see below) were alternatively treated as missing data or additional character state ("fifth base"). Bootstrap (1000 replications) using the search option heuristics of PAUP 3.1.1 were done. An estimation of the decay of parsimony analysis for each branch (Bremer 1988) was performed using the trees longer by 1, 2 and 3 steps than the most parsimonious trees. This decay analysis also

Table 2. Names and sequences of the primers for PCR and sequencing

Name	Sequence
trnS	5′(TACCGAGGGTTCGAATC)3′
rps5′	5′(ATGTCCCGTTATCGAGGACCT)3′
rps3′	5′(ATATTCTACAACTAACAACTC)3′
rpsM	5′(CCATTAACTAAAATATGTGT)3′
rpsM1	5′(CAATATCGTATTCGTCTAGAA)3′
rpsM2	5′(TTTTACTACAACTACTTGAGA)3′
rpsM′	5′(TAGACATATTTTAGTTAATGG)3′

used the program PAUP 3.1.1. Distance analyses were performed using the new program ANATAXIS (Bittar & Carter 1994, Bittar 1995, Nadot & al. 1995). This program was developed as an alternative to the quick but biased classical phenetic methods. It is a tree-compatibility distance method that groups taxa while taking into account both the possibility of homoplasy and of different rates of evolution within different branches. We used this obvious advantage in the phylogenetic reconstruction of monocots (Nadot & al. 1995) and in the present work on *Iridaceae*. The objective of ANATAXIS is to allow a quick analysis of a large set of data by using all possible information without having to analyse the possible evolutionary story of each site (as cladistic parsimony methods do), while nevertheless avoiding the phenetic pitfalls of widely differing evolutionary rates in the different lineages. It should be noted that the ANATAXIS trees indicate polychotomies when information is judged to be insufficient, in order to avoid the unnecessary display of uncertain or unsupported phylogenetic relationships. The program MacClade 3.0 (Maddison & Maddison 1992) was used for performing analysis of character evolution.

Results

With the aid of the deduced amino acid sequences which they encode, the *rps* 4 nucleotide sequences were aligned for all plants listed in Table 1, representing 34 *Iridaceae* and 32 distinct sequences, since those of *Iris pallida* and *I. lutescens* are identical, as are those of *Gladiolus communis* and *G. papilio*. Similarly, 8 outgroup species (represented by 7 distinct sequences, those of *Agave bracteosa* and *Yucca filamentosa* being identical) were also used. Including insertions (see below), the nucleotide sequence matrix encompasses 552 nucleotide sites (corresponding to 184 amino acids) out of the 648 of the complete reading frame.

The nucleotide sequence alignment displays 129 variables sites, 80 of which are cladistically informative; 14% of these changes occur at the first position of

Table 3. Calculated values of the consistency index (c.i.), retention index (r.i.) and rescaled consistency index (r.c.i. = c.i. × r.i.) for each codon position for the three illustrated in Fig. 3

Codon position	c.i.	r.i.	r.c.i.
1st	1.00	1.00	1.00
2nd	0.69	0.78	0.57
3rd	0.70	0.92	0.64

```
Outgroups
Furcraea gigantea                ATATTGCTTCCTCCCCTCAC------GAGGAATTGCCAAACCA
