International Journal of Bioinformatics Research, ISSN: 0975–3087, Volume 2, Issue 2, 2010, pp-42-60

Molecular phylogenetic study of *Luffa tuberosa* Roxb. (Cucurbitaceae) based on internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA and its systematic implication

Ajmal Ali M.[1]*, Karuppusamy S.[2] and Fahad M. Al-Hemaid[1]
[*1]Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, 11451, Saudi Arabia, ajmalpdrc@gmail.com
[2]Department of Botany, The Madura College, Madurai, Tamil Nadu, India

Abstract- The phylogenetic position of long been debatable species *Luffa tuberosa* was inferred in the present study using ITS sequence of nuclear ribosomal DNA data. The study sampled a total number of 16 accessions which include five accessions of *Luffa* (under four species i.e. *Luffa acutangula, L. cylindrica, L. aegyptiaca* and *L. tuberosa*), nine accessions of *Momordica* (under eight species i.e. *M. angustisepala*, *M. balsamina*, *M. cabraei*, *M. charantia*, *M. charantia* subsp. *macroloba*, *M. cissoides*, *M. cochinchinensis*, *M. dioica* and *M. foetida*) and two accessions of *Trichosanthes* under two species (i.e. *T. lepiniana* and *T. tricuspidata*). The sequence data analysis clearly reveals nesting of *Luffa tuberosa* within the clade of *Momordica*, thus, we herein support the inclusion of *Luffa tuberosa* into the genus *Momordica* as *M. tuberosa* (Roxb.) Cogn.
Key Words: *Luffa tuberosa*, *Momordica,* Cucurbitaceae, ITS, nuclear ribosomal DNA

INTRODUCTION

The genus *Luffa* Miller, belongs to Tribe Luffeae, Subfamily Cucurbitioideae, Family Cucurbitaceae are distributed mainly in tropical regions of the world [1]. Chakravarty, 1982 [2] recognized a total number of nine species of *Luffa* out of which seven species (*L. acutangula*, *L. cylindrica, L. echinata*, *L. graveolens*, *L. hermaphrodita*, *L. tuberosa* and *L. umbellata*) occur in India. *Luffa tuberosa* is distributed in Peninsular India (Andhra Pradesh, Karnataka, Madhya Pradesh and Maharashtra) and Tropical Africa. The species, *Luffa tuberosa* has been characterized morphologically by perennial climbing herb, monoecious; stems: slender, scandent, striate, pubescent, arising from a small perennial tuber; leaf-blade: reniform-orbicular or pentagonal in outline, cordate, obscurely sinuate to distinctly sinuate-toothed, glabrous or sparsely hairy especially on the nerves beneath, obscurely 5-lobed; petiole pubescent; tendrils simple; male flowers: subtended by a minute bract, petals yellow, stamens 3 or 2; female flowers: peduncle cylinder, ebracteate, stigma 2 or 3, bipartitie; ovary: fusiform, glabrous or pubescent; fruit: fleshy, fusiform, shortly beaked, pubescent, longitudinally ribbed; seeds: subglobose, rugose-appendaged at one end, testa smooth. The tender fruits of *L. tuberosa* are used as diuretic and laxative in the traditional system of medicine in India [2]. Ayyangar, 1976 [3] reported chromosome No. 2n= 22 in *Momordica tuberosa* (= *L. tuberosa*).

The genus *Momordica* L. belongs to Tribe Joliffieae, Subtribe Thladianthinae, Family Cucurbitaceae [1] comprises c. 45 species of annual or perennial climbing herbaceous or shrubby plants, natives of tropical and subtropical Africa, Asia and Australia. The herbaceous, tendril-bearing vine grows up to 5 meter. It bears simple, alternate leaves, 4-12 cm across, with 3-7 deeply separated lobes [1].

The taxonomic position of *Luffa* aroused much interest after Jeffrey (1962) created a new Subtribe Luffinae under Cucurbitaceae to provide separate rank of Subtribe to the genus *Luffa* [4]. Singh (1964) has justified the position of *Luffa* in Jeffrey's classification on the basis of comparative study of the endosperm haustorium in the family [5]. The species *Luffa tuberosa* was established by Roxburgh in 1832 [6]. Clarke (1879) transferred *Luffa tuberosa* to the genus *Momordica* under *Momordica cymbalaria* Fenzl [7]. Congiaux (1881) recognized *Momordica tuberosa* (Roxb.) Cogn., based on Roxburgh's *Luffa tuberose* [8]. The phylogenetic position and taxonomic status of *Luffa tuberosa* within the genus is debatable [9].

The aim of this study is to evaluate phylogenetic position of *Luffa tuberosa* by comparing sequences of the internal transcribed spacer regions of the nuclear ribosomal DNA with species of *Momordica*.

