Phytochemistry, 1970, Vol. 9, pp. 2425 to 2430. Pergamon Press. Printed in England.

DEOXYRIBONUCLEIC ACID BASE COMPOSITION OF SOME ANGIOSPERMS AND ITS TAXONOMIC SIGNIFICANCE

S. B. BISWAS and A. K. SARKAR

Department of Botany, University of Kalyani, Kalyani, Nadia, West Bengal, India

(*Received* 27 *October* 1969, *in revised form* 24 *December* 1969)

Abstract—DNA (deoxyribonucleic acid) base composition of 61 species belonging to 35 families of angiosperms, both dicotyledons and monocotyledons, were determined with a view to assessing its taxonomic and phylogenetic significance. The data indicated a considerable similarity amongst the forms investigated, the range of GC content being about 13%. When the dicotyledons were considered separately, the range was 6%. Though the variation was found to be statistically significant, no clear cut delimitation was discernible at any taxonomic level. The results are discussed in the light of recent researches on plant DNA.

INTRODUCTION

THE IMPORTANCE of DNA (deoxyribonucleic acid) base composition in taxonomic and phylogenetic evaluation of organisms has been amply demonstrated in the cases of bacteria,[1,2] actinomycetes[3] and blue-green algae.[4] However, no investigation has yet been carried out with angiosperms except that dealing with the systematics of the genus Gossypium.[5] Several other reports concerning plant DNA[6–8] have appeared recently, but are of little value in evaluating the taxonomic or phylogenetic significance of such data.

The present study was carried out with a view to determining the overall base composition of some angiosperms and assessing its taxonomic and systematic significance. Species with different degrees of relatedness, according to current concepts of phylogeny, were examined. DNA base composition was determined by the method of Wang and Hashagen.[9] Corroborative data were also obtained by the method of Marmur and Doty[10] in some cases.

RESULTS AND DISCUSSION

The overall, base composition data as determined by the methods of Wang and Hashagen and Marmur and Doty are summarized in Table 1.

The procedures employed for the determination of mean base composition were found to be reasonably satisfactory. The results obtained with certain DNA samples of known composition were in fair agreement with previous reports arrived at by other procedures.[11]

[1] J. MARMUR, S. FALKOW and M. MANDEL, *Ann. Rev. Microbiol.* **17**, 329 (1963).
[2] J. DE LEY, *Ann. Rev. Microbiol.* **18**, 17 (1964).
[3] E. M. TEWFIK and S. G. BRADLEY, *J. Bacteriol.* **94**, 1994 (1967).
[4] M. EDELMAN, D. SWINTON, J. A. SCHIFF, H. T. EPSTEIN and B. ZELDIN, *Bacteriol. Rev.* **31**, 315 (1967).
[5] D. ERGLE, F. R. KATTERMAN and W. RICHMOND, *Plant Physiol.* **39**, 139 (1964).
[6] Y. SUYAMA and W. D. BONNER, *Plant Physiol.* **41**, 383 (1966).
[7] B. R. GREEN and M. P. GORDON, *Biochim. Biophys. Acta* **145**, 378 (1967).
[8] F. QUETIER and E. GILLES, *Arch. Biochem. Biophys.* **124**, 3 (1968).
[9] S. Y. WANG and J. M. HASHAGEN, *J. Mol. Biol.* **8**, 333 (1964).
[10] J. MARMUR and P. DOTY, *J. Mol. Biol.* **5**, 109 (1962).
[11] J. BONNER, in *Plant Biochemistry* (edited by J. BONNER and J. VARNER), p. 46, Academic Press, New York (1965).

