Phylogeny and disjunct distribution: evolution of *Saintpaulia* (Gesneriaceae)

MICHAEL MÖLLER[1] AND QUENTIN C. B. CRONK[1,2*]

[1]*Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK* (m.moeller@rbge.org.uk)
[2]*Institute of Cell and Molecular Biology, The University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JH, UK* (q.cronk@rbge.org.uk)

SUMMARY

The molecular phylogeny of African violets (*Saintpaulia* H. Wendl.), based on ribosomal DNA internal transcribed spacer (ITS) sequences, follows the disjunct biogeography of the genus. Sequence analysis by parsimony of 19 accessions, representing 17 currently recognized *Saintpaulia* species, resulted in four trees of 182 steps. The first major division is between *S. goetzeana*, from the Uluguru Mts, Tanzania, and the rest of the genus. The basal position of *S. goetzeana*, and its putative primitive characters, may indicate an Uluguru origin for *Saintpaulia* and subsequent colonization of the more northerly mountains. Of the remainder, *S. teitensis*, from the Teita Hills of Kenya, is sister taxon to the other species (which occur mainly in the Usambara Mts of north-east Tanzania). A group of nine Usambaran species that we call the '*ionantha* complex' show minimal ITS genetic differentiation and are also taxonomically critical. Species diversity in the Usambara Mts appears to be the result of rapid, recent (possibly Pleistocene) radiation. This study reveals the limitations of ITS sequences for elucidating the radiation of poorly differentiated species (the *ionantha* complex). However, the molecular data strongly suggest that conservation of the Uluguru and Teita populations is essential for the protection of the full range of diversity within the genus.

1. INTRODUCTION

The flowering plant genus *Saintpaulia* H. Wendl. (African violet: Gesneriaceae, subfamily Cyrtandroideae) was first described from material collected at Tanga in north-east Tanzania in 1892 by Baron Walter von Saint Paul-Illaire (1860–1940), and sent, in 1893, to Hermann Wendland, then Director of the Royal Botanic Garden at Herrenhausen, Germany (Baatvik 1993). The plant quickly became popular because of its attractive free-flowering habit, and the ease of propagation from leaf cuttings, and is now the centre of a large horticultural industry. The two early introductions, *S. ionantha* and *S. confusa*, were crossed, creating considerable variation among hybrids from which cultivars could be selected. There is wide cross-fertility within the genus (Arisumi 1964) and later several other species were used in breeding efforts.

In 1958, B. L. Burtt recognized 19 species based on morphology of living and herbarium material. However, *S. amaniensis* was later included in *S. magungensis* as a synonym (Burtt 1964), and a re-examination of existing and new material resulted in an increase of the number of species to 20 by the recognition of *S. brevipilosa* and *S. rupicola* (Burtt 1964). Since then, little taxonomic revision has been carried out, and the pioneering treatment by B. L. Burtt is still the standard work on *Saintpaulia* systematics.

These species generally occur in very small, local populations, often in upland areas (800–2000 m), and some are known from a single locality only (Burtt 1958; 1964). However, in recent years, this picture has had to be modified, and it has become clear that some taxa are more widespread in lowland areas (Clarke 1998), although again in small colonies, typically occupying less than 100 m^2. Localities in Tanzania include the Mafi Hills, the Pangani Falls and the Gendagenda Forests (Handeni District) where *Saintpaulia tongwensis*, previously known only from Mt Tongwe, was discovered in 1991 (Burgess *et al*. 1992). What may prove to be a new species (or even two new species), with affinities to *S. rupicola*, has been discovered in Kenya at Kacharoni (Kalifi District) and at the Mwachi Forest Reserve (Kwale District) (Clarke 1998; Eastwood & Maunder 1995). More remarkable still are two discoveries at southerly locations in Tanzania: at the Sanje Falls (Ulanga District, 7°46′S) found by R. Polhill and J. Lovett, and at the Kiwengoma Forest Reserve (Matumbi massif, Rufiji District 8°21′S) by D. Shiel (Burgess *et al*. 1992). Neither of these plants can be satisfactorily distinguished from *S. ionantha* (B. L. Burtt, personal communication). These recent finds, while

*Author for correspondence.

Table 1. *Biogeographic distribution of species of* Streptocarpus *subgenus* Streptocarpella *and* Saintpaulia *at four localities in Tanzania and Kenya* (*Hilliard & Burtt 1971*)

	Teita Hills	Usambara Mts	Nguru Mts	Uluguru Mts
Saintpaulia (all spp.)	+	+	+	+
Streptocarpus				
S. bambuseti	–	–	+	–
S. buchananii	–	–	–	–
S. caulescens	+	+	–	+
S. euanthus	–	–	–	+
S. glandulosissimus	+	+	+	+
S. gonjaënsis	–	+	–	–
S. hirsutissimus	–	–	–	+
S. holstii	–	–	+	–
S. inflatus	–	–	–	–
S. kimbozanus	–	–	–	–
S. kirkii	+	+	–	+
S. parensis	–	?	–	–
S. saxorum	+	+	+	+
S. stomandrus	–	–	+	–

indicating that *S. ionantha* itself is more widespread than the other species, do not obscure the essentially Usambaran nature of the *S. ionantha* complex.

Recent molecular systematic work (Möller & Cronk 1997) on African Gesneriaceae revealed a close relationship between *Saintpaulia* and African caulescent *Streptocarpus* species, with *Saintpaulia* nested within *Streptocarpus* subgenus *Streptocarpella*. These data suggest the evolution of *Saintpaulia* from *Streptocarpus* subgenus *Streptocarpella*. The differences in flower and vegetative characters are thought to be due to ecological adaptation leading to a relatively rapid radiation of *Saintpaulia* (Möller & Cronk 1997).

The disjunct distribution of *Saintpaulia* is similar to that of many species of *Streptocarpus* subgenus *Streptocarpella* (table 1). *Streptocarpus saxorum* and *S. glandulosissimus* even have a wider species distribution than the entire *Saintpaulia* genus. *Saintpaulia* thus forms a geographically restricted aggregate, with four main areas (in Tanzania, the Uluguru Mts, Nguru Mts, and Usambara Mts; in Kenya: Teita Hills) and scattered further localities along the coast of Kenya and Tanzania between the Tana and Rufiji rivers. The centre of species diversity is in the Usambara Mts of north-east Tanzania (Burtt 1958; Iversen 1991) (figure 1). This particular disjunct distribution (Uluguru Mts–Nguru Mts–Usambara Mts–Teita Hills) is known outside *Saintpaulia* and is a common phenomenon (Lovett & Friis 1996).