MATERIALS AND METHODS

Present study sampled a total number of 16 accessions which include five accessions of *Luffa*, nine accessions of *Momordica* and two accessions of *Trichosanthes* (outgroup). The plant material of *Luffa acutangula*, *L. cylindrica*, *L. tuberosa, Momordica dioica*, *Trichosanthes lepiniana* and *T. tricuspidata* were collected in nature from the state Bihar, West Bengal and Andhra Pradesh of India. GenBank accession numbers along with voucher information of newly generated sequence for this study are listed in the Table 1. The voucher specimens submitted in BHAG (Herbarium, Tilka Manjhi Bhagalpur

University, Bhagalpur, Bihar, India) and SKU (Herbarium, Sri Krishnadevaraya University, Anantapur, Andhra Pradesh, India). To compare and analyse the ITS sequences of nuclear ribosomal DNA, eight accessions of *Momordica* under eight spices i.e. *M. angustisepala* (AM981080), *M. balsamina* (AM981071), *M. cabraei* (AM981084), *M. charantia* (AM981062), *M. charantia* subsp. *macroloba* (AM981061), *M. cissoids* (AM981079), *M. cochinchinensis* (AY606266) and *M. foetida* (AM981065) and two accessions of *Luffa* under two species i.e. *L. aegyptiaca* (AM981167) and *L. cylindrica* (AF013324) were retrieved from the NCBI GenBank database. The genus *Luffa*, *Momordica* and *Trichosanthes* belong to subfamily Cucurbitoideae [1]. Based on chloroplast DNA sequences [10] from two genes, one intron, two spacers (*rbcL, matK, trnL, trnL-trnF, rpl20- rps12*) and on ITS1 and ITS2 sequences of nrDNA [11] suggested a close relationship between Tribe Luffeae and Trichosantheae to which *Luffa* and *Trichosanthes* belong (73% bootstrap support) and in between Trichosantheae and Joliffieae to which *Trichosathes* and *Momordica* belong. Therefore, for the phylogenetic analysis, the species of *Trichosanthes* were selected as outgroup.

Total DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Amsterdam, Netherlands). ITS sequences of nuclear ribosomal DNA were amplified using primers ITS1 (Forward 5'-GTCCACTGAACCTTATCATTTAG-3') and ITS4 (Reverse 5'-TCCTCCGCTTATTGATATGC-3') [12] via the polymerase chain reaction (PCR) using the AccuPower HF PCR PreMix (Bioneer, Daejeon, South Korea) in 20 µL volumes containing 2 µL of 10X buffer, 300 µM dNTPs, 1 µL of a 10 pM solution of each primer, 1 unit of HF DNA polymerase. The initial denaturation at 94℃ for 5 min, and followed by 40 cycles of 94℃ for 1 min, 49℃ for 1 min, and 72℃ for 1 min, with a final extension step of 72℃ for 5 min. The PCR products were ligated into the pT7Blue cloning vector using Perfectly Blunt Cloning Kit (Novagen, Inc.) according to the manufacturer's instructions. Resulting recombinant plasmids were used to transform competent cells included in the kit. The transformation mix was incubated in 250 µl SOC medium for 1hour at 37℃ on a rotary shaker, then plated on LB agar with 50 µg/mL ampicillin. Colonies were randomly selected and were put into PCR buffer. The PCR products were purified with the SolGent PCR Purification Kit-Ultra (SolGent, Daejeon, South Korea) prior to sequencing. The purified fragments were directly sequenced using dye terminator chemistry following the manufacturer's protocol. Cycle sequencing was conducted using same primers used in amplification and BigDye vers. 3 reagents and an ABI PRISM 3730XL DNA Analyzer (Perkin-Elmer, Applied Biosystems) by following the manufacturer's instructions. Cycling conditions included an initial denaturing set at 94℃ for 5 min., followed by 30 cycles of 96℃ for 10 sec., 50℃ for 5 sec., and 60℃ for 4 minutes. Each sample was sequenced in the sense and antisense direction. The sequences were analyzed with ABI Sequence Analysis and ABI Sequence Navigator software (Perkin-Elmer/Applied Biosystems). Nucleotide sequences of both DNA strands were obtained and compared to ensure accuracy.

Initially the sequence alignments were performed using ClustalX version 1.81 [13] with gap opening penalty = 10 and gap extension penalty = 3.0. Sequence alignments were subsequently adjusted manually using BioEdit [14] and SeaView [15]. Insertion-deletions (Indels) were scored as single characters when we had confidence in positional homology (Annexure I). The boundaries between the ITS1, 5.8S, and ITS2 were determined by comparisons with earlier published sequences available at National Center for Biotechnology Information (NCBI) GenBank (www.ncbi.nlm.nih.gov). Gaps were treated as missing data in phylogenetic analyses. All sequences generated in the present study were deposited in GenBank and GenBank accession number included in Table 1. Parsimony analyses were performed with PAUP* 4.0b10 [16]. Heuristic searches were conducted using 10,000 random addition sequence replicates, holding 10 trees at each step, and with tree-bisection-reconnection (TBR) branch swapping, characters equally weighted, and gaps treated as missing data. Support for internal nodes was assessed using bootstrap analysis [17] of 1000 replicates with 100 random additions per replicate and holding 10 trees at each step. Phylogenetic and molecular evolutionary analyses (evolutionary divergence between sequences, the number of base substitutions per site from averaging evolutionary divergence over all sequence pairs, homogeneity test of substitution patterns between sequences, base composition bias difference between sequences, maximum composite likelihood estimate of the pattern of nucleotide substitution, codon-based test of neutrality for analysis between sequences, and Fisher's exact test of neutrality for sequence pairs) were conducted using MEGA version 4 [18-21]. Parsimony analyses were also performed using MEGA4. The result was verified with Maximum Likelihood method (using SeaView) and Baseyan analysis (Mr Bayes). For Bayesian analysis, the best-fit model of nucleotide evolution was found using jModelTest v1.0.1 [22]. Bayesian posterior probabilities for the clades were obtained using Metropolis-coupled Markov chain Monte Carlo analysis as implemented in MrBayes. Two simultaneous independent runs with four Markov chains were done for 5 million generations, and trees were

sampled every 100th generation, resulting in 50,000 trees. The first 10,000 trees were considered as the burn-in phase and discarded. A majority-rule consensus tree based on the remaining 40,000 trees was computed.

