Annals of Botany **91**: 31–38, 2003
doi:10.1093/aob/mcg005, available online at www.aob.oupjournals.org

First Nuclear DNA C-values for 28 Angiosperm Genera

LYNDA HANSON, REBECCA L. BROWN, AMY BOYD, MARGARET A. T. JOHNSON and MICHAEL D. BENNETT*

Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

Received: 22 August 2002 Returned for revision: 19 September 2002 Accepted: 25 September 2002 Published electronically: 31 October 2002

This paper reports first DNA C-values for 28 angiosperm genera. These include first DNA C-values for 25 families, of which 16 are monocots. Overall familial representation is 47·2 % for angiosperms, but is now much higher for monocots (75 %) and basal angiosperms (73·1 %) than for eudicots (38·7 %). Chromosome counts are reported for 22 taxa, including first records for six genera plus seven species. Unrepresented families will become increasingly enriched for monotypic taxa from obscure locations that are harder to access. Thus, completing familial representation for genome size for angiosperms may prove impossible in any short period, and progress towards this goal will become slower.

Key words: Nuclear DNA amounts, DNA C-values, angiosperm families, chromosome numbers, monocots, genome size.

INTRODUCTION

Within angiosperms, DNA C-values (corresponding to the DNA amount in an unreplicated gametic nucleus) range about 1000-fold from approx. 0·1 to 127·4 pg (Bennett *et al*., 2000). In recent years there has been increased interest in the causes and consequences of this huge range, with new research providing intriguing insights into the possible mechanisms that generate it (Vicient *et al*., 1999; Kirik *et al*., 2000; Shirasu *et al*., 2000; Bensasson *et al*., 2001; Petrov, 2001). However, available C-value data for angiosperms (approx. 3500 species) still correspond to only approx. 1·4 % of an estimated 250 000 species.

The need to identify major gaps in C-value data and to recommend targets for new work to fill them by international collaboration was confirmed at the Angiosperm Genome Size Workshop, held at the Royal Botanic Gardens, Kew (RBG, Kew) in 1997. C-values were then still unavailable for 68 % of angiosperm families recognized by the Angiosperm Phylogeny Group (APG) classification (APG, 1998). Consequently, a goal of complete familial representation by 2002 was agreed. However, Bennett *et al*. (2000) noted that in a sixth supplementary list of C-values for 691 species published or communicated since 1997, only 12 were also first estimates for families. As progress towards the goal of completing familial coverage was disappointing, new work to correct this was begun at RBG, Kew in 1999. Thus in 2001, Hanson *et al*. (2001*a*, *b*) reported first DNA C-values for 50 angiosperm families determined in the new project, and another targeted study recently added first C-values for a further five families in the basal angiosperms (Leitch and Hanson, 2002). Nevertheless, massive gaps still exist in our knowledge of C-values, and over 50 % of angiosperm families still have no reported C-value. The very size of this problem can be daunting and can act as a barrier to progress. In such cases it is useful to build on strengths and break the task into manageable parts. This approach was adopted when RBG, Kew recently elected to make monocots (which constitute 20 % of angiosperms) the key focus for work to complete a checklist for such species by 2007, and to link monocot species' names to the wide range of information available in one seamless database. As part of this strategy, the general target of achieving full familial representation for angiosperms, set at the 1997 workshop, now has a particular goal of achieving this for monocots as our key focus. The prime aim of the present work was therefore to obtain first DNA C-values for a further 25 families but with a particular emphasis on monocots.

* For correspondence. Fax +44 (0)20 8332 5320, e-mail m.bennett@rbgkew.org.uk

MATERIALS AND METHODS

Plant material

Table 1 lists the 28 perennial species studied in the present work and gives their origin, source and reference data. Following the familial circumscriptions of the Angiosperm Phylogeny Group (APG II, 2003), 25 of the species are from families (Table 2) for which no DNA C-values were previously included in the Angiosperm DNA C-values database (Bennett and Leitch, 2001) or in Hanson *et al*. (2001*a*, *b*). The remaining three (*Aphyllanthes monspeliensis*, *Triteleia laxa* and *Xanthorrhoea preisii*) were also originally selected as representing families for which no published DNA C-value was previously known, namely Aphyllanthaceae, Themidaceae and Xanthorrhoeaceae (APG, 1998). However, a new circumscription of families by the APG (APG II, 2003) has meant that Themidaceae and Aphyllanthaceae, which were previously recognized as separate families, are sunk together with five others into a newly circumscribed Asparagaceae. The circumscription of

TABLE 1. *Geographical region of origin, source of experimental material, RBG Kew identity number (ID no.), cytology number (Cyt. no.) and identification status for the 28 species studied in the present work*