Table 1. DNA base composition of some angiosperms

Family	Species‡	Moles, % of GC*	Moles, % of GC†	AT/GC
Monocotyledons				
Alismataceae	*Alisma plantago*	36	—	1·77
	Sagittaria sagittifolia	36·8	—	1·72
Gramineae	‡*Triticum vulgare*	47·5	48·2	1·09
	Zea mays	47·8	47·6	1·09
	Oryza sativa	—	49·0	1·04
Palmae	*Cocos nucifera*	39·5	—	1·59
Araceae	*Colocasia* sp.	38·0	—	1·63
Commelinaceae	*Commelina benghalensis*	38·5	—	1·6
Liliaceae	‡*Allium cepa*	36·8	—	1·72
Amaryllidaceae	*Zephyranthes rosea*	37·0	38·5	1·7
Musaceae	*Musa sapientum*	39·0	—	1·56
Cannaceae	*Canna indica*	40·6	—	1·49
Orchidaceae	*Vanda roxburghii*	41·0	—	1·44
Dicotyledons				
Piperaceae	*Piper betel*	38·0	—	1·63
Moraceae	*Ficus benghalensis*	38·8	—	1·58
Chenopodiaceae	*Beta vulgaris* var. *benghalensis*	38·0	—	1·63
Portulacaceae	*Portulaca oleracea*	36·8	—	1·72
Basellaceae	*Basella rubra*	37·8	—	1·64
Caryophyllaceae	*Dianthus caryophyllus*	38·0	—	1·63
Nymphaeaceae	*Nymphaea nouchali*	40·6	—	1·46
Ranunculaceae	*Ranunculus scleratus*	39·5	—	1·54
Cruciferae	*Iberis amara*	41·0	—	1·44
Crassulaceae	*Bryophyllum pinnatum*	41·4	—	1·42
Rosaceae	*Rosa* sp.	41·4	—	1·42
Leguminosae	*Phaseolus mungo*	38·8	—	1·57
	‡*Phaseolus aureus*	—	39·7	1·52
	‡*Pisum sativum*	38·5	37·1	1·64
	Lens culinare	—	38·7	1·58
	Vicia faba	—	40·0	1·5
	Cicer arietinum	—	40·2	1·49
	Glycine max	—	40·2	1·49
Oxalidaceae	*Oxalis corniculata*	39·5	—	1·54
Rutaceae	*Citrus aurantifolia*	40·4	—	1·48
Anacardiaceae	*Mangifera indica*	41·0	—	1·44
Rhamnaceae	*Zizyphus mauritiana*	40·6	—	1·46
Malvaceae	*Hibiscus rosa-sinensis*	38·8	—	1·58
Caricaceae	*Carica papya*	38·0	—	1·69
Apocynaceae	*Catharanthus roseus*	39·0	—	1·56
Verbenaceae	*Lantana aculeata*	39·5	—	1·54
	Clerodendrum viscosum	39·4	—	1·55
Solanaceae	*Solanum melongena*	38·5	36·8	1·66
	Datura metel	38·8	38·8	1·61
	Lycopersicum esculentum	—	37·3	1·68
Scrophulariaceae	*Bacopa monieri*	39·4	—	1·55
Acanthaceae	*Gendarussa vulgaris*	40·6	—	1·46
Cucurbitaceae	*Cucurbita maxima*	39·5	41·1	1·47
	Citrullus vulgaris	41·0	42·8	1·42
	‡*Lagenaria leucantha*	39·4	40·8	1·49
	Cucumis melo	41·4	39·5	1·47
	Benincasa hispida	41·0	41·9	1·42
	Cucumis trigonus	—	40·8	1·49
	Momordica charantia	—	41·1	1·44
	Cyclanthera pedata	—	41·7	1·4
	Luffa cylindrica	—	41·9	1·38

TABLE 1—*cont.*

Family	Species‡	Moles, % of GC*	Moles, % of GC†	AT/GC
	Luffa acutangula	—	42·2	1·37
	Trichosanthes dioica	—	42·2	1·37
	Trichosanthes anguinia	—	42·4	1·35
Compositae	*Tagetes* sp.	41·0	—	1·42
	Eupatorium odoratum	41·4	—	1·44
	Dahlia sp.	41·4	—	1·42
	Helianthus annus	40·6	—	1·46

* Method of Wang and Hashagen (1964).
† Method of Marmur and Doty (1962).
‡ See Ref. 11.