RESULTS

***Sequence Characteristics*:** The combined length of the entire ITS region (ITS1, 5.8S and ITS2) from taxa sampled in the present study ranged from 600-686 bp. The ITS1 region and %GC ranged from 201-243 bp in length and 61-69% respectively, the 5.8S gene was 164 bp, the ITS2 region and %GC ranged from 218-296 bp in length and 65-73% respectively (Table 2). Data matrix has a total number of 766 characters of which 500 characters are constant, 104 characters are variable but parsimony-uninformative and 153 characters are parsimony-informative. Insertions and deletions (indels) were necessary to align the sequences. Indels were ranged from 1 to 22 bp.

***Phylogenetic analyses*:** The parsimony analysis (using PAUP) of the entire ITS region resulted in eight maximally parsimonious trees (MPTs) with a total length of 486 steps, a consistency index (CI) of 0.7366 (0. 6484 excluding uninformative characters), a homoplasy index (HI) of 0.2634 (0. 3516 excluding uninformative characters), rescaled consistency index (RC) of 0.5195 and a retention index (RI) of 0.7051. The bootstrap values above the line in bootstrap strict consensus tree (Fig. 1) show the relative support of each clade. The number of base substitutions per site from analysis between sequences (evolutionary divergence between sequences) is shown in Table 3. The number of base substitutions per site from averaging evolutionary divergence over all sequence pairs was found 0.081.

Homogeneity test of substitution patterns between sequences: The probability of rejecting the null hypothesis that sequences have evolved with the same pattern of substitution, as judged from the extent of differences in base composition biases between sequences. A Monte Carlo test (1000 replicates) was used to estimate the *P*-values, which are shown in diagonal in the Table 4. *P*-values smaller than 0.05 are considered significant. The estimates of the disparity index per site are shown for each sequence pair above the diagonal.

Base composition bias difference between sequences: The difference in base composition bias per site is shown in table 5. Even when the substitution patterns are homogeneous among lineages, the compositional distance correlates with the number of differences between sequences.

***Maximum composite likelihood estimate of the pattern of nucleotide substitution**:* Each entry shows the probability of substitution from one base (row) to another base (column) instantaneously. Rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics. The nucleotide frequencies are 0.18 (A), 0.177 (T/U), 0.35 (C), and 0.292 (G). The transition/transversion rate ratios are k_1 = 1.692 (purines) and k_2 = 2.866 (pyrimidines). The overall transition/transversion bias is R = 1.494 (Table 6).

Codon-based test of neutrality for analysis between sequences: The probability of rejecting the null hypothesis of strict-neutrality ($d_N = d_S$) (Codon-based test of neutrality for analysis between sequences) is shown in Table 7 (below diagonal). Values of *P* less than 0.05 are considered significant at the 5% level. The test statistic (synonymous substitutions d_N - nonsynonymous substitutions d_S) is shown above the diagonal.

Fisher's exact test of neutrality for sequence pairs: The probability (*P*) of rejecting the null hypothesis of strict-neutrality in favor of the alternative hypothesis of positive selection is shown for each sequence pair (Table 8). *P* values smaller than 0.05 are considered significant at the 5% level.

DISCUSSION

The parsimony analysis (using PAUP) of ITS sequence of nrDNA data of species of *Luffa* and *Momordica* clearly reveals two major group i.e. *Luffa* group (100% bootstrap support) and *Momordica* group with *L. tuberosa* (94% bootstrap support). Both the groups show strong relationship to each other (99% bootstrap support). *L. tuberosa* nested as a polytomy within base of *Momordica* group (Fig. 1). The relationship was found consistent (Fig. 2-8) when the results were verified with Maximum Parsimony method (using MEGA), Maximum Likelihood method (using SeaView) and Baseyan analysis (using Mr Bayes).

Luffa tuberosa was established by Roxburgh in 1832 [6]. Clarke (1879) transferred *L. tuberosa* to the genus *Momordica* under *Momordica cymbalaria* Fenzl [7]. Cogniaux (1881) recognized *Momordica tuberosa* (Roxb.) Cogn. based on Roxburgh's *L. tuberose* [8]. According to Chakravarty (1982) the fruit is a specific character in *Luffa* and there is no reason to shift this species to *Momordica* which has either muriculate or echinate fruits but never angular

[2]. Roxburgh's note in the original description is interesting. He stated that the fruit is exactly like *L. amara* Roxb. (= *L. acutangula* var. *amara*) but without stopple. The leaves of all the species of *Momordica* contain true cystoliths on the lower surface [2]. Cystoliths are absent in this species. Foliaceous bracts which are common features within the genus *Momordica* are also absent in *L. tuberosa* [2]. The seed coat anatomy does not support the inclusion of *M. cymbalaria* (= *L. tuberosa*) under *Luffa* [23-24]. Based on seed fat characteristics [25] supported the retention of *L. tuberosa* under the genus *Momordica*. Seed fat of *M. tuberosa* (=*L. tuberosa)* contains a conjugated triene acid which is characteristic of seed fat of the genus *Momordica*, however, on the other hand, genus *Luffa* does not contain conjugated triene acid [25]. Owing to proximity of *Luffa tuberosa* with the genus *Momordica*, we herein also (based on nrDNA ITS sequences data analysis using Maximum Parsimony method, Maximum Likelihood method and Baseyan analysis) support inclusion of *Luffa tuberosa* into the genus *Momordica* as *M. tuberosa* (Roxb.) Cogn.