Entry no.	Taxon	Origin	Source of material[‡]	ID no.	Cyt. no.	Identification status[†]
Monocots						
Alismatales						
1	*Zostera marina* L.	Europe, Australia, New Guinea	RBG,K	–	01-136	b
Asparagales						
2	*Aphyllanthes monspeliensis* L.	Portugal to Italy, N. Africa	RBG,K	1990-2610	01-145	a
3	*Triteleia laxa* Benth.	West and North America	RBG,K	1960-67034	01-146	a
4	*Astelia fragrans* Colenso	Mascarenes, New Guinea, Australia, New Zealand, Polynesia to Hawaii and Chile	RBG,K	1989-2690	01-176	c
5	*Blandfordia punicea* Sweet.	E. Australia	RBG,K	2000-3965	01-121	c
6	*Doryanthes palmeri* W. Hill ex Benth.	E. Australia	RBG,K	1948-60704	01-175	a
7	*Ixiolirion ledebourii* Fisch. & Mey.	West and central Asia	RBG,K	1989-3153	01-126	c
8	*Odontostomum hartwegii* Torr.	California	RBG,K	1987-8225	01-135	d
9	*Xanthorrhoea preisii* Endl.	Australia	RBG,K	1990-2825	01-127	c
10	*Xeronema callistemon* W. R. B. Oliv.	New Caledonia and north New Zealand	RBG,K	1974-1783	01-150	c
Dioscoreales						
11	*Narthecium ossifragum* Huds.	Northern temperate Europe	JSB	2001-4165	01-161	b*
Liliales						
12	*Lapageria rosea* Ruiz & Pav.	Chile and Argentina	RBG,K	–	01-169	c
13	*Ripogonum papuanum* C. T. White	New Guinea, Australia, New Zealand	RBG,K	1987-8058	01-216	c
Commelinoids						
14	*Dasypogon hookeri* Drumm.	SW Australia	RBG,K	–	01-134	c
15	*Hanguana malayana* Merrill	Sri Lanka, SE Asia and Malaysia	RBG,K	1998-1475	00-23	c
Poales						
16	*Eriocaulon aquaticum* Druce	Tropical and warm areas	RBG,K	1998-3616	01-151	d
17	*Flagellaria guineensis* Schum.	Old world tropics	CBG	2002-747	02-73	c
18	*Rhodocoma gigantea* (Kunth) H. P. Linder	SW and East Cape	RBG,K	1996-2437	01-167	c
19	*Xyris gracilis* R.Br. ssp. *gracilis*[§]	Australia and Africa	RBG,K	1984-2761	01-147	b*
Core eudicots						
20	*Buxus sempervirens* L.	W. Europe, Mediterranean to S. Africa, temp. E. Asia, W. Indies and Central America	SP	–	02-14	b*
21	*Trochodendron aralioides* Siebold & Zucc.	Korea and Japan to Taiwan	RBG,K	2000-100	02-82	c
Higher eudicots						
Ericales						
22	*Myrsine africana* L.	Azores, Africa, Asia	SS	–	01-130	b*
23	*Planchonella eerwah* (F. M. Bailey) van Royen	Trop. America, Asia to Pacific and Africa	RBG,K	1986-2961	01-139	c
24	*Pterostyrax psilophylla* Diels ex Perkins	Burma to Japan	RBG,K	1999-4201	01-154	c
Euasterid I						
25	*Merrilliodendron megacarpum* (Hemsl.) Sleum.	Philippines and W. Pacific	RBG,K	1990-1136	02-15	c
Garryales						
26	*Garrya fremontii* Torr.	Washington to Panama and W. Indies	RBG,K	1998-2069	02-74	c
Solanales						
27	*Montinia caryophyllacea* Thunb.	S. Africa	SS	–	01-132	b*
Euasterid II						
28	*Escallonia rubra* Pers.	S. America, especially around the Andes	RBG,K	2000-2609	01-158	c

[†] Identification information: a, taxonomically verified and herbarium voucher prepared for species; b, no herbarium voucher, but species has been taxonomically verified; b*, species taxonomically verified, and is currently being grown on at RBG, Kew to prepare a herbarium voucher; c, species not taxonomically verified, but is currently being grown on at RBG, Kew to prepare a herbarium voucher; d, herbarium voucher prepared, but species has not been taxonomically verified.

[‡] Plant material obtained from RBG, Kew (RBG,K), Cambridge Botanic Gardens (CBG), Silverhill Seeds, S. Africa (SS), Syon Park Garden Centre (SP) or John Shipton Bulbs, Wales (JSB).

[§] Authority of species not known or unclear to present authors.

Xanthorrhoeaceae has also changed: in the classification of the APG (1998) the family contained just a single genus *Xanthorrhoea*, but recent phylogenetic studies have recognized that Xanthorrhoeaceae, together with Asphodelaceae and Hemerocallidaceae, form a strongly supported monophyletic group, so the new circumscription of Xanthorrhoeaceae also includes these other two families (APG II, 2003). As C-value data were already available for genera in Asphodelaceae and Hemerocallidaceae, Xanthorrhoeaceae is no longer an unrepresented family. Nevertheless, these materials are included in Table 2 as they represent first DNA C-value estimates for genera, if not for families.

The present sample, whilst highly diverse, is focused on four groups that are of particular current interest to RBG, Kew, but notably on monocots. Whereas monocots constitute approx. 20 % of angiosperm species, they form 68 % of the present sample (19 species), and provide 64 % (16 species) of the first values for 25 angiosperm families. A second focus is geographical, with over a third (42 %) of the

TABLE 2. *Chromosome number (*2n*), ploidy level (*x*), replicated genome size and nuclear DNA contents, calibration standard species and method used to estimate DNA C-values in 28 species from 25 families unrepresented in the Angiosperm DNA C-values database*