The values obtained by the two methods also agreed with each other and the variation observed was statistically insignificant, t being much lower than the value at 0·05 level of probability ($N = 10$). However, the relative proportions of the individual bases could not be determined with these methods. To that extent these methods fall short of the chromatographic or electrophoretic procedures.

The mean base compositions showed that the DNA of the angiosperms examined are of the AT type. Contrary to the case of microorganisms, where a considerable range of variation is observed in terms of moles % of GC, the variation in case of the angiosperms is very small. The range over all the angiosperms examined is 13%; from 36 to 49 moles % of GC. When the dicotyledons are considered separately, the range of variation is within 36·8 and 42·8 moles % of GC, only 6%. Even when the data are considered in terms of AT/GC the variation is not spectacular, being 0·73, from 1·04 to 1·77, for the angiosperms investigated. For dicotyledons it is 0·36, from 1·36 to 1·72. Nevertheless, the variations were found to be statistically significant. The critical differences calculated for monocotyledons ($N = 13$), dicotyledons ($N = 48$) and angiosperms examined as a whole ($N = 61$) were found to be 2·4, 0·52 and 0·51 respectively at 0·05 level of probability. The significance was equally pronounced when the AT/GC were analysed.

Though the base composition data have been an important taxonomic criterion in case of the microorganisms, their significance in respect of the angiosperms appears to be negligible, especially when the mean GC contents are considered. The results appear to support contentions discounting any such role.[12] It is possible that the relative proportion of individual bases could be more valuable guide in this regard. The distribution pattern of 5-methylcytosine has been shown to be of considerable taxonomic significance at infraspecific levels.[5] The present methods adopted do not allow specific measurements of any particular base and are inadequate to that extent.

The data particularly reflect the conservative nature of DNA of higher plants and are consistent with Sueoka's,[13] suggestion that the higher the organisms are on evolutionary scale, the smaller is the variation in their DNA base composition. However, as has been pointed out by Marmur *et al.*,[1] while a difference in DNA base composition indicates a

[12] R. E. ALSTON, in *Evolutionary Biology* Vol. 1 (edited by TH. DOBZHANSKY, M. K. HECHT and W. C. STEERE), p. 298.
[13] N. SUEOKA, in *The Bacteria*, Vol. V (edited by I. S. GUNSALUS and R. Y. STANIER), (1964).

difference in genetic information, a similarity does not necessarily indicate any genetic similarity. Therefore, it may be said that though the DNA base compositions point to a stable genomic constitution of the angiosperms, their evolutionary advancement and organizational complexity notwithstanding, it is possible that there exists a considerable amount of specificity with respect of the DNA of the individual species. This however, is not discernible at the level of base ratios despite the statistical significance of the data. It has been suggested that the DNA of higher organisms, including angiosperms have a significant proportion of repeating sequences[14] or duplicated genes.[15] If that be so it is quite possible that the total effect of evolutionary diversification on DNA has been considerably masked to be resolved at the level of base compositions.

In view of the facts that in the present study the total extractable DNA has been analysed and that it is likely to contain both mitochondrial[16] and chloroplastic[17] DNA, it could be that the base composition of overall DNA is somewhat different from that of nuclear DNA, since the particulate DNA are reported to have higher GC contents.[18] However, the alternate possibility that the nuclear and particulate DNA possess similar base ratios has also been contended.[19,20] Nevertheless, it is possible that the differences in base compositions of nuclear DNA are not reflected in that of the overall DNA to the desired extent. On the other hand, the differences due to these factors could be identical in case of angiosperms in general.

Statistical significance of the present data apart, recent researches amply substantiate the possibility of angiosperm DNA being more heterogeneous than would appear from base composition data. Reports of investigations concerning CsCl gradient density analyses,[6,7,21] thermal denaturation[8] and fractionation[21] of angiosperm DNA are on record. The existence of an intrinsic heterogeneity of the DNA molecules due either to interchain or intrachain organization has also been suggested.[8,22] Investigations employing DNA hybridization technique[23] have shown that there is little homology amongst the DNA of *Pisum sativum* and *Phaseolus vulgaris* although their mean base compositions are remarkably similar, both belonging to the family Leguminosae. The taxonomic importance, if any, of such heterogeneity would be a fascinating discovery.