ACKNOWLEDGEMENT

The senior author (MAA) would like to acknowledge sincere thank to Professor Arun K. Pandey, Department of Botany, Delhi University, Delhi, India who was responsible for career in plant systematics, through whom only it become possible to discover the joys of botanical field work and DNA sequence based experimental taxonomy - a passion that has continued to this day also. The senior author (MAA) also would like to acknowledge sincere thank to Dr Joongku Lee, International Biological Material Research Centre (IBMRC), Plant Diversity Research Centre (PDRC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, South Korea. This study was partially supported by Post Doctoral Fellowship from International Biological Material Research Centre (IBMRC), Plant Diversity Research Centre (PDRC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, South Korea to Dr. M. Ajmal Ali. We are grateful to anonymous reviewers for constrictive comments.

REFERENCES

[1] Jeffrey C. (2005) *Bot Zhurn,* 90: 332–335.
[2] Chakravarty H. L. (1982) *BSI*, Howrah, India.
[3] Ayyangar K. R. (1976) Today and Tomorrow's Printers & Publishers, New Delhi. pp. 85-116.
[4] Jeffrey C. (1962) *Kew Bull,* 15:337–371.
[5] Singh D. (1964) *Proc Ind Acad Sci,* 60: 399-413.
[6] Roxburgh W. (1832) 3: 717. Serampore.
[7] Clarke C. B. (1879) Reeve & Co., London, 2: 618.
[8] Cogniaux A. (1881) *Mon Phan,* 3: 325–951, 979–1008.
[9] Pandey A. K., Varma S. K. and Ali M. A. (2006) Pointer Publisher, Jaipur, India, pp. 260-270.
[10] Kocyan A., Zhang L. B., Schaefer H. and Renner S. S. (2007) *Mol Phylo Evol*, 44:553–577.
[11] Jobst J., King K. and Hemleben V. (1998) *Mol Phylo Evol,* 9:204-219.
[12] White T. J., Bruns T., Lee S. and Taylor J. (1990) San Diego, California: Academic Press. pp. 315-322.
[13] Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins D. G. (1997) *Nucl Acid Res,* 24: 4876–4882.
[14] Hall T. A. (1999). *Nucl Acid Symp Ser,* 41: 95-98.
[15] Gouy M., Guindon S., and Gascuel O. (2010) *Mol Bio Evol,* 27(2): 221-224.
[16] Swofford D. L. (2002) Sinauer, Sunderland, Massachusetts.
[17] Felsenstein J. (1985) *Evol,* 39: 783-791.
[18] Nei M. and Gojobori T. (1986) *Mol Phylo Evol,* 3:418-426.
[19] Kumar S. and Gadagkar S. R. (2001) *Genetics,* 158:1321-1327.
[20] Tamura K., Nei M. and Kumar S. (2004) *Proc Nat Acad Sci (USA),* 101:11030-11035.
[21] Tamura K., Dudley J., Nei M. and Kumar S. (2007) *Mol Bio Evol,* 24:1596-1599.
[22] Posada D. (2008) *Mol Biol Evol,* 25: 1253–1256
[23] Singh D. (1971) *J Ind Bot Soc,* 50: 208-215.
[24] Singh D. and Dathan A. S. R. (2001) Printwell Publishers Distributors, Jaipur, India, pp. 87–114.
[25] Azeemoddin G and Rao S. D. T. (1967) *Curr Sci*, 36:100.

Table 1- *Plant accessions used and ITS sequences generated for the molecular systematic study species of Luffa and Momordica*

Taxon	**Voucher**	**Geographic origin**	**GenBank Accession No.**
Luffa acutangula (L.) Roxb.	M. Ajmal Ali and A. K. Pandey 1061 (BHAG) Bihar,	India	GQ183044
L. cylindrica (L.) M. Roem.	M. Ajmal Ali and A. K. Pandey 1089 (BHAG) Bihar,	India	GQ183045
L. tuberosa Roxb.	S. Karuppusamy 28631 (SKU)	Andhra Pradesh, India	GQ183046
Momordica dioica Roxb. ex. Willd	M. Ajmal Ali and A. K. Pandey 1079 (BHAG) Bihar,	India	GQ183048
Trichosanthes lepiniana (Naudin) Cogn.	M. Ajmal Ali and A. K. Pandey 20052 (BHAG)	West Bengal, India	GQ183049
T. tricuspidata Lour.	M. Ajmal Ali and A. K. Pandey 1060 (BHAG) Bihar,	India	GQ183050

Table 2- *Length and GC Contents of ITS1 and ITS2*

Species	ITS 1		ITS2	
	Size (bp)	%GC	Size (bp)	%GC
Luffa cylindrica	201	61	258	65
L. aegyptiaca	201	63	257	66
L. acutangula	201	61	235	65
L. cylindrica	201	63	254	66
L. tuberosa	236	69	254	70
Momordica cabraei	219	67	296	70
M. balsamina	243	62	266	70
M. foetida	238	64	282	72
M. charantia subsp. *macroloba*	234	64	269	68
M. charantia	234	65	270	68
M. angustisepala	215	64	290	70
M. dioica	211	69	218	73
M. cissoides	213	66	276	70
M. cochinchinensis	207	67	276	69
Trichosanthes lepiniana	200	61	236	66
T. tricuspidata	199	61	235	66

Table 3- Evolutionary divergence between sequences of species of *Luffa* and *Momordica*