Entry no.	Taxon	Family	2*n*[†]	Ploidy level (*x*)	Genome size: 4C DNA amount/ploidy level (pg)	4C DNA amount ± s.d. (pg)	1C DNA amount (Mbp)[§]	Calibration standard species[‡]	Method[¶]
Monocots									
Alismatales									
1	*Zostera marina*	Zosteraceae	12*	2	0·63	1·26 ± 0·08	309	*Vigna*	Fe
Asparagales									
2	*Aphyllanthes monspeliensis*	Asparagaceae	~32	?	–	2·59 ± 0·04	635	*Vigna*	FC
3	*Triteleia laxa*	Asparagaceae	28*	4	10·65	42·59 ± 0·42	10 435	*Allium*	FC
4	*Astelia fragrans*	Asteliaceae	~60	8	0·63	5·06 ± 0·02	1240	*Oryza*	FC
5	*Blandfordia punicea*	Blandfordiaceae	68	4	8·13	32·53 ± 1·57	7970	*Pisum*	Fe
6	*Doryanthes palmeri*	Doryanthaceae	48	?	–	13·22 ± 0·04	3239	*Pisum*	FC
7	*Ixiolirion ledebourii*	Ixioliriaceae	~24	2	2·03	4·06 ± 0·31	995	*Vigna*	Fe
8	*Odontostomum hartwegii*	Techophilaeaceae	20	2	5·12	10·23 ± 0·09	2506	*Pisum*	FC
9	*Xanthorrhoea preissii*	Xanthorrhoeaceae	22	2	2·07	4·14 ± 0·31	1014	*Vigna*	Fe
10	*Xeronema callistemon*	Xeronemataceae	34	2 or 4	–	13·10 ± 0·06	3210	*Pisum*	FC
Dioscoreales									
11	*Narthecium ossifragum*	Nartheciaceae	26*	2	0·83	1·65 ± 0·03	404	*Vigna*	FC
Liliales									
12	*Lapageria rosea*	Philesiaceae	30 + 1B	2	13·56	27·12 ± 1·95	6644	*Pisum*	Fe
13	*Ripogonum papuanum*	Ripogonaceae	30	2	22·29	44·58 ± 2·87	10 922	*Pisum*	Fe
Commelinoids									
14	*Dasypogon hookeri*	Dasypogonaceae	14	2	0·87	1·74 ± 0·01	426	*Vigna*	FC
15	*Hanguana malayana*	Hanguanaceae	~170	?	–	6·58 ± 0·65	1612	*Hordeum*	Fe
Poales									
16	*Eriocaulon aquaticum*	Eriocaulaceae	32	4	4·19	16·74 ± 0·14	4101	*Pisum*	FC
17	*Flagellaria guineensis*	Flagellariaceae	38*	2	1·80	3·59 ± 0·03	880	*Oryza*	FC
18	*Rhodocoma gigantea*	Restionaceae	–	–	–	2·97 ± 0·04	728	*Vigna*	FC
19	*Xyris gracilis* ssp. *gracilis*	Xyridaceae	26*	2	14·02	28·03 ± 0·44	6867	*Allium*	FC
Core eudicots									
20	*Buxus sempervirens*	Buxaceae	28	2 or 4	–	3·24 ± 0·01	794	*Oryza*	FC
21	*Trochodendron aralioides*	Trochodendraceae	38	2	3·82	7·64 ± 0·02	1872	*Solanum*	FC
Higher eudicots									
Ericales									
22	*Myrsine africana*	Myrsinaceae	46	?	–	4·92 ± 0·25	1205	*Vigna*	Fe
23	*Planchonella eerwah*	Sapotaceae	~24	2	1·08	2·15 ± 0·15	527	*Vigna*	Fe
24	*Pterostyrax psilophylla*	Styracaceae	24	2	1·77	3·54 ± 0·32	867	*Vigna*	Fe
Euasterid I									
25	*Merrilliodendron megacarpum*	Icacinaceae	30	2	2·19	4·37 ± 0·07	1071	*Oryza*	FC
Garryales									
26	*Garrya fremontii*	Garryaceae	~20	2	3·04	6·08 ± 0·05	1490	*Solanum*	FC
Solanales									
27	*Montinia caryophyllacea*	Montiniaceae	24	2	1·13	2·26 ± 0·18	554	*Vigna*	Fe
Euasterid II									
28	*Escallonia rubra*	Escalloniaceae	24	2	0·85	1·69 ± 0·22	414	*Vigna*	Fe

[†] Chromosome numbers labelled with an asterisk were taken from literature, all others were determined for the present work.

[‡] Calibration standard used: *Oryza*, *Oryza sativa* IR36, 4C = 2·02 pg; *Vigna*, *Vigna radiata* 'Berken', 4C = 2·12 pg; *Solanum*, *Solanum lycopersicum* 'Gardener's Delight', 4C = 4·00 pg; *Pisum*, *Pisum sativum* 'Minerva Maple', 4C = 19·46 pg; *Hordeum*, *Hordeum vulgare* 'Sultan', 4C = 22·24 pg; *Allium*, *Allium cepa* 'Ailsa Craig', 4C = 67·1 pg.

[§] 1 pg = 980 Mbp.

[¶] Fe, Feulgen microdensitometry; FC, flow cytometry.