With respect to habitat, *Saintpaulia* species occur in a wide range of forest types, from montane to submontane and lowland, and appear to be more restricted by physical substrate than vegetation type. They grow on rocky mountain peaks amongst moss, or on wet rocks near water, often on cliffs or in gorges (Burtt 1958, 1964). While generally found on rock, either gneiss (Eastern Arc Mts) or limestone (coastal forests), they occasionally occur as epiphytes, for example *S. grotei* at Bamba in the Usambara Mts at 800 m, epiphytic on *Encephalartos* sp. and *Pandanus* sp. (D. Percy, personal communication). Some species are relatively drought tolerant, and in *S. ionantha* crassulacean acid metabolism (CAM) has been reported (Guralnick *et al*. 1986), which is often associated with drought tolerant succulents.

The Uluguru, Nguru, Usambara Mts and Teita Hills form part of a chain, known as the 'Eastern Arc' group of mountains (Lovett 1990), which also include the Pare, Ukaguru and Usagara Mts. However, with the exception of *Saintpaulia pusilla* in the Ukaguru Mts, *Saintpaulia* species have not been recorded from these latter mountains. The chain is formed of crystalline rocks, and the forests covering them are influenced by a fairly stable Indian Ocean climate (Lovett 1990).

The Uluguru Mts are just south of the town of Morogoro. The southern part, the Lukwangule Plateau, is the highest area, rising to 2668 m at Kimhandu. The vegetation is moist lowland forest, grading into moist submontane and afromontane rainforest, and high-altitude grassland on the Lukwangule Plateau (Davis *et al.* 1994). Different rainfall patterns occur, with 1300–2900 mm on the eastern side (no dry season), 3000 mm on the main ridge (1800–2600 m), and 800–2000 mm in rain-shadow on the western slopes and foothills (dry season of 2–5 months) (Pócs 1974). The forest at higher altitudes is largely undisturbed, but population pressure is increasingly threatening the lower slopes (Lovett 1990).

The Nguru Mts are formed by four major groups of mountains separated by deep rocky valleys, located in the Morogoro District (altitude 400–2400 m) with lowland moist and dry forest grading into high-altitude afromontane rainforest, and a unimodal rainfall pattern with a rainy season from November to May of 1250–2500 mm of rain (Davis *et al.* 1994; Lovett 1996). The threat to the indigenous vegetation is similar to that in the Uluguru Mts (Lovett 1990).

The East Usambara Mts are a group of low mountains in the Tanga Region, north-east Tanzania, ranging from 150–1506 m in altitude. The vegetation is

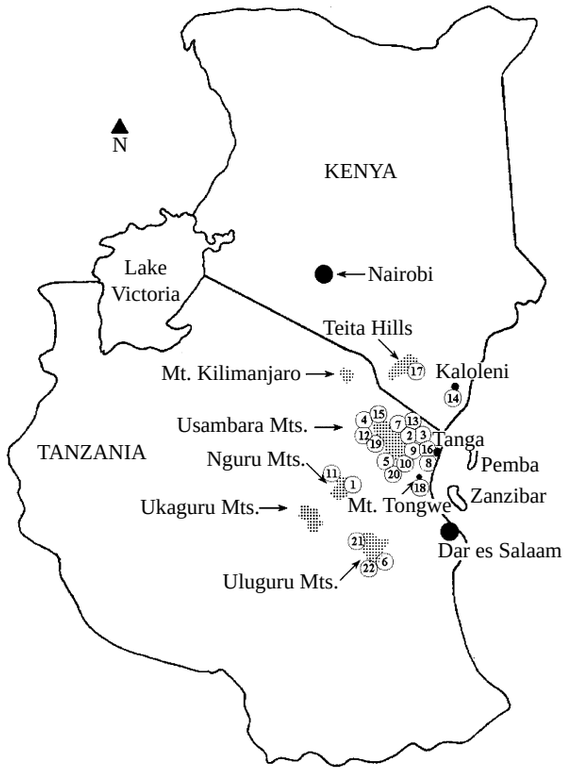

Figure 1. Guide to the geographical distribution of *Saintpaulia* species described to date (numbers 1–19 referring to *Saintpaulia* species listed in table 2 (number 20 = *Saintpaulia confusa*, no. 21 = *S. inconspicua*, no. 22 = *S. pusilla*). The position of the number is an indication of the (very approximate) general distribution of the species.

of lowland semi-deciduous and evergreen submontane forest with an annual rainfall of 1250–2500 mm (Davis *et al*. 1994; Lovett 1990). The West Usambara Mts are a large upland block rising nearly to 2300 m. The vegetation on this rain-shadow side of the Usambara Mts changes to woodland, thicket and scrub, with a bimodal rainfall peaking in November and April with 2000 mm of rain a year in the wettest areas, falling to less than 600 mm in the drier areas (Lovett 1996). According to Lovett (1990), the West Usambara Mts are threatened more by high population pressure than the East Usambara Mts. Past logging programmes have depleted large areas of Usambara forest, although there is now greater emphasis on forest conservation, particularly in water-catchment areas (Bjorndalen 1992).

The Teita Hills rise abruptly from the plains in the Teita–Taveta District of the Coast Province in Kenya. Most of the lower hills are under intensive cultivation, but between 1200–2228 m moist-to-semi-dry submontane afromontane rainforest still survives (Davis *et al.* 1994). The rainfall is 1600–2450 mm with a distinct dry season of 2–3 months, but mist or cloud is usually present throughout the year (Beentje 1990).

Because of the widespread interest in this genus and the difficult taxonomy (in particular, the problematic morphological species delimitation), we decided to extend the preliminary phylogenetic work previously reported (Möller & Cronk 1997) in order to (i) elucidate the pattern of diversification within the genus, (ii) investigate species relationships, and (iii) see whether the disjunct distribution of *Saintpaulia* is reflected in its phylogeny. Further sampling was therefore conducted to include *Saintpaulia* species from all the main areas of distribution listed above.

2. MATERIALS AND METHODS

The methods used follow Möller & Cronk (1997). These are summarized below.

(*a*) *Origin and choice of plant material*

Plant material was taken from living plants of the research collection held at the Royal Botanic Garden, Edinburgh (RBGE), except for material of *Saintpaulia teitensis* and *S. goetzeana*. Identifications were kindly confirmed by B. L. Burtt. For the taxa analysed, voucher herbarium specimens were prepared, flowers preserved in Kopenhagen mixture in a spirit collection, and photographs of flowering specimens taken and deposited in the RBGE library. Material of *S. teitensis* was taken from a herbarium specimen held at the herbarium in Edinburgh (*E*), and silica gel-dried *S. goetzeana* material was kindly provided by Maryjane Evans of the American Gloxinia and Gesneriad Society, Inc.

The outgroup taxa should be systematically close enough to the taxa under observation to allow sequence alignment and yet distantly enough related to enable unequivocal rooting of the tree. As a result of the original work on the phylogenetic relationship between *Saintpaulia* and *Streptocarpus* (Möller & Cronk 1997), *Streptocarpus saxorum* and *Streptocarpus caulescens* were chosen as outgroups. African species of *Streptocarpus* subgenus *Streptocarpella* were found to be the closest relatives to *Saintpaulia* and the two *Streptocarpus* species chosen represented two different groups of those caulescent *Streptocarpella*. They share an identical chromosome number ($2n = 30$) and seed morphological features (verruculose seeds) with *Saintpaulia*.