At the present level of investigations the information available lend support to the concept of unity amongst angiosperms rather than that of their diversity. It is probable, as has been contended, that changes in angiosperm DNA during evolution, which presumably occurred at random, have been reflected in the appearance of newer sequences side by side with a core of more conservative ones.[24,25] While the conservative or "redundant" sequences accounted for the unity, the newer unique sequences were responsible for the observed diversity.[23,26] Further investigations on these aspects, involving a large number of plants, at more refined levels are expected to bring out many useful and interesting observations.

[14] R. J. BRITTEN and D. KOHNE, *Science* **161**, 529 (1968).
[15] M. NEI, *Nature* **221**, 41 (1969).
[16] A. GIBOR and S. GRANICK, *Science* **145**, 890 (1964).
[17] D. J. L. LUCK and E. REICH, *Proc. Natl. Acad. Sci. U.S.* **52**, 931 (1964).
[18] H. G. DUBUY and F. L. RILEY, *Proc. Natl. Acad. Sci. U.S.* **57**, 790 (1967).
[19] J. N. LYTTLETON and G. B. PETERSEN, *Biochem. Biophys Acta* **80**, 394 (1964).
[20] R. BAXTER and J. T. O. KIRK, *Nature* **222**, 273 (1969).
[21] E. GUILLE, F. QUETIER and J. CAU, *Bull. Soc. Franc. Physiol. Vegetale* **13**, 221 (1967).
[22] E. P. GEIDUSCHEK, *J. Mol. Biol.* **4**, 467 (1964).
[23] A. J. BENDICH and E. T. BOLTON, *Plant Physiol.* **42**, 959 (1967).
[24] B. J. MCCARTHY, *Bacteriol. Rev.* **31**, 215 (1967).
[25] B. J. MCCARTHY, in *Prog. Nucl. Acid Res.* **4**, 129 (1965).
[26] R. J. BRITTEN and M. WARING, in *Carnegie Inst. Wash. Yearbook*, p. 316, Garamound/Pridemark Press, Baltimore, Maryland (1965).

EXPERIMENTAL

Plant Material

Sixty-one species belonging to 35 families, monocotyledons as well as dicotyledons, were examined. The species investigated are listed in Table 1. The families are arranged in phylogenetic order, according to Engler and Prantl.[27] Usually DNA was isolated from seeds, fruit mesocarps or leaves depending on availability. Seeds of the following species were purchased from Sutton & Co., Calcutta: *Lagenaria leucantha, Cucurbita maxima, Citrullus vulgaris, Benincasa hispida, Cucumis melo, Luffa acutangula, Luffa cylindrica, Momordica charantia, Pisum sativum, Phaseolus mungo, Vicia faba, Glycine max, Cicer arietinum* and *Phaseolus aureus.*

The rest of the forms investigated were collected from the University campus. In some cases fruits were purchased from the local markets. Seeds of some cereals were also obtained from the University farms.

Preparation of DNA

DNA samples were prepared by the methods of Bendich and Bolton,[23] Katterman[28] and Kirby[29] with several modifications. No distinction was made with respect to the methods of isolation while interpreting the data. In all cases tissues were primarily processed for the removal of fats and pigments according to the method of Katterman[28] with some modifications. Deproteinization of the samples was carried out by shaking with $CHCl_3$–*iso*-amyl alcohol (24:1) or $CHCl_3$–octanol (4:1) or H_2O–saturated phenol. RNA was removed by treatment with RNase (20 μg/ml final concentration), solution previously heated at 100° for 4 min to inactivate any contaminant (DNase) and incubated at 37° for 45 min. Alternative procedures for the removal of RNA such as selective precipitation with iso-propanol or treatment with 0·3N KOH, final concentration, were also used depending upon the method of base composition determination. In some of the cases the procedures were used conjointly.