TABLE 3. *Minimum (min.), maximum (max.), mean, mode and range of 4C DNA values in the major subdivisions of angiosperms*

	No. of species with C-values	Min. (pg)	Max. (pg)	Mean (pg)	Mode (pg)	Range
Basal angiosperms	67	1·7	35·4	7·97	3·20	20·8
Monocots	1498	0·6	509·6	41·73	3·80	849·3
Eudicots	1978	~0·4	317·3	12·85	2·80	793·3
All angiosperms	3543	~0·4	509·6	24·97	2·20	1274·0

Data taken from the Angiosperm DNA C-values database (Bennett and Leitch, 2001) and Hanson *et al*. (2001*a*, *b*)

species being from south-east Asia and Australasia (mainly Australia). The third focus is on species of economic utility in a broad sense: wood of *Buxus sempervirens* (common box) is used to make rulers, musical instruments and croquet balls; dried leaves of *Zostera marina* are used for matting; stems of *Flagellaria guineensis* are used for basketry and fish traps; resin of *Xanthorrhoea preissii* is used to varnish or lacquer metals; and *Narthecium ossifragum* (bog asphodel) has been used as a substitute for saffron in Scotland. Cultivated garden ornamentals include *Astelia fragrans*, *Blandfordia punicea*, *Triteleia laxa*, *Xyris gracilis*, *Myrsine africana*, *Pterostyrax psilophylla* and *Trochodendron aralioides*. Other species of horticultural interest include the popular garden plants *Buxus sempervirens* and *Escallonia rubra* that are used as hedge plants. A fourth focus concerns conservation status: the Queensland rainforest tree *Planchonella eerwah*, which was not seen from the time of its naming in 1894 until its rediscovery in 1980, is endangered; *Pterostyrax psilophylla* is thought to be vulnerable in the wild; and *Blandfordia punicea* (whose flowers were found in the gut of the first emu shot in Australia in 1788) is protected in the wild (Mabberley, 1997). These foci are not mutually exclusive, and many materials were chosen to contribute to more than one. However, their prime interest is as previously unrepresented families and genera to provide useful further additions to the Angiosperm DNA C-values database.

Growth of plants

Actively growing root tips or young leaves were collected from established plants that were either potted or grown in beds (all sources listed in Table 1). *Myrsine africana* and *Montinia caryophyllacea* were grown from seed. Prior to germination, seeds were sterilized, stratified and scarified as necessary, and placed on 1 % agar in a Petri dish.

Estimating total nuclear DNA C-values

DNA C-values of the test species were estimated using either flow cytometry or Feulgen microdensitometry. The method used for each taxon is shown in Table 2 together with the calibration standard used. Several different calibration standards were used to cover the range of 4C-values encountered. The 4C-values used to convert arbitrary units into absolute values were taken from Bennett and Leitch (1995) except for *Solanum lycopersicum* L. 'Gardener's Delight' (4C = 4·00 pg), which was determined by Obermayer *et al*. (2002).

Flow cytometry. Young healthy leaf tissue was collected from the test species and calibration standard, and was co-chopped in isolation buffer. The solution was filtered through a 30-μm nylon mesh, then digested with RNase and stained with the non-base specific DNA stain propidium iodide, as described in Obermayer and Greilhuber (1999). Samples for 16 test species were analysed on a Partec PA II flow cytometer with a 100 W high pressure mercury lamp, a ×40 gel objective and a high-quality red sensitive photomultiplier. The optical bench set-up and filter types are described in Obermayer *et al*. (2002). The linearity of the machine was checked on a regular basis using chicken red blood cells. For each test species, three preparations of unknown and standard material were usually made and each was analysed five times (5000 nuclei per run). Coefficients of variation (CVs) were usually less than 3 %, otherwise the number of preparations was increased. Absolute 4C DNA values were calculated using the following formula: mean peak ratio × 4C-value of calibration standard used.

Feulgen microdensitometry. DNA C-values for 12 test species were estimated using a Vickers M85a microdensitometer using the methods described in Hanson *et al*. (2001*a*).

Chromosome counts

Chromosome counts were obtained using a standard root tip squash technique as previously described in Hanson *et al*. (2001*a*). Photographs of metaphase cells were taken on a Zeiss Photomicroscope III using Pan F film. Microscope preparations and photographs are stored at the Jodrell Laboratory, RBG, Kew. In five instances when it was not possible to obtain chromosome counts from living material, chromosome counts were taken either from Fedorov (1969) or from the 'Indexes to plant chromosome numbers' series published by the Missouri Botanical Garden (Goldblatt and Johnson, 2002).

RESULTS

Nuclear DNA amounts

Table 2 gives 4C DNA amounts estimated for the 28 taxa studied; these ranged from 1·26 pg in *Zostera marina*

(Zosteraceae) to 44·58 pg in *Ripogonum papuanum* (Ripogonaceae). Thus, these 4C values differed 35-fold, which is a narrow range restricted to the lowest 9 % of the approx. 1000-fold variation known for angiosperms as a whole. Moreover, the mean 4C DNA amount of the present sample (10·63 pg) is also low compared with that for the 3543 species (24·97 pg) in the Angiosperm DNA C-values database (Bennett and Leitch, 2001; Hanson *et al*., 2001*a*, *b*). The mean 4C value for 17 diploids (9·07 pg) was significantly lower ($P = 0{\cdot}05$) than that for four polyploids (24·23 pg) (NB seven species were excluded from this comparison as their ploidy level was unknown or unclear). Further analysis showed that replicated mean genome size (calculated as 4C DNA value/ploidy level) in 17 diploids (4·54 pg) was smaller than that in four polyploids (5·90 pg), but not significantly so ($P = 0{\cdot}68$).

The prime focus of the present work was monocots. The mean 4C-value for the 19 monocot species (13·78 pg) was significantly larger ($P = 0{\cdot}01$) than that for nine eudicot species (3·99 pg). Similarly, replicated mean genome size (2C) in 14 monocots (6·20 pg) was larger than that in seven eudicots (1·98 pg), and significantly so ($P = 0{\cdot}04$). Further analysis showed that the overall difference between monocots and eudicots was also seen in diploids alone. Thus, the mean 4C value for ten diploid monocots (12·64 pg) was larger than that for seven diploid eudicots (3·96 pg), but not significantly so ($P = 0{\cdot}10$). As the present sample did not include polyploid eudicots, a similar comparison for polyploids alone was not possible.