Agave, Yucca                     ATATTGCTTCCTCCCCTCAC------GAGGAATTGCCAAACCA
Hymenocallis littoralis          CTATTGCTTCAGCCCCTCGC------GAGGAATTGCCAAACCA
Nerine bowdenii                  TTATTGCTTCATCCCCTCAC------GAGGAATTGCCAAACCA
Haemanthus magnificus            TTATTGCTTCAGCCCCTCGC------GAGGAATTGCCAAACCA
Narcissus odorus                 CGATTGCTTCAGCCCCTCGC------GAGGAATTGCCAAACCA
Asparagus scaber                 CTATCGCTTCGTCCCCTCAC------GAGGAATTGCCAAACCA

Iridoideae
Trimezia stayermarkii            ATATTGCTTCATCCTCCCAC------GACCAATTACCAAACCA
Alophia veracruzana              ATATTGCTTCTACCCCCCAC------GCGGAATTACCAAACCA
Belamcanda chinensis             ATATTGCTTTGTCCCCCCAC------GAGGAATTGCCAAACCA
Bobartia gladiatus               ATACTGCTTCGTCCCCCCAC------GAGGAGTTGCCAAACCA
Cypella sp.                      ATATTGCTTCATCCTCCCAC------GAGGAATTACCAAACCA
Dietes grandiflora               ATATTGCTTCGTCCCCCCAC------GAGGAGTTGCCAAACCA
Dietes robinsoniana              ATATTGCTTCGTCCCCCCAC------GAGGAGTTGCCAAACCA
Iris ensata                      ATTTGGCTTGGTCCCCCCAC------GAGGAATTGCCAAACCA
Iris lutescens, I. pallida       ATATTGCTTCGTCCCCCCAC------GAGGAATTACCAAACCA
Moraea spathulata                ATATTGCTTCGTCCCCCCAA------GAGGAGTTGCCAAACCA
Neomarica sp.                    ATATTGCTTCATCCCCCCAC------GAGGAATTACCAAACCA
Sisyrinchium striatum            CTATTGCCTCGTCCCCCCAC------GAGGAATTACCAAATCA
Sisyrinchium sp.                 CTATTGCCTCGTCCCCCCAC------GAGGAATTACCAAATCA
Tigridia sp.                     ATATTGCTTCATCCCCCCAC------GAGGAATTACCAAATCA
Libertia formosa                 CTATTGCTTCGTCCCCCCAC------GAGGAATTACCAAACCA

Isophysidoideae
Isophysis tasmanica              ATATTGCTTCATCCCCCCAC------GAGGAATTGCCAAACCA

Ixioideae
Watsonia angusta                 TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Sparaxis sp.                     TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Pillansia templemannii           TTATTGCTTCATACCCCCAC---CAGGAGGAATTGCCAAACCA
Romulea revelieri                TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Lapeirousia neglecta             TTATTGCTTCATCCCCCCAA---CAGGAGGAATTGCCAAACCA
Gladiolus communis, G. papilio   TTATTGCCTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Gladiolus murielae               TTATTGCCTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Babiana stricta                  TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Crocosmia sp.                    TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Crocus nudiflorus                TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA
Freesia sp.                      TTATTGCTTCATCCCCCCGC---CAGGAGGAATTGCCAAACCA
Micranthus junceus               TTATTGCTTCATCCCCCCAC---CAGGAGGAATTGCCAAACCA

Nivenioideae
Aristea platycaulis              ATATTGCTTCATCCCCCCAC---CACGAGGAATTGCCAAACCA
Nivenia corymbosa                CTATTGCTTCCTCCCCCCAC---CAGGAGGAATTGCCAAACCA
Patersonia sp.                   ATATTGCTTCATCTACCCAACCCCAGGAGGAACTGCCAAACCA
Patersonia fragilis              ATATTGCTTCATCTACCCAGCCCCAGGAGGAACTGCCAAACCA
```