Nineteen accessions of *Saintpaulia*, representing 17 species, were chosen as the ingroup for this study (table 2). *S. magungensis* var. *minima*, formerly described as *S. amaniensis* but later included in *S. magungensis* (Burtt 1964), and *S.* 'Sigi Falls', (assumed to be *S. ionantha* or a close relative) were included to investigate their phylogenetic rank within the genus. *Saintpaulia pusilla* and *S. inconspicua*, from the Uluguru Mts, could not be included because they are not in cultivation, and DNA extraction and amplification from herbarium material failed repeatedly. The only other recognized species not included is *S. confusa*. This is believed to be very closely related to *S. ionantha* and other species in this study (the *ionantha* complex).

(*b*) *DNA extraction, PCR and sequencing*

Fresh leaf material or dry herbarium material of one plant representing each accession was used for total DNA extraction using a modified CTAB procedure of Doyle & Doyle (1987) with no further purification.

The complete ITS region was amplified using the polymerase chain reaction (PCR), and for each taxon forward and reverse sequencing reactions were performed using the PCR primer 'ITS 3P': GCA TCG ATG AAG AAC GTA GC, primer 'ITS 2G'; GTG ACG CCC AGG CAG ACG T, primer 'ITS 5P'; GGA AGG AGA AGT CGT AAC AAG G, primer ITS 8P': CAC GCT TCT CCA GAC TAC A, for sequence confirmation.

Table 2. *Accessions of* Streptocarpus *and* Saintpaulia *examined for ITS 1 and ITS 2 sequence variation*

no.	taxon	origin: distribution / altitude	RBGE accession no.[a]
1	*Saintpaulia brevipilosa* B. L.Burtt	Tanzania: Nguru Mts, Lulaga, Mt Kanga	1970 0909
2	*Saintpaulia difficilis* B. L.Burtt	Tanzania: E. Usambara Mts, Sigi River, Monga, 900 m	1987 2176
3	*Saintpaulia diplotricha* B. L.Burtt	Tanzania: NE Usambara Mts, Maweni, Tanga, 1000 m	1987 2172B
4	*Saintpaulia grandifolia* B. L.Burtt	Tanzania: W. Usambara Mts, Lutindi	1985 0678
5	*Saintpaulia grotei* Engl.	Tanzania: E. Usambara Mts, Amani, Mt Mlinga, 1080 m	1987 2171
6	*Saintpaulia goetzeana* Engl.	Tanzania: Uluguru Mts, Lukwangule Plateau, 1300–2000 m	1997 1201
7	*Saintpaulia intermedia* B. L.Burtt	Tanzania: E. Usambara Mts, Kigongoi	1997 0101
8	*Saintpaulia* cf. *ionantha* H. Wendl.	Tanzania: Tanga, Sigi Caves,	1971 0860
9	*Saintpaulia magungensis* E. Roberts	Tanzania: E. Usambara Mts, Magunga, Mt Mlinga	1992 3187
10	*Saintpaulia magungensis* var. *minima* B. L.Burtt	Tanzania: E. Usambara Mts, Mavoera estate, Amani,	1959 4352
11	*Saintpaulia nitida* B. L.Burtt	Tanzania: Nguru Mts, Mkobwe, Turiani, *ca*. 1000 m	1992 3186
12	*Saintpaulia orbicularis* var. *purpurea* B. L.Burtt	Tanzania: W. Usambara Mts, Ambangulu, 1060–1200 m	1958 3586
13	*Saintpaulia pendula* var. *kizarae* B. L.Burtt	Tanzania: NE Usambara Mts, Mt Mtai, Kizara	1997 0103
14	*Saintpaulia rupicola* B. L.Burtt	Kenya: Kaloleni	1997 0094
15	*Saintpaulia shumensis* B. L.Burtt	Tanzania: W. Usambara Mts, Shume, 1900–1950 m	1996 2088
16	*Saintpaulia* Sigi Falls	Tanzania: Tanga, Sigi River	1992 3183
17	*Saintpaulia teitensis* B. L.Burtt	Kenya: Teita Hills, Mbololo Hill	C 3771[b]
18	*Saintpaulia tongwensis* B. L.Burtt	Tanzania: E. Usambara Mts, Tongwe Mts, 600 m	1985 0668
19	*Saintpaulia velutina* B. L.Burtt	Tanzania: W. Usambara Mts, Balangai, 900 m	1987 2179
20	*Streptocarpus caulescens* Vatke	Tanzania: W. Usambara Mts, Uluguru Mts; Kenya: Teita Hills, 1800 m	1971 1199
21	*Streptocarpus saxorum* Engl.	Tanzania: Usambara, Nguru, and Uluguru Mts, Kenya: Teita Hills, 900 m	1972 1499

[a]These numbers were also used as voucher numbers.
[b]Herbarium specimen no.

(*c*) *Sequence and phylogenetic analysis*

Sequence boundaries of both ITSs of all taxa were determined by comparison with published rDNA sequence data for *Daucus carota* and *Vicia faba* (Yokota *et al*. 1989). Both ITS regions were aligned using the CLUSTAL option in the multiple alignment program Sequence Navigator™ v. 1.0.1 software package (Perkin Elmer, Applied Biosystems Division, Foster City, CA, USA), with manual adjustments. The G+C content was determined by inspection, and transition-transversion ratios using MacClade v. 3.01 (Maddison & Maddison 1992). Sequence divergence among taxa was calculated using the DISTANCE MATRIX option in PAUP v. 3.1.1 (Swofford 1993). All sequences used in this study are available from the authors on request.

Phylogenetic analysis by parsimony and statistical analysis (bootstrap) (*BS*; Felsenstein 1985), decay index (*DI*; Bremer 1988; Donoghue *et al*. 1992) and descriptive statistics (consistency index (*CI*); Kluge & Farris 1969), retention index (*RI*; Farris 1989), rescaled consistency index (*RC*; Swofford 1993)) were performed as described previously (Möller & Cronk 1997), except for the *DI* which was only calculated for trees up to two steps longer than the shortest tree, due to the limit of trees (32 767) PAUP is capable of storing. Only combined ITS 1 and ITS 2 sequence data were subjected to phylogenetic analyses. For all analyses of sequence data, gaps (indels) were treated as missing data (Soltis & Kuzoff 1995; Susanna *et al*. 1995; Downie & Katz-Downie 1996). Indels were scored as separate presence–absence characters. An analysis with and without gap matrix was carried out. Character state changes were weighted equally, except for one analysis where transitions over transversions were weighted by a factor of 1.3 : 1, as described previously (Möller & Cronk 1997).