Chromosome counts

Table 2 gives a count (or in six cases, an estimate) of chromosome number for 22 of the 28 listed materials studied, and counts for the five remaining taxa taken from the literature. The chromosome number for *Rhodocoma gigantea* is unknown. Table 2 also gives ploidy levels for 21 of the 28 listed materials, based on a comparison of the count for the present material and all counts reported for related material(s) in the same genus, or by analysing karyotypes. Chromosome preparations for nine of the species are shown in Fig. 1, arranged in order of increasing DNA amount from *Escallonia rubra* (4C = 1·69 pg; Fig. 1A) to *Ripogonum papuanum* (4C = 44·58 pg; Fig. 1I).

A search through Fedorov (1969) and the 'Indexes to plant chromosome numbers' (Goldblatt and Johnson, 2002) showed that there were six new generic records: (1) *Dasypogon hookeri* (Dasypogonaceae; Fig. 1B); (2) *Montinia caryophyllacea* (Montiniaceae; Fig. 1C); (3) *Merrilliodendron megacarpum* (Icacinaceae; Fig. 1E); (4) *Hanguana malayana* (Hanguanaceae); (5) *Lapageria rosea* (Philesiaceae; Fig. 1G); and (6) *Ripogonum papuanum* (Ripogonaceae; Fig. 1I). *Lapageria rosea* had $2n = 30 + 1B$; this is probably the first record of B chromosomes in Philesiaceae, in which this genus is placed (APG II, 2003), although B chromosomes occur frequently in genera belonging to the order Liliales (Jones and Rees, 1982), which includes Philesiaceae.

There are also seven new species records: (1) *Blandfordia punicea* was counted with $2n = 68$ (Fig. 1H). While no previous count of 68 has been reported for other species of *Blandfordia*, a count of $2n = 34$ obtained for several species suggests that the present material is tetraploid. (2) *Eriocaulon aquaticum*: numerous counts have been reported for *Eriocaulon* species, ranging from $2n = 24$ to approx. 110. A first count of $2n = 32$ for *E. aquaticum* is reported here, agreeing with published counts for *E. septangulare* With. and *E. pellucidum* Michaux by Moldenke (1969) and Löve and Löve (1982), respectively. (3) The first count of $2n = 24$ for *Pterostyrax psilophylla* (Fig. 1D) agrees with one other published count for *P. corymbosa* by Manshard (1936). (4) *Astelia fragrans*: although there are counts for 11 other species of *Astelia*, with $2n = 16$, 60, 64, approx. 80 and 210 (Scottsberg, 1955; Wheeler, 1966), the count of $2n =$ approx. 60 is the first for *A. fragrans*. (5) *Ixiolirion ledebourii*: previous work has recorded $2n = 24$ for *I. tataricum* and *I. montanum* and $2n = 72$ in *I. tataricum*, so the first count of $2n =$ approx. 24 for *I. ledebourii* reported here agrees with previous counts for other species. (6) *Garrya fremontii*: a first count of $2n =$ approx. 20 is reported for *Garrya fremontii*. Counts previously reported for *G. elliptica* Dougl. (Meurman, 1930) and *G. lindheimeri* Torr. (Turner, 1960) had $2n = 22$. (7) The count of $2n =$ approx. 24 for *Planchonella eerwah* is related to the only other count for a *Planchonella* species, i.e. *P. sandvicensis* by Scottsberg (1955), which was tetraploid with $2n = 48$. The counts obtained for the nine remaining species all agreed with previously published counts including that for *Doryanthes palmeri* ($2n = 48$; Fig. 1F).

A chromosome count is available for only about 25 % of angiosperm species (Bennett, 1998). Although the basic technique is well established, obtaining chromosome counts for a previously unstudied taxon is neither trivial nor certain within a short period, especially if the material is rare, growth is strongly seasonal or tissue that is suitable for cytological work is in limited supply for conservation or other reasons. Clearly, obtaining a chromosome count remains highly desirable for the six species studied in this work for which no exact chromosome count could be made (see Table 2), and this remains our aim (see recommendations of the Angiosperm Genome size meeting at http://www.rbgkew.org.uk/cval/conference.htmloutline).
However, the Convention on Biological Diversity (Programme UNE, 1992) noted the need to make biodiversity data available, despite imperfections; a view that we support. Consequently, we are determined to make C-value estimates for the taxa still lacking counts available quickly, whilst attempting to obtain and publish counts for difficult materials as they become available. By the same token, publishing good estimates of chromosome number (as has been done for five species) seems highly worthwhile since this information is potentially useful until the exact number is determined or confirmed.

DISCUSSION

Survey of available DNA C-values

As part of an ongoing programme to increase familial representation in the Angiosperm DNA C-values database