* ↓															
Coch															
Acut	0.093														
Bals	0.071	0.088													
Foet	0.069	0.086	0.021												
Ciss	0.038	0.095	0.080	0.075											
Char	0.069	0.085	0.045	0.033	0.069										
Chma	0.071	0.087	0.048	0.036	0.071	0.002									
Cyli	0.088	0.006	0.086	0.084	0.088	0.083	0.085								
Tric	0.084	0.066	0.081	0.079	0.095	0.087	0.090	0.064							
Aegy	0.088	0.006	0.086	0.084	0.088	0.083	0.085	0.000	0.064						
Lepi	0.086	0.068	0.082	0.082	0.097	0.090	0.092	0.066	0.010	0.066					
Dioi	0.037	0.099	0.086	0.084	0.050	0.090	0.092	0.095	0.090	0.095	0.093				
Tube	0.137	0.165	0.134	0.129	0.132	0.134	0.136	0.163	0.158	0.163	0.163	0.136			
Angu	0.040	0.093	0.077	0.073	0.038	0.075	0.077	0.088	0.084	0.088	0.086	0.048	0.120		
Cygb	0.093	0.000	0.088	0.086	0.095	0.085	0.087	0.006	0.066	0.006	0.068	0.099	0.165	0.093	
Cabr	0.063	0.092	0.066	0.062	0.065	0.075	0.077	0.086	0.073	0.086	0.071	0.071	0.127	0.060	0.092

*** Abbreviation of Taxon (Annexure II)**

Table 4- Test of the homogeneity of substitution patterns between sequences of *Luffa* and *Momordica*

* ↓																
Ciss		0.000	0.240	0.142	0.073	0.071	0.069	0.150	0.223	0.150	0.169	0.023	0.050	0.069	0.240	0.000
Coch	1.000		0.157	0.061	0.006	0.006	0.008	0.073	0.142	0.073	0.106	0.033	0.000	0.010	0.157	0.000
Acut	0.014	0.051		0.000	0.121	0.015	0.000	0.008	0.000	0.008	0.000	0.451	0.213	0.013	0.000	0.173
Bals	0.043	0.121	1.000		0.033	0.000	0.000	0.000	0.000	0.000	0.000	0.228	0.050	0.000	0.000	0.084
Foet	0.114	0.358	0.066	0.054		0.015	0.031	0.036	0.081	0.036	0.077	0.058	0.012	0.000	0.121	0.061
Char	0.102	0.376	0.284	1.000	0.237		0.000	0.000	0.000	0.000	0.010	0.132	0.000	0.000	0.015	0.000
Chma	0.114	0.322	1.000	1.000	0.130	1.000		0.000	0.000	0.000	0.000	0.154	0.000	0.000	0.000	0.000
Cyli	0.047	0.143	0.121	1.000	0.227	1.000	1.000		0.000	0.000	0.000	0.309	0.134	0.000	0.008	0.106
Tric	0.016	0.029	1.000	1.000	0.123	1.000	1.000	1.000		0.000	0.000	0.409	0.202	0.004	0.000	0.184
Aegy	0.038	0.153	0.109	1.000	0.234	1.000	1.000	1.000	1.000		0.000	0.309	0.134	0.000	0.008	0.106
Lepi	0.040	0.073	1.000	1.000	0.109	0.348	1.000	1.000	1.000	1.000		0.365	0.223	0.002	0.000	0.169
Dioi	0.237	0.111	0.000	0.012	0.158	0.048	0.046	0.009	0.000	0.004	0.000		0.025	0.140	0.451	0.038
Tube	0.220	1.000	0.070	0.229	0.346	1.000	1.000	0.113	0.075	0.128	0.060	0.295		0.000	0.213	0.000
Angu	0.035	0.325	0.323	1.000	1.000	1.000	1.000	1.000	0.365	1.000	0.387	0.006	1.000		0.013	0.002
Cygb	0.011	0.031	1.000	1.000	0.068	0.337	1.000	0.102	1.000	0.097	1.000	0.001	0.074	0.316		0.173
Cabr	1.000	1.000	0.034	0.070	0.116	1.000	1.000	0.077	0.019	0.071	0.015	0.185	1.000	0.389	0.033	

*** Abbreviation of Taxon (Annexure II)**

Table 5- Base composition bias difference between sequences of *Luffa* and *Momordica*

* ↓															
Ciss															
Coch	0.013														
Acut	0.328	0.244													
Bals	0.217	0.129	0.056												
Foet	0.144	0.071	0.202	0.054											
Char	0.136	0.071	0.096	0.017	0.048										
Chma	0.136	0.075	0.079	0.019	0.065	0.002									
Cyli	0.232	0.155	0.013	0.023	0.115	0.052	0.042								
Tric	0.311	0.221	0.006	0.036	0.155	0.083	0.071	0.006							
Aegy	0.232	0.155	0.013	0.023	0.115	0.052	0.042	0.000	0.006						
Lepi	0.259	0.186	0.017	0.054	0.154	0.094	0.081	0.008	0.010	0.008					
Dioi	0.071	0.069	0.543	0.309	0.136	0.217	0.240	0.397	0.493	0.397	0.451				
Tube	0.169	0.121	0.359	0.171	0.129	0.090	0.106	0.278	0.342	0.278	0.367	0.148			
Angu	0.106	0.048	0.100	0.021	0.036	0.004	0.006	0.048	0.083	0.048	0.083	0.186	0.106		
Cygb	0.328	0.244	0.000	0.056	0.202	0.096	0.079	0.013	0.006	0.013	0.017	0.543	0.359	0.100	
Cabr	0.035	0.025	0.259	0.148	0.121	0.067	0.067	0.186	0.253	0.186	0.236	0.106	0.065	0.060	0.259