3. RESULTS

(*a*) *ITS sequence matrix*

Alignment of ITS sequences of the 21 taxa analysed resulted in a 485 bp-long data matrix (given in the electronic appendix to this paper, which can be accessed at `http://www.pubs.royalsoc.ac.uk/publish/pro_bs/dec97pb.htm`); their characteristics (including G+C content) are given in table 3. Several taxa had polymorphic rDNA copies with one base duplication, detected by double peaks in the sequence data file (e.g. *S. pendula* nucleotide T at position 387, *S. magungensis* base C at position 65, and *S. magungensis* var. *minima* G at position 60, and both

Table 3. *Sequence characteristic of ITS 1 and ITS 2 regions of 21 taxa of* Streptocarpus *and* Saintpaulia

parameter	ITS 1	ITS 2	ITS 1 and ITS 2
length range (total), bp	239–243	225–233	466–473
length mean (total), bp	241.0	226.5	467.5
length range (ingroup), bp	241–243	225–228	466–471
length mean (ingroup), bp	241.2	225.8	467.0
length range (outgroup), bp	239–240	233	472–473
length mean (outgroup), bp	239.5	233	472.5
aligned length, bp	248	237	485
G+C content range (%)	50.2–55.3	54.4–57.1	52.2–55.3
G+C content mean (%)	51.5	55.2	53.3
sequence divergence (ingroup), %	0–17.6	0–13.9	0–15.8
sequence divergence (in/outgroup), %	9.2–15.9	10.6–12.9	9.9–14.5
number of indels	12	10	22
size of indels, bp	1–3	1–4	1–4
number of constant sites (%)	181 (73.0)	178 (75.1)	359 (74.0)
number of variable sites (%)	67 (27.0)	59 (24.9)	126 (26.0)
number of autapomorphic sites (%)	40 (16.1)	37 (15.6)	77 (15.9)
number of informative sites (%)	27 (10.9)	22 (9.3)	49 (10.1)
transitions (minimum)	44	37	81
transversions (minimum)	37	30	67
transitions/transversions	1.23	1.27	1.24
average number of steps per character	0.343	0.295	0.320

S. magungensis accessions base T at position 387, respectively). In each case inclusion of all sequence types resulted in neighbouring position of the respective sequence variants. Therefore only one type was included in the present analysis.

The length of ITS 1 and ITS 2 was, on average, 241.0 and 226.5 bp. Alignment of all taxa required the insertion of 22 gaps of 1–4 bp in length, 12 in ITS 1 and 10 in ITS 2. Out of those ten were potentially informative. The lengths of aligned ITS 1 and ITS 2 regions were 248 bp and 237 bp, respectively. Of those 485 unambiguously aligned sites, 359 (74.0%) were constant, 49 (10.1%) were potentially informative phylogenetically, and 77 (15.9%) were autapomorphies, unique to individual taxa (table 3).

Within the *Saintpaulia* accessions, sequence divergence of ITS 1 ranged from 0–17.6%, and from 9.2–15.9% between ingroup taxa and outgroup taxa. ITS 2 was less variable with 0–13.9% divergence observed between ingroup taxa, and 10.6–12.9% sequence divergence in pairwise comparisons between ingroup taxa and outgroup taxa. Pairwise comparisons of individual taxa across both spacer regions revealed 0–15.8% sequence divergence within the ingroup, and 9.9–14.5% divergence between ingroup and outgroup taxa analysed (table 3). The maximum sequence variation between *Saintpaulia* accessions was 15.8% (73 character changes) between *S. goetzeana* and *S. nitida*. Sequences of *S. grandifolia*, *S. grotei*, *S. magungensis*, *S. magungensis* var. *minima*, *S.* Sigi Falls, *S. tongwensis* and *S. velutina* were identical.

(*b*) *Phylogenetic analysis*

Parsimony analysis of aligned ITS sequences yielded four most parsimonious trees of 182 steps when all uninformative characters were included, and 92 steps with autapomorphies excluded, with *CI* values of 0.890 and 0.783, respectively. These values are considerably higher than the theoretical value of 0.460 calculated for 20 taxa from the study by Sanderson & Donoghue (1989). The *RI* was 0.843, and thus the *RC* was 0.750 with, and 0.659 without, uninformative characters.

The average number of nucleotide substitutions per character was low, with 0.32 indicating a low saturation of base substitution. The homoplasy index (*HI*) of the present data matrix was low (*HI*=0.110). The bootstrap values for individual clades ranged from 56 to 100% (figure 2).

Eleven character changes separated the outgroup taxa *Streptocarpus caulescens* and *Streptocarpus saxorum* from the *Saintpaulia* taxa (*BS*=78%, *DI*=2). *S. goetzeana* is separated from the rest of the *Saintpaulia* species by 23 character changes (*BS*=100%, *DI*>2), and *S. teitensis* by 17 changes (*BS*=100%, *DI*=2) (figure 3). The other *Saintpaulia* taxa formed two main groups, with *S. intermedia*, *S. pendula* and *S. rupicola* with unresolved relations, with a separate clade containing *S. brevipilosa* and *S. nitida* (*BS*=95%, *DI*>2) on one side, and the rest of the *Saintpaulia* species in a separate clade, the *ionantha* complex (*BS*=59%; *DI*=1), with *S. shumensis* (*BS*=56%, *DI*=1) separating them. *S. difficilis* and *S. orbicularis* var. *purpurea* were separated from the rest of the *ionantha* complex by two character changes (*BS*=82%, *DI*=2). Of the ten potentially informative indels, six were congruent with the tree topology of the strict consensus tree.

Exclusion of the gap matrix from the combined ITS 1 and ITS 2 data matrix resulted in four most parsimonious trees of 155 steps (77 steps excluding uninformative characters; *CI*=0.903; *RI*=0.853; *RC*=0.771). The strict consensus tree differed from the strict consensus tree obtained with the addition of a