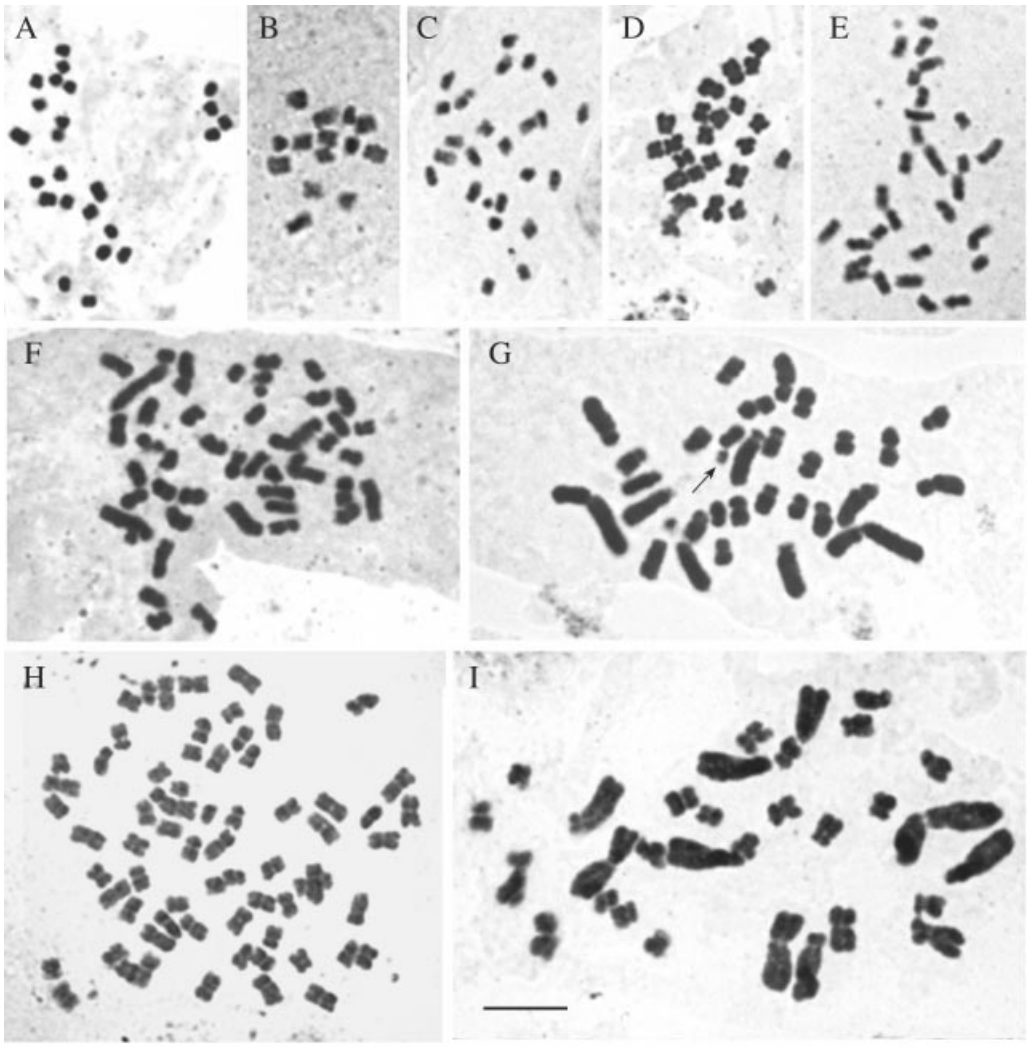

FIG. 1. Somatic chromosomes arranged (A–I) in ascending order of 4C DNA amount. A, *Escallonia rubra*, 2*n* = 24, 4C = 1·69 pg. B, *Dasypogon hookeri*, 2*n* = 14, 4C = 1·74 pg. C, *Montinia caryophyllacea,* 2*n* = 24, 4C = 2·26 pg. D, *Pterostyrax psilophylla,* 2*n* = 24, 4C = 3·54 pg. E, *Merrilliodendron megacarpum,* 2*n* = 30, 4C = 4·37 pg. F, *Doryanthes palmeri,* 2*n* = 48, 4C = 13·22 pg. G, *Lapageria rosea,* 2*n* = 30 + 1B (arrow), 4C = 27·12 pg. H, *Blandfordia punicea,* 2*n* = 68, 4C = 32·53 pg. I, *Ripogonum papuanum,* 2*n* = 30, 4C = 44·58 pg. Scale bar = 5 μm.

(Bennett and Leitch, 2001), the present paper reports C-values for a further 25 previously unrepresented families, with 16 (64 %) belonging to the monocots. The overall 4C-values differed 35-fold, with both the smallest and largest species (1·26 pg in *Zostera marina* and 44·58 pg in *Ripogonum papuanum*) being represented by monocots.

Previous work has noted that monocots may have a C-value profile very different from that of dicots, with a clear tendency for monocots to have a significantly larger DNA amount than dicots (Bennett and Leitch, 2000). The realization that angiosperms are now divided into three main groups (i.e. basal angiosperms, monocots and eudicots; APG, 1998) has not changed this trend as an analysis shows that monocots have the largest mean, mode and range of 4C-values of these major subdivisions (Table 3). Thus, while both monocots and eudicots contain species at the lowest end of the range, the largest-known monocot 4C-value (509·6 pg) is nearly 40 % bigger than the largest-known eudicot value (4C = 317·3 pg). This trend is reflected within the sample of 25 families reported here, with monocots being characterized by a significantly larger mean and range of 4C-values than eudicots. Thus, the mean for 19 monocots (mean 4C = 13·78 pg, range 1·26–44·58 pg) is nearly three times larger than that for nine eudicots (mean 4C = 3·99 pg, range 1·69–7·64 pg). Even if replicated genome size is analysed rather than 4C-values, the mean for 14 monocots (6·20 pg) is still significantly larger than that for seven eudicots (1·98 pg). Interestingly, the nine highest C-values in Table 2 are for monocot families belonging to Liliales, Asparagales or commelinoids. These three orders of monocots were shown to be the only ones that contained species with very large C-values (defined as 4C $\geqslant$ 140 pg; Leitch *et al*., 1998).