*** Abbreviation of Taxon (Annexure II)**

Table 6- Maximum composite likelihood estimate of the pattern of nucleotide substitution of *Luffa* and *Momordica*

	A	T	C	G
A	-	*4.1*	*8.13*	**11.48**
T	*4.19*	-	**23.29**	*6.78*
C	*4.19*	**11.76**	-	*6.78*
G	**7.08**	*4.1*	*8.13*	-

Table 7-Codon-based test of neutrality for analysis between sequences of *Luffa* and *Momordica*

* ↓																
Ciss		-1.214	-1.184	-1.536	-1.206	-0.981	-0.822	-1.399	-1.157	-1.399	-1.038	-1.735	-1.199	-1.215	-1.184	-1.925
Coch	0.227		-0.017	-1.312	-0.364	-0.870	-0.693	-0.248	-0.409	-0.248	-0.275	-0.787	-1.621	-0.353	-0.017	-1.203
Acut	0.239	0.986		-1.913	-1.061	-1.564	-1.448	1.449	-0.354	1.449	-0.220	-0.292	-1.162	-0.979	0.000	-1.249
Bals	0.127	0.192	0.058		-1.200	-1.295	-1.138	-2.060	-2.070	-2.060	-1.928	-1.760	-0.712	-2.035	-1.913	-1.624
Foet	0.230	0.717	0.291	0.233		-0.345	-0.172	-1.285	-1.248	-1.285	-1.106	-1.009	-0.590	-1.202	-1.061	-0.739
Char	0.329	0.386	0.121	0.198	0.731		1.044	-1.812	-1.465	-1.812	-1.355	-1.203	-1.805	-1.255	-1.564	-1.822
Chma	0.413	0.490	0.150	0.258	0.864	0.298		-1.692	-1.336	-1.692	-1.226	-1.049	-1.706	-1.088	-1.448	-1.669
Cyli	0.164	0.805	0.150	0.042	0.201	0.073	0.093		-0.352	0.000	-0.216	-0.525	-1.151	-1.110	1.449	-1.454
Tric	0.249	0.683	0.724	0.041	0.214	0.145	0.184	0.725		-0.352	1.754	-1.460	-1.388	-1.435	-0.354	-1.257
Aegy	0.164	0.805	0.150	0.042	0.201	0.073	0.093	1.000	0.725		-0.216	-0.525	-1.151	-1.110	1.449	-1.454
Lepi	0.301	0.784	0.826	0.056	0.271	0.178	0.222	0.830	0.082	0.830		-1.328	-1.708	-1.301	-0.220	-1.404
Dioi	0.085	0.433	0.770	0.081	0.315	0.231	0.296	0.600	0.147	0.600	0.187		-0.750	-1.284	-0.292	-1.668
Tube	0.233	0.108	0.248	0.478	0.556	0.074	0.091	0.252	0.168	0.252	0.090	0.454		-1.621	-1.162	-1.212
Angu	0.227	0.725	0.330	0.044	0.232	0.212	0.279	0.269	0.154	0.269	0.196	0.202	0.108		-0.979	-1.417
Cygb	0.239	0.986	1.000	0.058	0.291	0.121	0.150	0.150	0.724	0.150	0.826	0.770	0.248	0.330		-1.249
Cabr	0.057	0.231	0.214	0.107	0.461	0.071	0.098	0.149	0.211	0.149	0.163	0.098	0.228	0.159	0.214	

*** Abbreviation of Taxon (Annexure II)**

Table 8- Fisher's exact test of neutrality for sequence pairs of *Luffa* and *Momordica*

* ↓															
Ciss															
Coch	1.000														
Acut	1.000	1.000													
Bals	1.000	1.000	1.000												
Foet	1.000	1.000	1.000	1.000											
Char	1.000	1.000	1.000	1.000	1.000										
Chma	1.000	1.000	1.000	1.000	1.000	0.712									
Cyli	1.000	1.000	0.513	1.000	1.000	1.000	1.000								
Tric	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000							
Aegy	1.000	1.000	0.513	1.000	1.000	1.000	1.000	1.000	1.000						
Lepi	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.366	1.000					
Dioi	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000				
Tube	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000			
Angu	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
Cygb	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.513	1.000	0.513	1.000	1.000	1.000	1.000	
Cabr	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