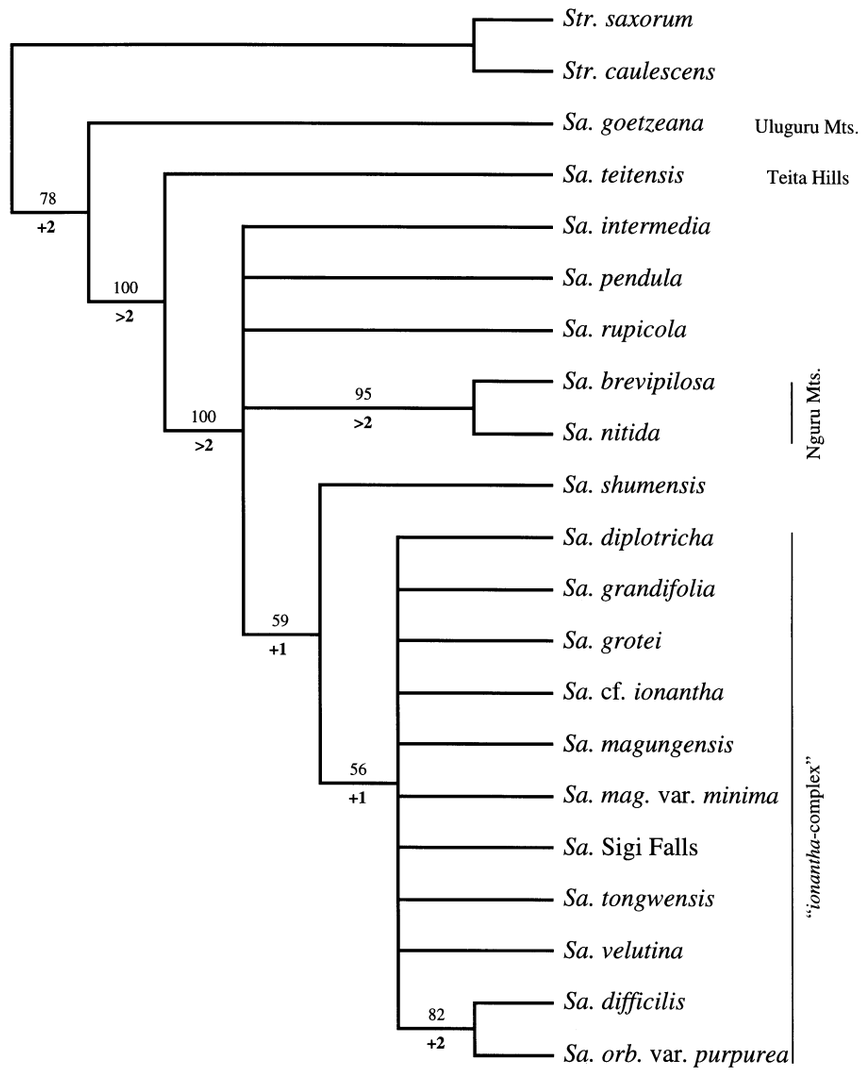

Figure 2. Strict consensus tree based on the four most parsimonious trees for 19 *Saintpaulia* and two *Streptocarpus* taxa of 182 steps in length based on parsimony analysis of the combined ITS 1 and ITS 2 sequence data plus the alignment gap matrix. Upper numbers are bootstrap values of 1000 replicates. Lower numbers (in bold type) are decay indices (the number of steps necessary to cause collapse of monophyletic groups). (CI=0.890; RI=0.843; RC=0.750.)

gap matrix in the collapse of the branch separating *S. shumensis* from the *ionantha* complex, and the collapse of the branch separating *S. difficilis* and *S. orbicularis* var. *purpurea* in the *ionantha* complex.

The transition–transversion ratio was 1.23 for ITS 1 and 1.27 for ITS 2, and 1.24 for the combined data matrix. Altering the character weights to 1.3 : 1 to accommodate the transition ratio and reanalysing the data (gap matrix excluded) in a weighted parsimony analysis gave four most parsimonious trees. The resulting strict consensus tree had an identical topology to the strict consensus tree (gap matrix excluded) for the unweighted parsimony analysis.

4. DISCUSSION

(*a*) *Molecular evolution of ITS in* Saintpaulia*, and its limitations for phylogenetic reconstruction*

The ITS of ribosomal DNA of the *Saintpaulia* species investigated evolved both by base substitutions and by insertion–deletion events (22 alignment gaps). Maximum sequence divergence within *Saintpaulia* was 17.6% for ITS 1 and 13.9% for ITS 2. The overall levels of sequence variation in *Saintpaulia* are similar to infrageneric levels found in other angiosperms. In genera of Asteraceae subtribe Madiinae, for instance, sequence divergence ranged from 0.4–19.2% in ITS 1 and 0–12.9% in ITS 2 (Baldwin 1992). However, *Saintpaulia* is unusual in that a large group of species (11 accessions: the *ionantha* complex) had sequence divergence too low for clear phylogenetic resolution (figures 2 and 3). Seven accessions had an identical sequence. The single resolved node in this group (*S. difficilis* and *S. orbicularis* var. *purpurea* as sister taxa) is not very well supported (although a *DI* of 2 was calculated) as the only differences shared by these species were at indel positions, and the branch supporting the two species collapsed when the gap matrix was excluded.

Unresolved groups in ITS phylogenies are known from other studies and are attributed to rapid radiation,