Fig. 2. Alignment of positions 477 to 521 of the *rps* 4 nucleotide sequences of the *Iridaceae* species and of the outgroups showing the insertions found in the genera *Patersonia, Aristea* and *Nivenia* and in the *Ixioideae*

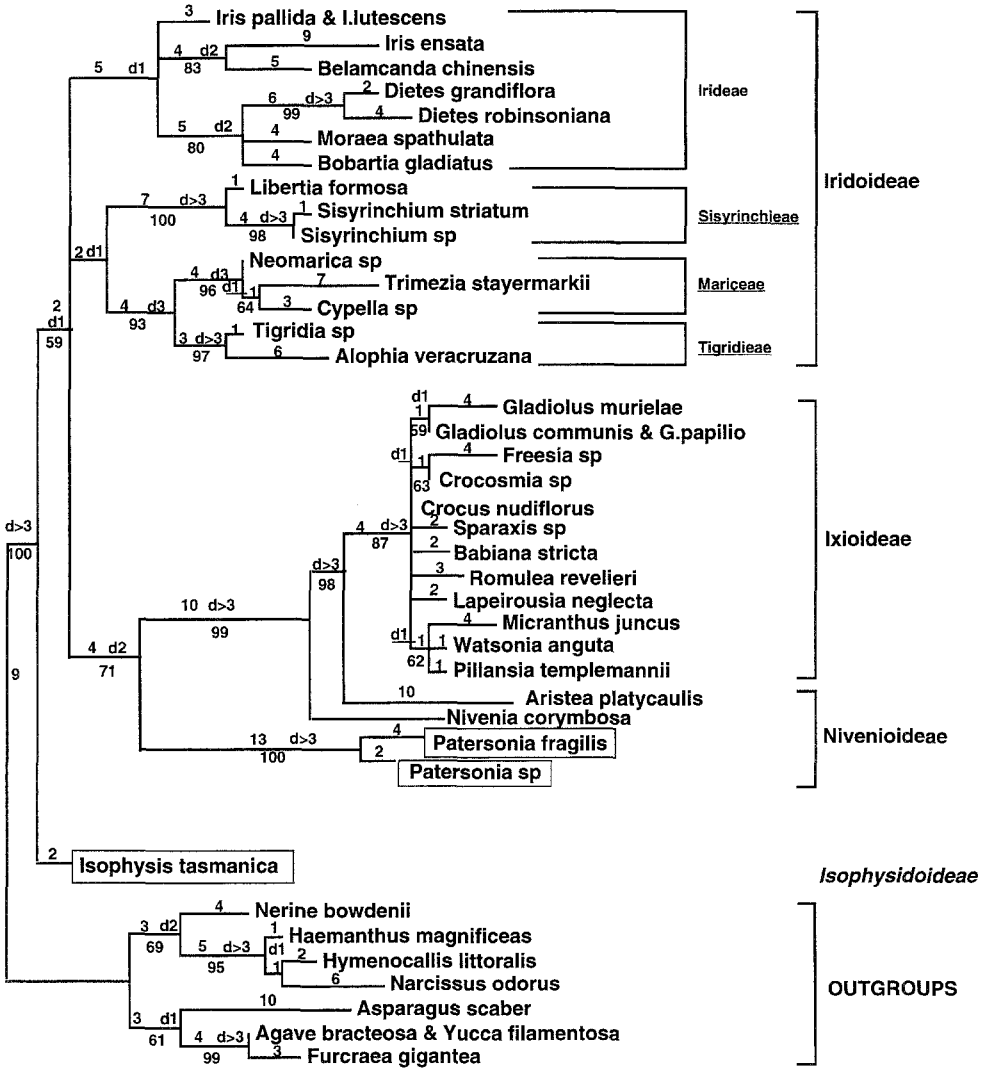

Fig. 3. Strict consensus tree of 223 steps calculated from 6 most parsimonious trees of 221 steps resulting of a branch-and-bound search using PAUP 3.1.1. (gap treated as fifth base). The bootstrap value (indicated if > 50%) of the 1000 replicates appears below the branch. The lengths of the branches as well as their decay values are indicated above the branches. c.i. = 0.758 (0.671 when excluding the cladistically uninformative sites); r.i. = 0.891. The subfamilies (and tribes within the *Iridoideae*) are indicated. American tribes within *Iridoideae* are indicated by underlining and the Oceanian species are shown using boxes

codons, 24% at the second position, and 62% at the third position as determined with MacClade 3.0. The values for consistency, retention and rescaled indices per codon were also calculated with MacClade 3.0 and are indicated in Table 3.

The most variable regions were found to be situated in the 5′ and 3′ regions of the sequence, as already shown by Nadot & al. (1994). Three supplementary nucleotides forming an extra codon were found close to the 3′ end in the *rps* 4 sequences of members of subfamily *Ixioideae* as well as in the sequences of the

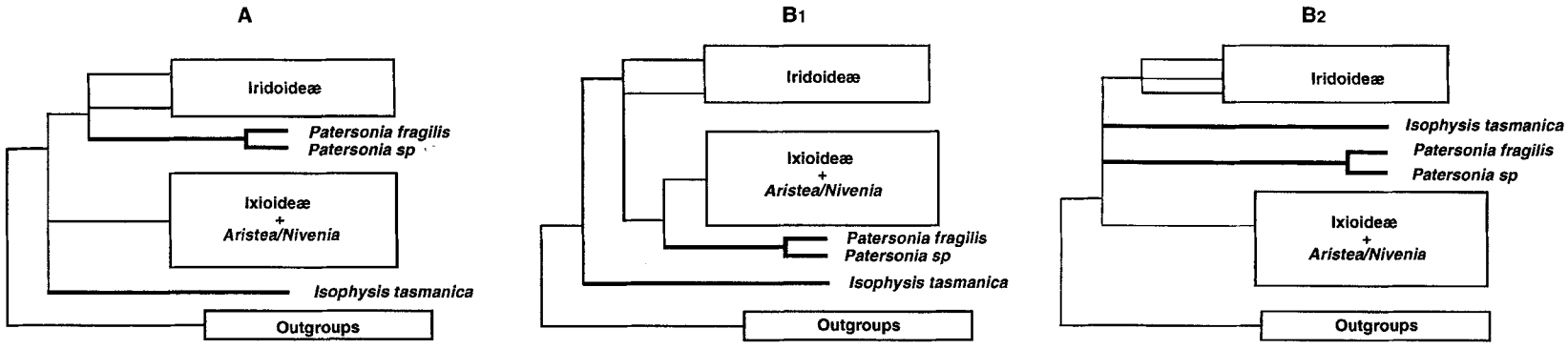

Fig. 4. The three different positions of *Isophysis* and *Patersonia* in the phylogenetic reconstructions of *Iridaceae* according to the way gaps were treated. *A* Gap treated as "missing data"; *B1* gap treated as "fifth base" using PAUP; *B2* gap treated as "fifth base" using the distance method ANATAXIS