Recent research is progressing our understanding of how genome size may increase or decrease. Together with polyploidy, the role of retrotransposition (SanMiguel *et al*.,

TABLE 4. *Cumulative proportion of angiosperm families with C-value data represented in the Angiosperm DNA C-values database, Hanson* et al. *(2001*a, b*), Leitch and Hanson (2002), and the present work*

Source of data	Cumulative number of families represented	Cumulative proportion of all 453 families represented (%)	Cumulative proportion of 27 basal families represented (%)	Cumulative proportion of 345 eudicot families represented (%)	Cumulative proportion of 81 monocot families represented (%)
Angiosperm DNA C-values database (release 3·0, Sept. 2001)	135	29·8	51·8	24·0	46·9
Hanson *et al*. (2001*a*)	160	35·3	51·8	30·7	49·4
Hanson *et al*. (2001*b*)*	184	40·6	51·8	36·2	55·6
Leitch and Hanson (2002)	189	41·9	70·3	36·2	55·6
Current paper	214	47·2	70·3	38·8	75·3

* Although Hanson *et al*. (2001*b*) listed C-values for 25 families, the new familial circumscription of Asparagaceae by APG II (2003) now encompasses Haemerocallidaceae (listed in Table 3 of Hanson *et al*., 2001*b*), so this family is excluded from the above table.

1996) is now recognized to be a major factor leading to increases in DNA C-value. Bennetzen and Kellogg (1997) suggested that these processes, if left unchecked, could lead to ever-expanding 'obese' genomes. Other mechanisms are being elucidated that can bring about a reduction in DNA amount. In the work of Petrov *et al*. (2000), the rate at which DNA was lost from grasshopper genomes was slower in species with larger genomes compared with that in species having smaller genomes. Similarly, work on the repair of experimentally induced double-stranded breaks in DNA showed that in the small genome of arabidopsis, double-stranded breaks were more often repaired with larger deletions and fewer insertions than those in tobacco, which has a genome nearly 60 times larger (Kirik *et al*., 2000). Thus, genome size itself may play a role in determining the rate of DNA loss from a genome. It is tempting to speculate that beyond a critical genome size (which may vary depending on the species), it becomes increasingly difficult for the genome to 'go on a diet' and lose DNA, as the mass of DNA itself progressively prevents/inhibits the action of the mechanisms causing DNA loss. Left unchecked, this could be one way in which the truly 'obese' genomes in some monocots have evolved.

How many angiosperm families are there?

As noted previously (Hanson *et al*., 2001*a*), authorities differ as to how many angiosperm families they recognize, ranging from 200 to 533 for the eight different systems of classification listed by Brummitt (1992). Moreover, the number of families recognized can also vary with time, as new families are created and previously recognized families are split or sunk on the basis of new data (APG, 1998; APG II, 2003). Even during the course of the current work, the circumscriptions of the families Themidaceae, Xanthorrhoeaceae and Aphyllanthaceae have changed (as noted in Materials and Methods). Such changes complicate the endeavour of tracking how many, and what proportion (%) of families are represented in the database, as previously noted (Hanson *et al*., 2001*a*). The present work follows the names and circumscriptions of the APG II (2003), which recognizes 453 families comprising 27 basal angiosperms, 81 monocots and 345 eudicots. While this revised classification broadly agrees with the 1998 publication of the Angiosperm Phylogeny Group (APG, 1998) used in Hanson *et al*. (2001*a*, *b*), the overall trend of their work has been to reduce the total number of recognized families from 462 in 1998 to 453 in 2002.

Progress towards completing familial representation of DNA C-values

Combining the data in the Angiosperm DNA C-values database (Bennett and Leitch, 2001) with that in Hanson *et al*. (2001*a*, *b*) and Leitch and Hanson (2002) brings the total number of families with C-value data to 214, corresponding to 47·2 % of angiosperm families recognized by APG II (2003; Table 4). Given the specific focus on monocots, progress in this group is also assessed.

APG II (2003) recognizes 81 monocot families. A survey of available C-value data shows that prior to the current highly targeted approach of increasing familial representation in the C-values database (i.e. Hanson *et al*., 2001*a*, *b*; this paper), less that half of the moncot families (38 in total) were represented. The present paper with first C-values for 16 unrepresented monocot families, together with the work of Hanson *et al*. (2001*a*, *b*), has increased this number to 61. Thus, 75 % of all monocot families now have at least one known C-value (Table 4). Representation for monocot families is currently much higher than that for eudicots, which comprise 345 families with data available for only 38·7 % of these. The only other group of angiosperms where familial representation approaches that in monocots is the basal angiosperms, where there is at least one DNA C-value estimate for 19 of the 27 families (70·3 %) (Leitch and Hanson, 2002). Both these high familial representations are the result of specific targeted studies, demonstrating, once again, the value of this approach for improving familial representation.