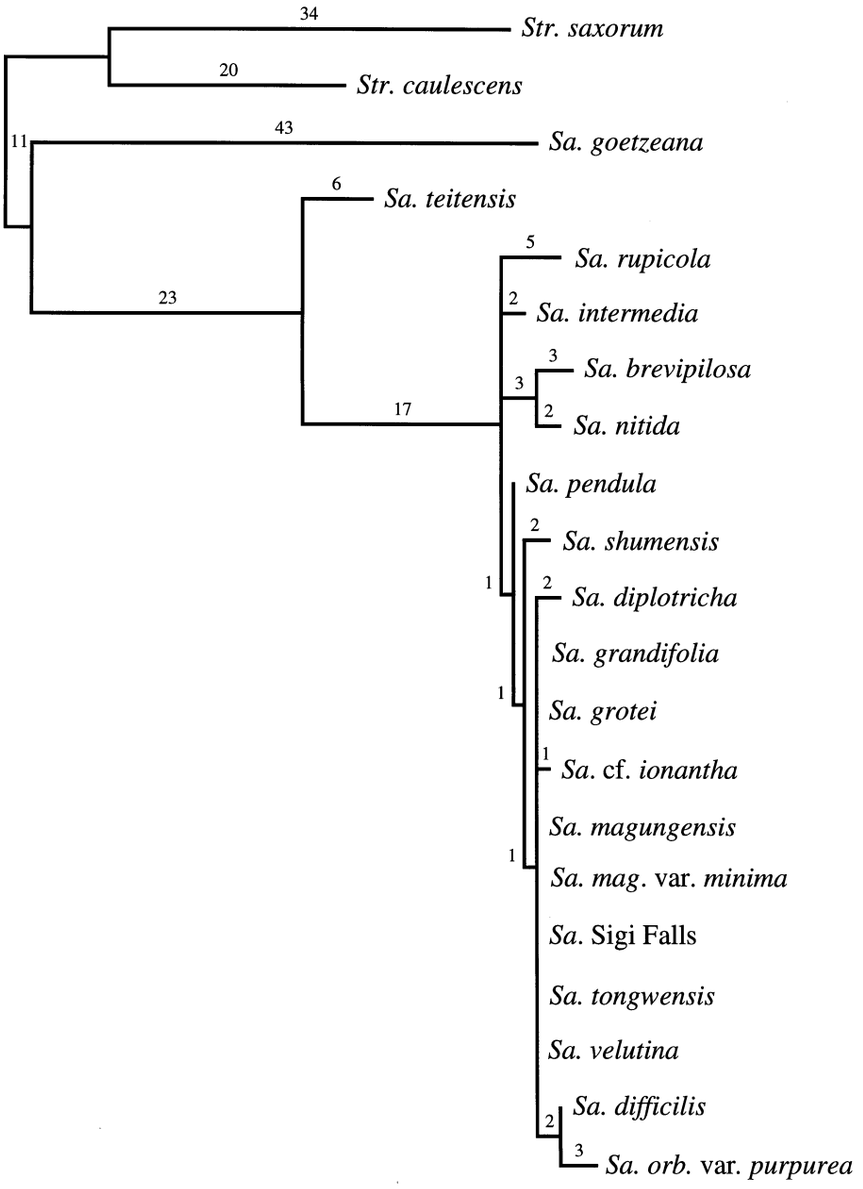

Figure 3. Phylogram of one of four most parsimonious trees for 19 *Saintpaulia* and two *Streptocarpus* taxa of 182 steps in length based on parsimony analysis of the combined ITS 1 and ITS 2 sequence data plus the alignment gap matrix. Numbers along branches indicate the number of character changes shared amongst taxa (branch length), including autapomorphic changes.

typically on islands. There are many reports of unresolved polytomies at the base of island clades, implying rapid radiation after colonization (Sang *et al*. 1994; Kim *et al*. 1996). The Eastern Arc Mts are habitat archipelagoes, and the fact that speciation has outstripped ITS differentiation may result from processes similar to those of other island radiations.

Although the spacers are thought to be important in post-transcriptional processing, and thus conserved to some extent (Liu & Schardl 1994; Van der Sande *et al*. 1992; Van Neus *et al*.1994), ITS sequences are known to vary very extensively at higher levels of the taxonomic hierarchy. It is therefore surprising that speciation (as in the *ionantha* complex) has been able to outstrip variation in the fast-evolving ITS region. However, ITS conservation at very low taxonomic levels (such as within species complexes) can be explained by the fact that nuclear rDNA is multi-copy and new variants have to spread within the genome by unequal crossing-over (Copenhaver *et al*. 1995), a process that may take many generations before fixation. This process of concerted evolution and gene homogenization is even slower between different chromosomal loci (Schlotterer & Tautz 1994). The rDNA copy number and the number of loci are not known for *Saintpaulia*, but in *Arabidopsis thaliana* there are approximately 1500 rRNA genes arrayed in tandem at two nucleolus organizer regions (Copenhaver & Pickaard 1996).

We therefore suggest that at divergence times that are small compared to rDNA homogenization rates, ITS variation will appear highly conservative. However, where divergence times are long compared to rDNA

homogenization rates, ITS variation will appear disproportionately extensive. This uneven predicted behaviour of ITS sequence variation may explain the fact that while there is virtually no ITS sequence variation between species of the *ionantha* complex, *S. goetzeana* is separated from the other species by a very long branch.

(*b*) *Biogeography and dispersal*

The genus *Saintpaulia* contains a number of geographically disjunct groups. We have shown that there is a strong correlation between the biogeography and the major clades of *Saintpaulia*. It is clear that the isolation of these populations has been a major factor in the differentiation of groups. The disjunct distribution of the genus may have two possible sources: either contraction from a formerly more widespread distribution (with spread and contraction presumably driven by climatic change (Coetzee 1978)) or long-distance dispersal (White 1983). With regard to the expansion and contraction of range, spread along river systems may be of significance, with the microclimate of riverine vegetation providing refuge from climatic vicissitudes.

All that is currently known about dispersal in *Saintpaulia* suggests that long-distance dispersal is highly unlikely. The capsule becomes dry and dehiscent and is unlikely to be consumed by birds for endozoochorous transport. The seeds, although small and lacking in endosperm, are not dust-like as in orchids, and are unlikely to blow far by wind. Furthermore, Saintpaulias are understorey herbs of places sheltered from wind. Coupled to this, there are several species of *Streptocarpus* with the same or similar distributions as the genus *Saintpaulia* (table 1). *Streptocarpus* also have no mechanisms for long-distance dispersal. It is highly unlikely that the same pattern of long-distance dispersal could be responsible for the distribution of a whole range of plants unadapted for dispersal. It is therefore likely that the distribution of *Saintpaulia* has been different in the past. Although the present *Saintpaulia* sites are now separated by areas of inhospitable dry woodland and cultivations, there is evidence that in the Tertiary the climate of the region was much wetter and that moist forest was very widespread until progressive aridification during the Pliocene. Moist forest was also more widespread during pluvial periods in the Pleistocene (Livingstone 1982). While the disjunct distribution of *Streptocarpus* and *Saintpaulia* may be Tertiary in origin, the apparently recent radiation of the *ionantha* complex in the Usambara Mts may be the result of periodic spread and isolation of populations related to climatic change in the Pleistocene.

(*c*) *Origin and evolution*

In a previous paper (Möller & Cronk 1997) we showed that *Saintpaulia* originated from within the *Streptocarpus* lineage, and is sister group to those East African members of *Streptocarpus* subgenus *Streptocarpella* sequenced so far. If *Saintpaulia* originated in one of the present areas of distribution, we believe that the basal position of the Uluguru Mountain species relative to the Teita Hills and Usambara Mountain species is consistent with an Uluguru origin for *Saintpaulia* followed by a subsequent northern migration of a lowland form as far as the Teita Hills and the Usambara Mts. The Uluguru Mts are a centre of distribution for *Streptocarpus* subg. *Streptocarpella* in East Africa (table 1), and in this connection it is interesting that *S. goetzeana* shows considerable similarity in many (apparently plesiomorphic) ITS sequence features with *Streptocarpus caulescens* (e.g. gap 12, 15, 19—see the electronic appendix).

The sequence divergence between *S. goetzeana* and the rest of the genus is a striking feature of this analysis, but there are no particularly remarkable morphological differences. However, both *S. pusilla* and *S. goetzeana* have somewhat campanulate flowers reminiscent of *Streptocarpus* which may be a plesiomorphic feature. Furthermore, *S. goetzeana* has the long internodes and opposite leaves characteristic of caulescent *Streptocarpus*. It also has striking bicolorous flowers, with the two adaxial corolla lobes being of dark purple, while the rest of the corolla is pale.

The Ulugurus are high mountains and Möller & Cronk (1997) and Cronk & Möller (1997) suggest that loss of specialized pollinators with elevation on mountains may provide the selective pressure for pollinator switching. Pollinator switching has occurred in the evolution of *Saintpaulia* from *Streptocarpus* (Möller & Cronk 1997; Cronk & Möller 1997). Against this, however, is the fact that the Ulugurus have an extremely rich biota that is compressed and stratified on a small mountain block, and until survey work is carried out it cannot be assumed that there is a deficiency of particular insect groups at altitude (R. Polhill, personal communication). Another possible explanation for the basal position of *S. goetzeana* is that the Uluguru Mountain population became isolated by aridity some time prior to the isolation of the Teita Hills and Usambara populations (a vicariance explanation).