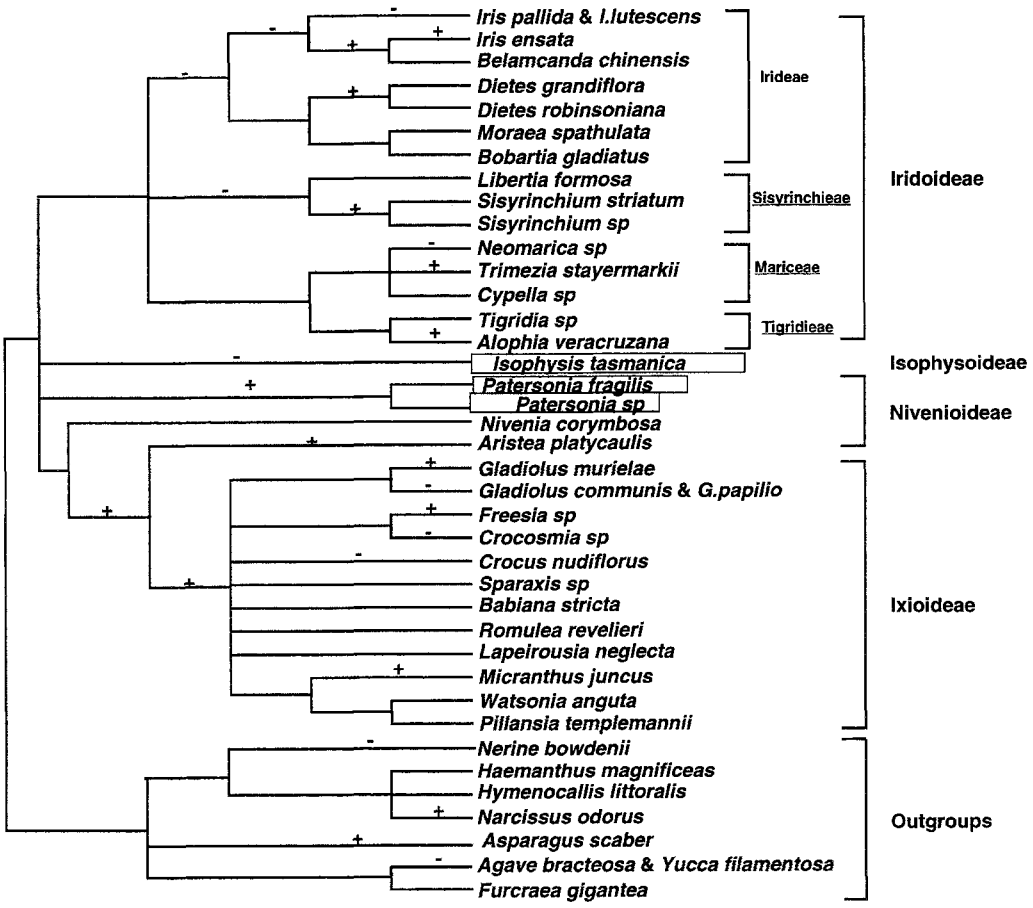

Fig. 5. Distance tree obtained using ANATAXIS (gap treated as fifth base). The subfamilies and the tribes of *Iridoideae* are indicated. Putative fast (+) or slow (–) evolution rates are indicated for all branches. The subfamilies (and tribes within the *Iridoideae*) are indicated. American tribes within *Iridoideae* are indicated by underlining and the Oceanian species are shown using boxes

genera *Aristea* and *Nivenia*. At this same location the *rps* 4 sequences of the two *Patersonia* species display not one but two supplementary codons. These insertions are illustrated in the alignment shown in Fig. 2.

In a preliminary analysis the amino acid sequences were aligned and used for phylogenetic reconstruction. This led to three MP trees of 86 steps with a strict consensus tree of 87 steps (not shown). Because of the lack of discrimination within the clades and of the many polychotomies found, further work was carried out using the nucleotide sequences.

Seven distinct *rps* 4 sequences from species belonging to the order *Asparagales* (see legend to Fig. 3), the closest order to *Liliales* to which *Iridaceae* belong, were chosen to form an outgroup, because preliminary tests (see Material and methods) with PAUP and ANATAXIS confirmed this order to be within the monocots the best outgroup to *Iridaceae*.

When the gaps in the alignment were treated as "fifth base", with a combina-

tion of branch-and-bound and heuristic searches (Tree-Bisection-Reconnection branch-swapping) 6 most parsimonious trees of 221 steps were obtained. The strict consensus tree (Fig. 3) consisted of 223 steps, with a consistency index of 0.758 (0.671, if the cladistically uninformative sites are excluded) and a retention index of 0.891. The robustness of the trees has been estimated both by bootstrapping (Felsenstein 1985) and by decay analysis (Bremer 1988). The corresponding values are given in Fig. 3.

When gaps were alternatively treated as "missing data", no essential differences in the trees were observed for the genus *Patersonia*, which was then found associated to the *Iridoideae* rather than to the *Nivenia/Aristea/Ixioideae* clade; these varying positions of *Patersonia* are illustrated in Fig. 4. The basal position of *Isophysis* to the rest of *Iridaceae* was not strongly supported.

The distance analysis with ANATAXIS was also used twice, alternatively taking into account or leaving out the insertions in *Patersonia, Nivenia, Aristea* and the *Ixioideae*. In the first option, *Patersonia* appears basal in the tree topology (Fig. 5), which places it in a different position from that proposed by PAUP (Fig. 3) using the same option, while in the second option it is joined to the *Iridoideae* as it does with PAUP (see Fig. 4). It can be noted that the species of *Iridoideae* form a clade in which the tribes are resolved according to Goldblatt (1990): *Irideae, Sisyrinchieae* and *Mariceae/Tigridieae*. The second main clade in the ANATAXIS tree clearly groups the *Ixioideae* together with *Aristea* and *Nivenia*.

With 12 synapomorphic sites, both the MP and ANATAXIS results seem therefore to confirm that the species of *Iridaceae* do form a clade. Within this taxon, *Isophysis tasmanica*, a species from Tasmania, appears basal to all other *Iridaceae* that are divided into three main clades grouping species with identical geographical distribution. The first one includes the species of *Iridoideae* from the Old World. The second clade groups the genera that are distributed within the tribes *Tigridieae, Mariceae* and *Sisyrinchieae* of the *Iridoideae* (Goldblatt 1990) and that occur in South and Central America. The third clade, consisting of *Nivenia, Aristea*, and of the subfam. *Ixioideae* Old World species, is very well supported and appears in 99% of the bootstrap replicates with a decay value of its supporting branch greater than 3. Its internal cluster of all *Ixioideae* (Goldblatt 1990) is also well supported with a bootstrap value of 87 though *rps* 4 provides poor resolution within this group. The very variable position of the Australian genus *Patersonia* was previously emphasized (Fig. 4).