*** Abbreviation of Taxon (Annexure II)**

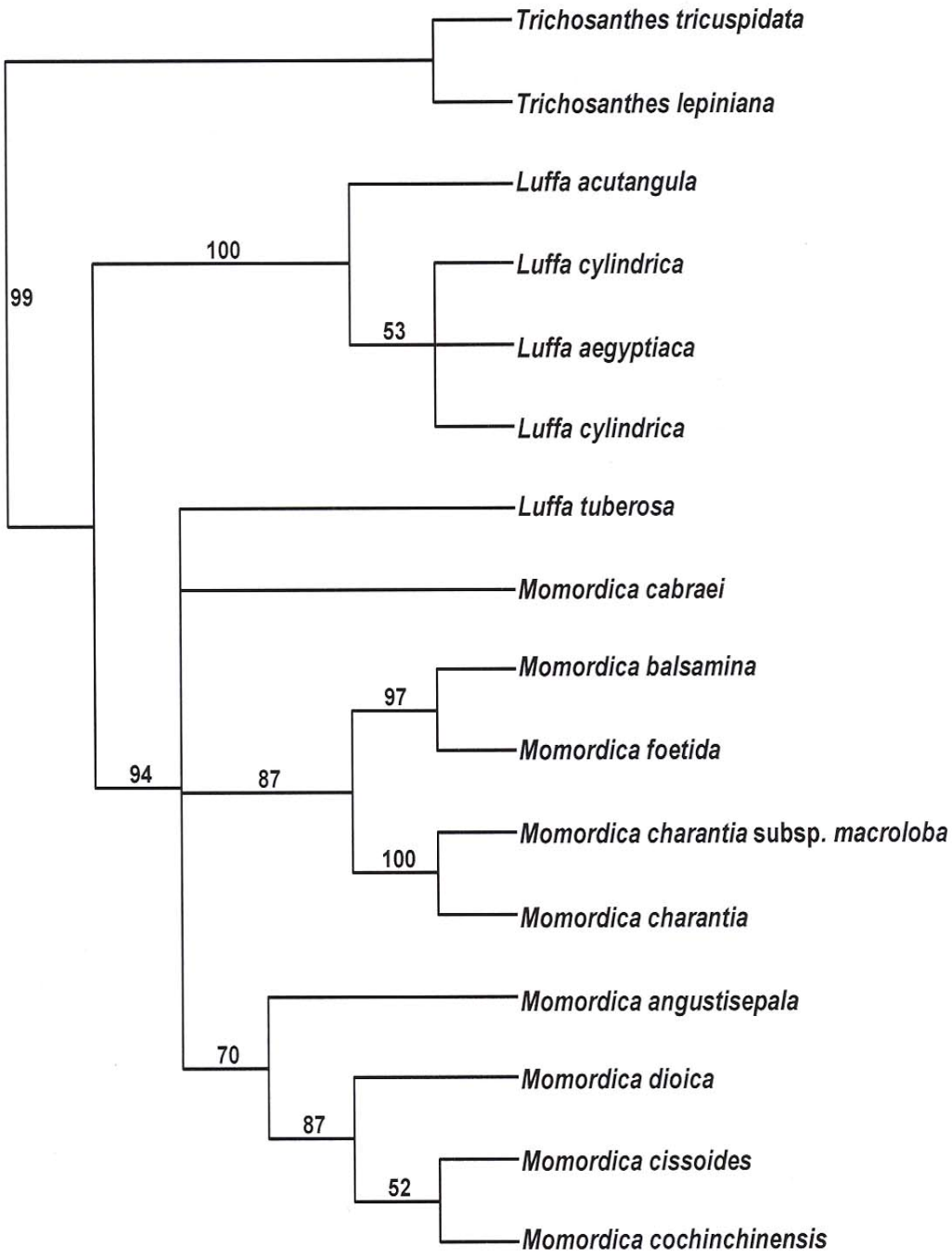

Fig. 1- The bootstrap strict consensus tree of eight maximally parsimonious trees based on the ITS sequences of nuclear ribosomal DNA data set with gaps being treated as missing data (486 steps, a consistency index (CI) of 0.7366, a homoplasy index (HI) of 0.2634, rescaled consistency index (RC) of 0.5195 and a retention index (RI) of 0.7051. the bootstrap values greater than 50% in 1000 bootstrap replicates are shown above lines