The *ionantha* complex from the Usambara Mts is an important group of species that shows little gene sequence differentiation and is taxonomically complex (Burtt 1958, 1964). In this complex, the ITS region has completely failed to resolve the relationships among the ten taxa. Further work is needed to find different genetic techniques to differentiate among this group. Nine taxa (of the *ionantha*-complex) have been included in this study. A tenth, *S. confusa*, has not been available to us, but as it is very close morphologically to *S. ionantha* and other species in the complex, we believe it belongs here. The only other species that have been omitted from this study are *S. inconspicua* and *S. pusilla*, both from the Uluguru Mts. These do not appear to be in cultivation, but would be extremely interesting to obtain for study. We think it is likely that both these species belong to the *S. goetzeana* 'Uluguru Mountains clade'. *Saintpaulia pusilla* has been confused with *S. goetzeana* in the past, and it has similar bicolorous flowers (Burtt 1958). *S. inconspicua* is morphologically interesting and more isolated. If it proves to be related to *S. goetzeana* by a deep branch, this would be powerful evidence for the long presence of *Saintpaulia* in the

Ulugurus and the possible origin of *Saintpaulia* there. What is clear, however, is that the centre of species diversity (East Usambara Mts) is not strongly implicated as a possible centre of origin. Rather, the close relationship of all these species is evidence for no more than a short residency of *Saintpaulia* in the Usambaras. Similarly, the morphologically divergent species in the Uluguru Mts may indicate a long residency there.

(*d*) *Conservation in relation to phylogeny*

The phylogenetically basal position, and comparatively long branch lengths, of *S. goetzeana* and *S. teitensis* make these species particularly important for conservation of the full range of genetic diversity of the genus (Faith 1994; Moritz 1995). The *S. goetzeana* lineage, being divergent, contributes the most unique DNA characters, at least as indicated by ITS variation (see Humphries *et al*. (1995) for a review of the concepts). It is likely that this pattern of variation is carried over into other regions of the genome.

The Uluguru Mts are comparatively undisturbed at high elevations (although there is severe population pressure lower down (Lovett 1990)), and the *Saintpaulia* species here all occur over 1300 m in altitude (Burtt 1958). Johansson (1978) reports that *S. goetzeana* is common and widespread in the Uluguru Mts, *S. pusilla* is rarer but it is also reported from one locality in the Ukaguru Mts. However, the forests of the Teita Hills are fragmented, and the region is densely populated (Davis *et al*. 1994), which must give rise to concern over conservation as *Saintpaulia teitensis* is restricted to one population in the Mbololo forest, Teita Hills (Eastwood & Maunder 1995).

On the other hand, it appears that species in the *ionantha* complex may be in a state of active evolutionary change. This is of particular interest, and it therefore makes them candidates for special-purpose conservation. Indeed, Linder (1995) goes as far as to propose an 'evolutionary fate' criterion in conservation. Phylogenetic information does not therefore entirely solve the 'agony of choice' involved in setting conservation priorities (Vane-Wright *et al*. 1991), but *Saintpaulia* does provide a good example of how phylogeny gives a further dimension to conservation planning.

The authors thank B. L. Burtt for confirming the identities of the accessions; B. L. Burtt, C. Lindqvist, S. Simuyu and Professor Bill Hill FRS for helpful discussions; A. Robertson, R. Polhill and J. Lovett for advice, information and comment; H. Sluiman, J. Preston, N. Preston and C. Guihal for technical assistance; and U. Gregory, D. Mitchell and J. Main for the maintenance of the Gesneriaceae collection. We thank the Regius Keeper and staff at the Royal Botanic Garden, Edinburgh (RBGE), for use of research facilities and the Institute of Cell and Molecular Biology, University of Edinburgh, for access to sequencing facilities. Special acknowledgement and thanks goes to the many collectors and growers who have contributed to the RBGE *Saintpaulia* collection: in particular the late I. C. (Sylva) Mather, H.-E. Wichert, Herrenhäuser Gärten, Hannover, Germany, and Maryjane Evans, American Gloxinia and Gesneriad Society, Inc., Randolph, NJ, USA. The receipt of a Leverhulme Trust Award, no. F/771/B, is gratefully acknowledged.

REFERENCES

Arisumi, T. 1964 Interspecific hybridization in African Violets. *J. Hered.* **55**, 181–182.

Baatvik, S. T. 1993 The genus *Saintpaulia* (Gesneriaceae) 100 years: history, taxonomy, ecology, distribution and conservation. *Fragm. Flor. Geobot. Suppl.* **2**(1), 97–112.

Baldwin, B. G. 1992 Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molec. Phylogenet. Evol.* **1**, 3–16.

Beentje, H. J. 1990 The forests of Kenya. *Mitt. Inst. Allg. Bot. Hamburg* A **23**, 265–286.

Bjorndalen, J. E. 1992 Tanzania's vanishing rain-forests—assessment of nature conservation values, biodiversity and importance for water catchment. *Agr. Ecosyst. Environ.* **40**, 313–334.

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

Burgess, N. D., Mwasumbi, L. B., Hawthorne, W. J., Dickinson, A. J. & Doggett, R. A. 1992 Preliminary assessment of the distribution, status and biological importance of coastal forests in Tanzania. *Biol. Conserv.* **62**, 205–218.

Burtt, B. L. 1958 Studies on the Gesneriaceae of the Old World XV: the Genus *Saintpaulia*. *Notes R. Bot. Gdn Edinb.* **22**, 547–568.

Burtt, B. L. 1964 Studies on the Gesneriaceae of the Old World XXV: additional notes on *Saintpaulia*. *Notes R. Bot. Gdn Edinb.* **25**, 191–195.

Clarke, G. P. 1998 Plants in peril: notes on lowland African Violets (*Saintpaulia*) in the wild. *Kew Mag.* (In the press.)

Coetzee, J. A. 1978 Phytogeographical aspects of the montane forests of the chain of mountains on the eastern side of Africa. In *Proceedings of the Symposium of the International Geographical Union Commission on High Altitude Geo-Ecology, November 21–23, 1974, Mainz* (ed. C. Troll & W. Lauer), pp. 482–494. Wiesbaden: Franz Steiner Verlag GmbH.

Copenhaver, G. P., Doelling, J. H., Gens, J. S. & Pickaard, C. S. 1995 Use of RFLPs larger than 100 kbp to map the position and internal organisation of the nucleolus organizer region on chromosome 2 in *Arabidopsis thaliana*. *Plant J.* **7**, 273–286.