Discussion

In this work we have examined the phylogeny of *Iridaceae* using the *rps* 4 gene as a phylogenetic tool. The results presented here confirm that despite the relatively small size of the *rps* 4 gene (making for easier and faster sequencing of the samples), the amount of information displayed by this gene is sufficient to reconstruct a phylogeny of the *Iridaceae*. In addition, the sequences showed a low level of homoplasy within the gene, as shown here by the indices (Table 3, Fig. 3). For this reconstruction we used in parallel two methods: the maximum parsimony principle as implemented within PAUP (Swofford 1993) and an improved distance method, ANATAXIS (Bittar & Carter 1994, Bittar 1995). Globally equivalent trees were obtained with both the PAUP and ANATAXIS programs.

A parsimony analysis was also performed using only the first and the second bases because the third codon positions may evolve more rapidly than the second and first codon positions in protein-coding genes. We obtained one tree of 74 steps (not shown), the topology of which was almost identical to that obtained with PAUP when treating gaps as fifth base. The few differences occurred mostly in the subfam. *Iridoideae*, which treated in this way were not resolved as a clade. The conservation of trees topology when only the first and the second bases are used had been shown already by RITLAND & CLEGG (1987) with the other plastid genes *rbc*L and *atp*B.

As mentioned in the Material and methods section, all tests were performed using eight species of *Asparagales* as outgroups because in preliminary tests *Iridaceae* appeared to be closer to *Asparagales* than to the order *Liliales* in which they are usually included. This is in agreement with other works based either on molecular (CHASE & al. 1993, 1995; DUVALL & al. 1993) or morphological (RUDALL 1994) data. Should convergent data confirm these links with *Asparagales*, then *Iridaceae* might be taken out from *Liliales* and their systematic position revised.

Although variable bootstrap values were found according to the selection of outgroups used, both reconstruction methods indicated a monophyletic origin of *Iridaceae*, in agreement with other systematic and phylogenetic studies of this family. Another important feature of our results is the almost perfect correspondence of the clusters of genera found in our phylogenetic trees with the classification proposed by GOLDBLATT (1990). Apart from a few exceptions or positional uncertainties discussed below, it may be said that molecular criteria fit perfectly with the results of morphological and anatomical studies.

Because of its basal (parsimony trees) or nearly-basal ANATAXIS positions in *Iridaceae, Isophysis* belongs to the family and may have diverged at an early stage from primitive *Iridaceae*. However, in a phylogeny by CHASE & al. (1995) using the *rbc*L gene and including six *Iridaceae, Isophysis tasmanica* appears as a sister group to a cluster grouping five other *Iridaceae* plus representatives of *Ixioliriaceae* and *Doryanthaceae*. If *Isophysis* really is a member of the *Iridaceae*, CHASE's results suggest that the family is paraphyletic. This contradicts our results and the classifications of DAHLGREN (1985) and GOLDBLATT (1990) and may originate either from the smaller sampling used by CHASE & al. (1995) or the fact that certain representatives were not included in our study (e.g., *Ixioliriaceae* or *Doryanthaceae*).

The trees do not confirm the monophyly of the *Nivenioideae* and emphasize the links of the three selected genera of this subfamily, *Nivenia, Patersonia* and *Aristea* with *Ixioideae*. *Nivenioideae* appear therefore to be paraphyletic and to consist of early emerging species sharing a common ancestor with *Ixioideae*. One of the reasons for this clustering is the presence of a similar insertion at the same place in the *rps* 4 sequence.

Insertions in this gene are not uncommon in higher plants, as for instance in its 3′ end in *Cycas revoluta* THUNB., *Taxus baccata* L., *Luzula sylvatica* (HUDSON) GODIN (NADOT 1994) and *Epifagus virginianus* BART. (WOLFE & al. 1992) or in its 5′ region as in *Cyperus vegetus* BART., *Taxus baccata* and *Luzula sylvatica* (NADOT 1994). In this work an identical or almost identical 3-nucleotide insertion was found in the *rps* 4 sequence at the same precise location in the *Ixioideae*, in *Aristea* and in *Nivenia*; in *Patersonia* a comparable 6-nucleotide insertion occurs at

the same position (Fig. 2). As mentioned in the Results section, the evolutionary implication of one (two for *Patersonia*) supplementary aminoacid in the putative protein led us to carefully examine how these 3- (or 6-) nucleotides insertions (resulting in gaps for other *Iridaceae* in the alignments) should be considered in the reconstruction programs.