The future

We plan to estimate first DNA C-values for a further 50 angiosperm families in 2002. Thus, work carried out at

RBG, Kew should have resulted in first estimates of DNA C-values for over 100 angiosperm families since 1999, increasing familial representation to over 58 %. Together, these data represent significant progress towards achieving the goal set in 1997. However, it is important to realise that working towards completing familial representation will become progressively more difficult, as the remaining unrepresented families increasingly comprise monotypic taxa from obscure locations that are difficult to access. Hitherto, obtaining materials of unrepresented families was not difficult, but it has recently become noticeably harder to access such materials for new DNA C-value estimations. For this reason, completing familial representation for angiosperms in general, and monocots in particular, may prove impossible in any short period, and progress towards this goal will become slower. Whilst collecting material of at least ten of the 20 remaining unrepresented monocot families seems feasible, many of the remaining families may be difficult to obtain including, for example, Triuridaceae (Pandanales), all members of which are achlorophyllous mycotrophic parasites, and Burmanniaceae (Dioscoreales), which are very difficult to grow in culture as they depend on adequate fungus partners which themselves are mycorrhizal partners of other green plants. It will also be difficult to collect other monocot families that are not widely in cultivation [e.g. Mayacaceae (Poales), Cymodoceaceae (Alismatales), Posidoniaceae (Alismatales)]. It will therefore be important to set realistic goals that allow for the law of diminishing returns in proportion to the effort expended. Be that as it may, progress is still likely to increase familial representation to approx. 60 % of angiosperm families by 2002. Whilst this falls far short of the target set in 1997, the achievements of the last 5 years will, nevertheless, have equalled those of the previous 40.

LITERATURE CITED

APG. 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* **85**: 531–553.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Botanical Journal of the Linnean Society* (in press).

Bennett MD. 1998. Plant genome values: How much do we know? *Proceedings of the National Academy of Sciences of the USA* **95**: 2011–2016.

Bennett MD, Leitch IJ. 1995. Nuclear DNA amounts in angiosperms. *Annals of Botany* **76**: 113–176.

Bennett MD, Leitch IJ. 2000. Variation in nuclear DNA amount (C-value) in monocots and its significance. In: Wilson KL, Morrison DA, eds. *Monocots: systematics and evolution*. Melbourne: CSIRO, 137–146.

Bennett MD, Leitch IJ. 2001. Angiosperm DNA C-values database (release 3.1, Sept. 2001). http://www.rbgkew.org.uk/cval/homepage.html

Bennett MD, Bhandol P, Leitch IJ. 2000. Nuclear DNA amounts in angiosperms and their modern uses—807 new estimates. *Annals of Botany* **86**: 859–909.

Bennetzen JL, Kellogg EA. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell* **9**: 1509–1514.

Bensasson D, Petrov DA, Zhang DX, Hartl DL, Hewitt GM. 2001. Genomic gigantism: DNA loss is slow in mountain grasshoppers. *Molecular Biology and Evolution* **18**: 246–253.

Brummitt RG. 1992. *Vascular plant families and genera*. Kew: Royal Botanic Gardens, Kew.

Fedorov A. 1969. *Chromosome numbers of flowering plants*. Leningrad: Nauka Publishers.

Goldblatt P, Johnson DE. 2002. Index to plant chromosome numbers. http://mobot.mobot.org/W3T/Search/ipcn.html

Hanson L, McMahon KA, Johnson MAT, Bennett MD. 2001*a*. First nuclear DNA C-values for 25 angiosperm families. *Annals of Botany* **87**: 251–258.

Hanson L, McMahon KA, Johnson MAT, Bennett MD. 2001*b*. First nuclear DNA C-values for another 25 angiosperm families. *Annals of Botany* **88**: 851–858.

Jones RN, Rees H. 1982. *B chromosomes*. London: Academic Press.

Kirik A, Salomon S, Puchta H. 2000. Species-specific double-strand break repair and genome evolution in plants. *EMBO Journal* **19**: 5562–5566.

Leitch IJ, Hanson L. 2002. DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. *Botanical Journal of the Linnean Society* **140**: 175–179.

Leitch IJ, Chase MW, Bennett MD. 1998. Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. *Annals of Botany* **82**: 85–94.

Löve A, Löve D. 1982. IOPB chromosome number reports LXXVII. *Taxon* **31**: 766–768.

Mabberley DJ. 1997. *The plant book*. Cambridge: Cambridge University Press.

Manshard E. 1936. Embryologische Untersuchungen an *Styrax obassia* Seib. et Zucc. *Planta* **25**: 364–386.

Meurman O. 1930. Chromosome numbers in the family Cornaceae. *Memoranda Societatis Fauna et Flora Fennica* **6**: 95–100.

Moldenke HN. 1969. Additional notes on the Eriocaulaceae XXII. *Phytologia* **18**: 344–396.

Obermayer R, Greilhuber J. 1999. *Genome* size in Chinese soybean accessions—stable or variable? *Annals of Botany* **84**: 259–262.

Obermayer R, Leitch IJ, Hanson L, Bennett MD. 2002. Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. *Annals of Botany* **90**: 209–217.

Petrov DA. 2001. *Evolution* of genome size: new approaches to an old problem. *Trends in Genetics* **17**: 23–28.

Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL. 2000. Evidence for DNA loss as a determinant of genome size. *Science* **287**: 1060–1062.

Programme UNE. 1992. Convention on Biological Diversity. United Nations Environment Programme, Nairobi, Kenya.

SanMiguel P *et al.* 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**: 765–768.

Scottsberg C. 1955. Chromosome numbers in Hawaiian flowering plants. *Arkiv fur Botanik* **3**: 63–70.

Shirasu K, Schulman AH, Lahaye T, Schulze-Lefert P. 2000. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Research* **10**: 908–915.

Turner BL. 1960. Meiotic chromosome numbers in Texas species of the genus *Coreopsis* (Compositae – Heliantheae). *Southwestern Naturalist* **5**: 12–15.

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH. 1999. Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *Plant Cell* **11**: 1769–1784.

Wheeler JM. 1966. Cytotaxonomy of the large Asteliads (Liliaceae) of the North Island of New Zealand. *New Zealand Journal of Botany* **4**: 95–113.