Copenhaver, G. P. & Pickaard, C. S. 1996 Two dimensional RFLP analyses reveal megabase-sized clusters of ribosomal RNA gene variants in *Arabidopsis thaliana* suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. *Plant J.* **9**, 273–282.

Cronk, Q. C. B. & Möller, M. 1997 Genetics of floral symmetry revealed. *Trends Ecol. Evol.* **12**, 85–86.

Davis, S. D., Heywood, V. H. & Hamilton, A. C. (eds) 1994 *Centres of plant diversity. A guide and strategy for their conservation*. Oxford: Information Press.

Donoghue, M. J., Olmstead, R. G., Smith, J. F. & Palmer, J. D. 1992 Phylogenetic relationships of Dipsacales based on *rbc*L sequences. *Ann. Mo. Bot. Gdn* **79**, 333–345.

Downie, S. R. & Katz-Downie, D. S. 1996 A molecular phylogeny of Apiaceae subfamily Apioideae: evidence from nuclear ribosomal DNA internal spacer sequences. *Am. J. Bot.* **83**, 234–251.

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

Eastwood, A. & Maunder, M. 1995 *A conservation assessment of* Saintpaulia *taxa cultivated within the living collections Royal Botanic Gardens Kew.* Species Assessment Report 2. Royal Botanic Gardens, Kew, UK: Conservation Projects Development Unit.

Faith, D. P. 1994 Phylogenetic pattern and the quantification of organismal biodiversity. *Phil. Trans. R. Soc. Lond.* B **345**, 45–58.

Farris, J. S. 1989 The retention index and homoplasy excess. *Syst. Zool.* **38**, 406–407.

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

Guralnick, L. J., Ting, I. P. & Lord, E. M. 1986 Crassulacean acid metabolism in the Gesneriaceae. *Am. J. Bot.* **73**, 336–345.

Hilliard, O. M. & Burtt, B. L. 1971 *Streptocarpus*: an African plant study. Pietermaritzburg: University of Natal Press.

Humphries, C., Williams, P. H. & Vane-Wright, R. I. 1995 Measuring biodiversity value for conservation. *A. Rev. Ecol. Syst.* **26**, 93–111.

Iversen, S. T. 1991 The Usambara Mountains, NE Tanzania: phytogeography of the vascular plant flora. *Symb. Bot. Upsalienses* **14**, 45–62.

Johansson, D. R. 1978 *Saintpaulias* in their natural environment with notes on their present status in Tanzania and Kenya. *Biol. Conserv.* **14**, 45–62.

Kim, S. C., Crawford, D. J., Francisco Ortega, J. & Santos Guerra, A. 1996 A common origin for woody *Sonchus* and five related genera in the Macaronesian islands—molecular evidence for extensive radiation. *Proc. Natn. Acad. Sci. USA* **93**, 7743–7748.

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

Linder, H. P. 1995 Setting conservation priorities—the importance of endemism and phylogeny in the southern African orchid genus *Herschelia*. *Conserv. Biol.* **9**, 585–595.

Liu, J.-S. & Schardl, C. L. 1994 A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Molec. Biol.* **26**, 775–778.

Livingstone, D. A. 1982 Quarternary geography of Africa and the refuge theory. In *Biological diversification in the tropics* (ed. G. T. Prance), pp. 523–536, New York: Columbia University Press.

Lovett, J. C. 1990 Classification and status of the moist forests of Tanzania. *Mitt. Inst. Allg. Bot. Hamburg* A **23**, 287–300.

Lovett, J. C. 1996 Elevational and latitudinal changes in tree associations and diversity in the Eastern Arc Mountains of Tanzania. *J. Trop. Ecol.* **12**, 629–650.

Lovett, J. C. & Friis, I. 1996 Patterns of endemism in the woody flora of north-east and east Africa. In *The biodiversity of African plants* (ed. L. J. G. Van der Maesen), pp. 582–601. Dordrecht, The Netherlands: Kluwer.

Maddison, W. P. & Maddison, D. R. 1992 *MacClade, v. 3.01.* Sunderland, MA: Sinauer.

Möller, M. & Cronk, Q. C. B. 1997 Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. *Am. J. Bot.* **84**, 956–965.

Moritz, C. 1995 Uses of molecular phylogenies for conservation. *Phil. Trans. R. Soc. Lond.* B **349**, 113–118.

Pócs, T. 1974 Vegetation mapping in the Uluguru Mountains (Tanzania, East Africa). *Boissiera* **24**, 477–498.

Sanderson, M. J. & Donoghue, M. J. 1989 Patterns of variation in levels of homoplasy. *Evolution* **43**, 1781–1795.

Sang, T., Crawford, D. J., Kim, S. C. & Stuessy, T. F. 1994 Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands—evidence from sequences of the ITS regions of nuclear ribosomal DNA. *Am. J. Bot.* **81**, 1491–1501.

Schlotterer, C. & Tautz, D. 1994 Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Curr. Biol.* **4**, 777–783.

Soltis, D. E. & Kuzoff, R. K. 1995 Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* **49**, 727–742.

Susanna, A., Jacas, N. G., Soltis, D. E. & Soltis, P. S. 1995 Phylogenetic relationships in tribe Cardueae (Asteraceae) based on ITS sequences. *Am. J. Bot.* **82**, 1056–1068.

Swofford, D. L. 1993 *PAUP: phylogenetic analysis using parsimony, v. 3.1.1.* Champaign, IL: Illinois Natural History Survey.

Van Der Sande, C. A. F. M., Kwa, M., Van Neus, R. W., Van Heerikhuizen, H., Raué, H. A. & Planta, R. J. 1992 Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. *J. Molec. Biol.* **223**, 899–910.

Vane-Wright, R. I., Humphries, C. J. & Williams. P. H. 1991 What to protect? Systematics and the agony of choice. *Biol. Conserv.* **55**, 235–254.

Van Neus, R. W., Rientjes, J. M. J., Van Der Sande, C. A. F. M., Zerp, S. F., Sluiter, C., Venema, J., Planta, R. J., & Raué, H. A. 1994 Separate structural elements within internal transcribed spacer 1 of *Saccharomyces cerevisiae* precursor ribosomal RNA direct the formation of 17S and 26S ribosomal RNA. *Nucl. Acids Res.* **22**(6), 912–919.

White, F. 1983 Long-distance dispersal and the origins of the Afromontane flora. *Sonderb. Naturwiss. Vereins Hamburg* **7**, 87–116.

Yokota, Y., Kawata, T., Iida, Y., Kato, A. & Tanifuji, S. 1989 Nucleotide sequences of the 5.8S rDNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. *J. Molec. Evol.* **29**, 294–301.

Received 8 August 1997; accepted 5 September 1997

As this paper exceeds the maximum length normally considered for publication in *Proceedings B*, the authors have agreed to make a contribution towards production costs.