Whichever way the gaps were treated ("missing data" or "fifth base"), *Ixioideae* and *Aristea* plus *Nivenia* remain associated, while this is not true for *Patersonia*. When the insertions are not taken into account, parsimony and distance trees suggest that the insertions found in the cluster *Nivenia* + *Aristea* + *Ixioideae* and the one found in *Patersonia* would result from a convergence. When the gaps are considered as a fifth base, ANATAXIS and the parsimony consensus trees are still consistent with the idea of a clade grouping *Nivenia, Aristea* and *Ixioideae* (Fig. 3) whereas differences exist in both PAUP and ANATAXIS trees concerning *Patersonia*: this genus is included by the parsimony consensus tree into the clade formed by *Ixioideae, Aristea* and *Nivenia* whereas it remains in the unresolved basal polychotomy displayed by ANATAXIS. Despite this discrepancy, the evolutionary importance of an insertion for the resulting *rps* 4 protein leads us to favour the "fifth base" option and to hypothesise an unique mutational event linking *Nivenia, Aristea*, the *Ixioideae* and, most probably, *Patersonia*.

Both parsimony and distance trees show that *Ixioideae* are in a terminal position relative to the other species of *Iridaceae*. This confirms the analysis done by Goldblatt (1990), but disagrees with Rudall's (1994) suggestion that *Ixioideae* have a basal position. The poor resolution within *Ixioideae* suggests a rapid radiation of the subfamily and indicates a evolutionary scenario different from that of *Iridoideae*; analysis of the clade with ANATAXIS suggests a few evolution rate at the origin of the radiation (Fig. 5).

Whichever phylogeny program was used, the molecular trees differentiate the species of the subfam. *Iridoideae* by their geographical distribution. The New and Old World species belong to different clades even if the phylogenetic relationships between them may vary slightly according to the program used. However, the presence of a basal trichotomy for the *Iridoideae* in the ANATAXIS tree, together with indications of fast evolution at this level, suggest that the divergence between all these groups may have occurred over a relatively short period of time.

The closeness of *Bobartia* to *Moraea* and *Dietes* is supported with statistical methods (bootstrap value of 80%, decay value of 2). This disagrees with Goldblatt's (1990) positioning of *Bobartia* in the tribe *Sisyrinchieae* sensu Goldblatt, and with Rudall's (1994) suggestion that *Bobartia* forms a sister group to the *Sisyrinchieae*, although Goldblatt & Rudall (1992) agreed that the position of *Bobartia* was controversial. Dahlgren & al. (1985) had suggested that *Bobartia* might have its closest relatives in *Iridoideae* near *Dietes*. Our results support Dahlgren's hypothesis on the closeness of *Bobartia* and *Dietes*.

One may speculate on the biogeographical evolution of *Iridaceae* correlating molecular data (especially the existence of the insertion in the *rps* 4 sequence), phylogenetical reconstructions, paleogeography of the Southern hemisphere and the geographical distribution of extant members of the family. One may for instance hypothesize that South African ancestors to extant *Iridaceae* generated first an early South American branch and that later on some of these ancestors

underwent a 3-nucleotide insertion event that led to the group from which present-day *Nivenioideae* (except *Pattersonia* whose position is problematic) and *Ixioideae* emerged, with no further possibility of migrating to South America. Both *Iridoideae* and *Ixioideae* lineages from this South African stock spread northwards to generate the existing *Iridaceae* of Europe, Asia and North America. An alternative hypothesis to explain the origin of *Iridaceae* might involve waves of migration of species from an original lineage in Oceania of which *Isophysis* and *Patersonia* would be current day representatives.

In conclusion, the phylogeny of the *Iridaceae* based on the *rps* 4 nucleotide sequences presented here confirms the monophyletic origin of the *Iridaceae* proposed in other systematic and phylogenetic works (Dahlgren & al. 1985, Goldblatt 1990, Rudall 1994); *Isophysis tasmanica* could be the first actual emerging species of the *Iridaceae*, which might support the proposal of a subfam. *Isophysioideae* by Dahlgren & al. (1985) and Goldblatt (1990). The subdivisions within *Iridoideae* agree with the observed geographical distribution of its species, the tribe *Irideae* being separated from the American tribes *Mariceae, Tigridieae* and *Sisyrinchieae*. The position in the tribe *Irideae* is very well supported and excludes *Bobartia* from the American *Sisyrinchieae* in which it has been placed. The subfam. *Ixioideae* is undoubtedly a natural clade that emerged recently within the *Iridaceae* and seems to be very strongly associated to *Aristea*, this last genus being clearly separated from *Nivenia* and *Patersonia*. This work has allowed us to confirm that the *rps* 4 gene can be used in molecular phylogenies to solve taxonomic problems at the intra-familial level and it is probable that other plant phylogenies can be based on sequences of this plastid gene.

We wish to thank Dr P. Goldblatt and Dr M. Chase for their comments on this work and for their considerable help in the sampling. Plant material was also provided by Mr M. Boussard whom we thank here. Helpful comments and suggestions on the manuscript from James Tregear, Donal O'Sullivan, Hervé Philippe and Marie-France Jubier are acknowledged. Tatiana T. de Souza Chies is supported by a fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES (Brasília, Brazil). This research has been supported by funds from the French Ministère de l'Enseignement Supérieur et de la Recherche – URA CNRS 1128 (Paris, France).