In all the experiments concerning helix–coil transition, purified DNA samples (20–35 μg/ml) were used. RNA as impurity did not exceed 5% in any of the cases. Protein was not detectable by standard biuret test. In samples prepared for analyses after bromination, the concentration of DNA varied from 20 μg to 40 μg/ml. The samples were purified as far as possible and generally RNA and protein remained undetected by orcinol[30] and biuret tests respectively. DNA in the samples was estimated by the diphynylamine test of Burton.[31]

DNA Base Composition Determination

(i) *Bromination method of Wang and Hashagen.* This method, founded upon the difference in reactivity between adenine and the other bases towards bromine reagents[32,33] was used, with *N*-bromoacetamide (Fluka, Switzerland) as the brominating agent. The method as prescribed by Wang and Hashagen was followed rigorously with only slight adjustments and the calibration chart prepared by them was utilized.

The DNA samples, purified as mentioned above were reprecipitated in 2 vol. of cold 95% EtOH, dissolved in standard saline citrate and dialysed against H_2O for 36 hr with several changes. The dialysed solution was either reprecipitated in 95% EtOH and the precipitate obtained then dissolved in N H_2SO_4 by constant shaking, or diluted with N H_2SO_4, to give extinction values of 0·4–0·8 at the u.v. maxima. Insoluble impurities were removed earlier by filtration.

The absorbancy readings were recorded at 270 mm and 300 mm, the difference giving the absorptivity reading before bromination (O.D.B.). 3 ml Aliquots of the DNA solution and the reference N H_2SO_4 were allowed to react for 2 hr with 0·1 and 0·075 ml of a 6 mM solution of *N*-bromoacetamide respectively. The absorbancy readings were noted again at the two wave lengths. The percentage absorbancy remaining (P) was calculated from these readings and the values referred to a previously prepared calibration chart to obtain the % values of GC. At least three readings were taken for each DNA sample and averaged. The results were calculated for the regular bases only.

(ii) *From Helix–coil transition profiles.* The helix-coil transition profiles were studied by the method of Marmur and Doty. DNA samples were dissolved in a solvent of low ionic strngth, 0·01 M phosphate buffer containing 0·001 M EDTA, pH 7. Purity of samples was checked by taking absorbancy readings at 260, 280 and 320 nm.

All the experiments were carried out with a Hilger-Watts spectrophotometer equipped with special holders for the cells. Water with gradually increasing temperature was circulated through the holders from a thermostatically controlled water bath. The temperature of the solutions in the cells was measured with the help of a thermistor bead. The absorptivity of the sample solution was noted at gradually increasing temperatures till

[27] A. ENGLER and K. PRANTL, *Die Naturlichen Pflanzenfamilien*, Leipzig (1926).
[28] ID., ERGLE, F. KATTERMAN and W. RICHMOND, *Plant Physiol.* **36**, 811 (1961).
[29] ID., S. KIRBY, *Biochem. J.* **66**, 495 (1967).
[30] W. MEJBAUM, *Z. Physiol. Chem.* **258**, 117 (1939).
[31] K. BURTON, *Biochem. J.* **62**, 325 (1956).
[32] H. ISHIHARA and N. SUZUKI, *Nature* **182**, 1302 (1958).
[33] A. S. JONES and D. L. WOODHOUSE, *Nature* **183**, 1603 (1959).

these became stationary. The values were corrected for volume expansion at all temperatures. The 50% increase in relative absorbance was noted and the corresponding temperature, T_m, determined from the melting profiles. GC content was then calculated from the T_m values according to the empirical formula of Marmur and Doty.

Acknowledgement—The authors are grateful to Prof. S. P. Sen, Head of the Department of Botany, University of Kalyani, for his constant interest and helpful suggestions throughout the work. They are also thankful to him for going through the manuscript. One of them (S. B. B.) thankfully acknowledges the award of a scholarship to him by the University of Kalyani. The authors are also thankful to Drs. A. Sen and S. Mitra and Mr. J. Dutta of the Bose Research Institute, Calcutta for providing facilities for spectrophotometric studies.