References

Bittar, G., 1995: Phylogénie: histoire, concepts, principes et présentation de nouvelles méthodes d'analyse numérique de l'évolution moléculaire. – Ph.D. Dissertation, Université de Genève, Switzerland.

– Carter, L., 1994: New probabilistic and numerical phenetics methods for inferring natural groupings and phylogenesis. – In: Proceedings of the International Meeting "Ecology and statistical methods", Niort, France, pp. 145–150.

Bremer, K., 1988: The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. – Evolution **42**: 795–803.

Chase, M. W., Soltis, D. W., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y. L., Kron, K. A., Retting, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Systma, K. J., Michaels, 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. J., 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. Missouri Bot. Gard. **80**: 528–580.

– Duvall, M. R., Hills, H. G., Conran, J. G., Cox, A. V., Eguiarte, L. E., Hartwell, J., Fay, M. F., Caddick, L. R., Cameron, K. M., Hoot, S., 1995: Molecular phylogenetics of *Lilianae*. – In Rudall, P. J., Cribb, P. J., Cutler, D. F., Humphries, C. J., (Eds): Monocotyledons: systematics and evolution, **1**: pp. 109–137. – Kew: Royal Botanic Gardens.

Clegg, M. T., Zurawski, G., 1992: Chloroplast DNA and the study of plant phylogeny: present status and future prospects. – In Soltis, P. S., Soltis, D. E., Doyle, J. J., (Eds): Molecular systematics of plants, pp. 1–13. – New York: Chapman & Hall.

– Gaut, B. S., Learn, G. H. J., Morton, B. R., 1994: Rates and patterns of chloroplast DNA evolution. – Proc. Natl. Acad. Sci. USA **91**: 6795–6801.

Cronquist, A., 1981: An integrated system of classification of flowering plants. – New York: Columbia University Press.

Dahlgren, R. M. T., Clifford, H. T., Yeo, P. F., 1985: The families of monocotyledons. – Berlin, Heidelberg, New York: Springer.

Doyle, J. J., Doyle, J. L., 1987: A rapid DNA isolation procedure for small quantities of fresh leaf tissue. – Phytochem. Bull. **19**: 11–15.

Duvall, M. R., Clegg, M. T., Chase, M. W., Clark, W. D., Kress, W. J., Hills, H. G., Eguiarte, L. E., Smith, J. F., Gaut, B. S., Zimmer, E. A., Learn, G. H., 1993: Phylogenetic hypothesis for the monocotyledons constructed from *rbc*L sequence data. – Ann. Missouri Bot. Gard. **80**: 607–619.

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

Gielly, L., Taberlet, P., 1994: The use of chloroplast DNA to resolve plant phylogenies: non-coding versus *rbc*L sequences. – Molec. Biol. Evol. **11**: 769–777.

Goldblatt, P., 1990: Phylogeny and classification of *Iridaceae*. – Ann. Missouri Bot. Gard. **77**: 607–627.

– 1991: An overview of the systematics, phylogeny and biology of the African *Iridaceae*. – Contr. Bolus Herb. **13**: 1–74.

– Rudall, P., 1992: Relationships of *Bobartia*. – S. African J. Bot. **58**: 304–309.

– Mannig, J. C., Bernhart, P., 1995: Pollination biology of *Lapeirousia* subgenus *Lapeirousia* (*Iridaceae*) in southern Africa; floral divergence and adaptation for long-tongued fly pollination. – Ann. Missouri Bot. Gard. **82**: 517–534.

Ham, R. C. H. J. v., Hart, H., Mes, T. H. M., Sandbrink, J. M., 1994: Molecular evolution of non-coding regions of the chloroplast genome in the *Crassulaceae* and related species. – Curr. Genet. **25**: 558–566.

Hoot, S. B., Culham, A., Crane, P. R., 1995: The utility of *atp*B gene sequences in resolving phylogenetic relationships: comparison with *rbc*L and 18S ribosomal DNA sequences in the *Lardizabalaceae*. – Ann. Missouri Bot. Gard. **82**: 194–207.

Johnson, L. A., Soltis, D. E., 1994: *mat*K DNA sequences and phylogenetic reconstruction in *Saxifragaceae* s.str. – Syst. Bot. **19**: 143–156.

– – 1995: Phylogenetic inference in *Saxifragaceae* sensu stricto and *Gilia* (*Polemoniaceae*) using *mat*K sequences. – Ann. Missouri Bot. Gard. **82**: 149–175.

Lamppa, G. K., Bendich, A. J., 1979: Chloroplast DNA sequence homologies among vascular plants. – Pl. Physiol. **63**: 660–668.

Lavin, M., Doyle, J. J., Palmer, J. D., 1990: Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the *Leguminosae* subfam. *Papilionoideae*. – Evolution **44**: 390–402.

Maddison, W. P., Maddison, D. R., 1992: MacClade: analysis of phylogeny and character evolution. Version 3.0. – Sunderland, Mass.: Sinauer.

Manen, J. F., Natali, A., Ehrendorfer, F., 1994: Phylogeny of *Rubiaceae-Rubieae* inferred from the sequence of a cpDNA intergene region. – Pl. Syst. Evol. **190**: 195–211.

Mes, T. H. M., Hart, H., 1994: *Sedum surculosum* and *S. jaccardianum* (*Crassulaceae*) share a unique 70 bp deletion in the chloroplast DNA *trn*L (UAA) – *trn*F (GAA) intergenic spacer. – Pl. Syst. Evol. **193**: 213–221.

NADOT, S., 1994: Utilisation de l'outil moléculaire pour la reconstruction d'une phylogénie des plantes monocotylédones en général et de la famille des *Poaceae* (Graminées) en particulier. – Ph.D. in Sciences – Dissertation, Université de Paris XI, Orsay, France.

– BAJON, R., LEJEUNE, B., 1994: The chloroplast gene *rps* 4 as a tool for the study of *Poaceae* phylogeny. – Pl. Syst. Evol. **191**: 27–38.

– BITTAR, G., CARTER, L., LACROIX, R., LEJEUNE, B., 1995: A phylogenetic analysis of monocotyledons based on the chloroplast gene *rps* 4, using parsimony and a new numerical phenetics method. – Molec. Phylogenet. Evol. **4**: 257–282.

OLMSTEAD, R. G., REEVES, P. A., 1995: Evidence for the polyphyly of the *Scrophulariaceae* based on chloroplast *rbc*L and *ndh*F sequences. – Ann. Missouri Bot. Gard. **82**: 176–193.

PALMER, J. D., ZAMIR, D., 1982: Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicum*. – Proc. Natl. Acad. Sci. USA **79**: 5006–5010.

– STEIN, D. B., 1986: Conservation of chloroplast genome structure among vascular plants. – Curr. Genet. **10**: 823–833.

RITLAND, K., CLEGG, M. T., 1987: Evolutionary analysis of plant DNA sequences. – Amer. Naturalist **130**: S74–S100.

RUDALL, P., 1994: Anatomy and systematics of *Iridaceae*. – Bot. J. Linn. Soc. **114**: 1–21.

SOLTIS, D. E., SOLTIS, P. S., NESS, B. D., 1989: Chloroplast-DNA variation and multiple origins of autopolyploidy in *Heuchera micrantha* (*Saxifragaceae*). – Evolution **43**: 650–656.

STEELE, K. P., VILGALYS, R., 1994: Phylogenetic analyses of *Polemoniaceae* using nucleotide sequences of the plastid gene *mat*k. – Syst. Bot. **19**: 126–142.

SWOFFORD, D. L., 1993: PAUP 3.1: phylogenetic analysis using parsimony. – Washington, D.C.: Smithsonian Institution Press.

TABERLET, P., GIELLY, L., PAUTOU, G., BOUVET, J., 1991: Universal primers for amplification of three non-coding regions of chloroplast DNA. – Pl. Molec. Biol. **17**: 1105–1109.

THOMPSON, J. D., HIGGINS, D. G., GIBSON, T. J., 1994: Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. – Nucl. Acids Res. **22**: 4673–4680.

WOLFE, K. H., MORDEN, C. W., PALMER, J. D., 1991: Ins and outs of plastid genome evolution. – Curr. Opinion Genet. Developm. **1**: 523–529.

– – EMS, S. C., PALMER, J. D., 1992: Rapid evolution of the plastid translational apparatus in a non-photosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. – J. Molec. Evol. **35**: 304–317.

ZURAWSKI, G., CLEGG, M. T., 1987: Evolution of higher-plant chloroplast DNA-encoded genes: implications for structure-function and phylogenetic studies. – Annu. Rev. Pl. Physiol. **38**: 391–418.

Addresses of the authors: TATIANA T. SOUZA-CHIES, SOPHIE NADOT, EVELYNE BESIN, BERNARD LEJEUNE (correspondence), Institut de Biotechnologie des Plantes, Batiment 630, Université de Parix XI, F-91405 Orsay, Cedex, France. – GABRIEL BITTAR, LEIGH CARTER, Département de Chimie Physique, Université de Genève, 30 quai Ansermet, CH-1205, Genève, Switzerland.

Accepted April 5, 1996 by F. EHRENDORFER

Verleger: Springer-Verlag KG, Sachsenplatz 4-6, A-1201 Wien. — Herausgeber: Prof. Dr. Friedrich Ehrendorfer, Institut für Botanik und Botanischer Garten der Universität Wien, Rennweg 14, A-1030 Wien. Redaktion: Rennweg 14, A-1030 Wien. — Hersteller: Adolf Holzhausens Nfg., Kandlgasse 19-21, A-1070 Wien, Verlagsort: Wien. — Herstellungsort: Wien. — Printed in Austria

Offenlegung gem. §25 Abs. 1 bis 3 Mediengesetzes.

Unternehmensgegenstand: Verlag von wissenschaftlichen Büchern und Zeitschriften.

An der Springer-Verlag KG ist beteiligt: Dr. Konrad F. Springer, Sachsenplatz 4-6, A-1021-Wien, als Kommanditist zu 52,38%. Geschäftsführer: Dr. Konrad F. Springer, Prof. Dr. Dietrich Götze, Ing. Wolfram F. Joos, Dipl.-Kfm. Claus Michaletz und Rudolf Siegle. Alle Sachsenplatz 4-6, A-1201 Wien