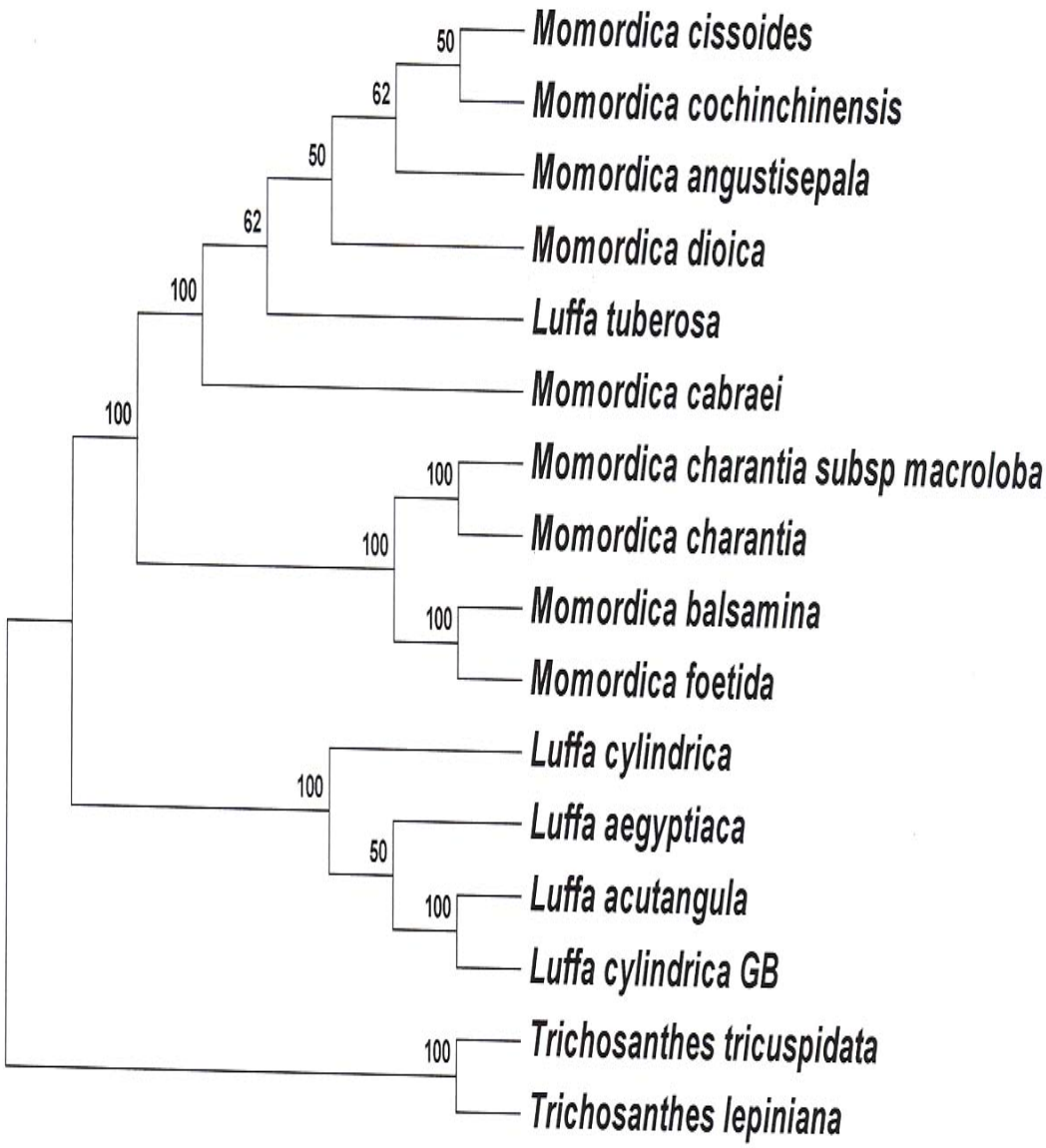

Fig. 2- 50% Majority rule tree inferred from internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed in MEGA4 after multiple alignment in ClustalX.

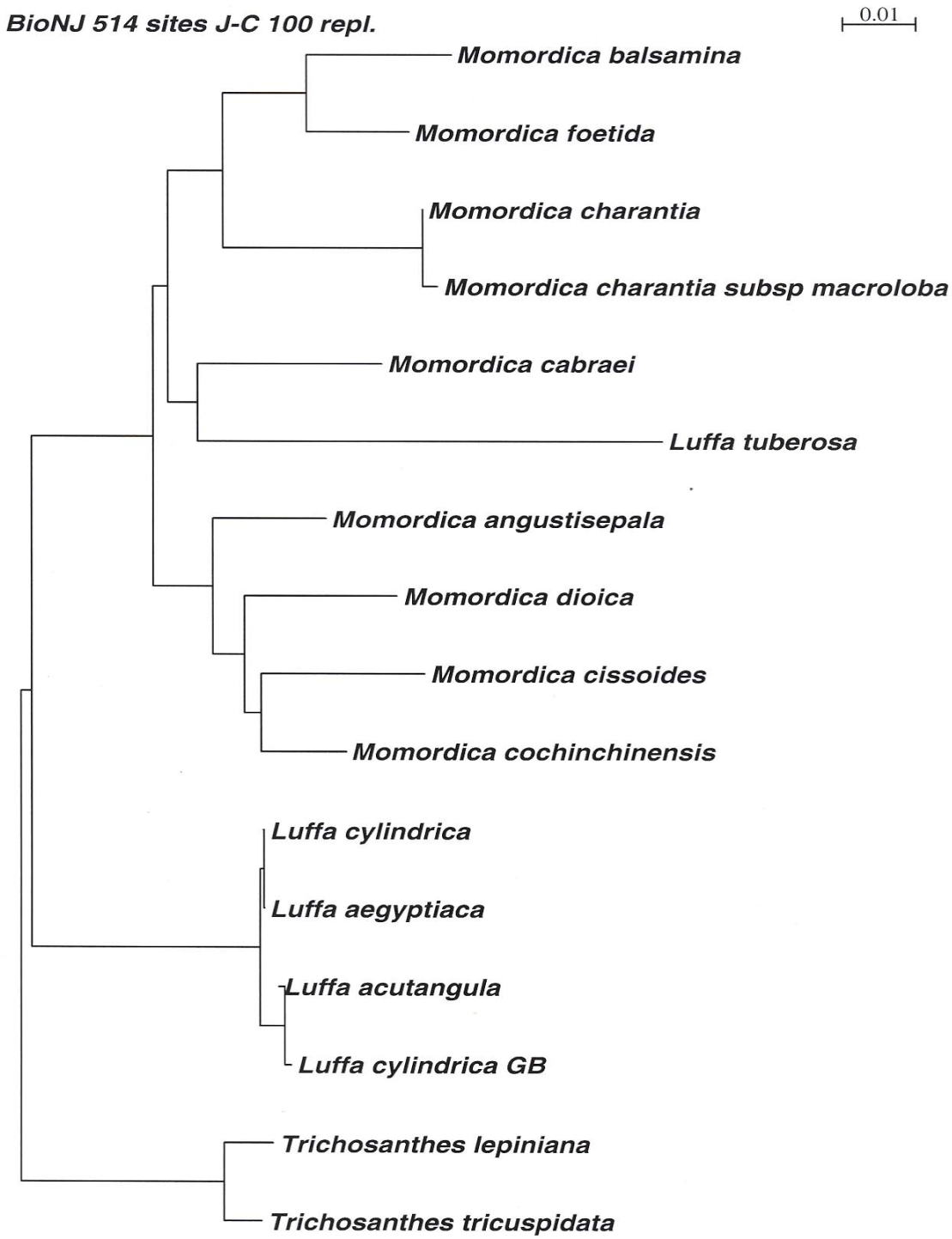

Fig. 3- BioNJ tree inferred from internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed using SeaView after multiple alignment in MUSCLE. The scale bar indicates relative length of the branch

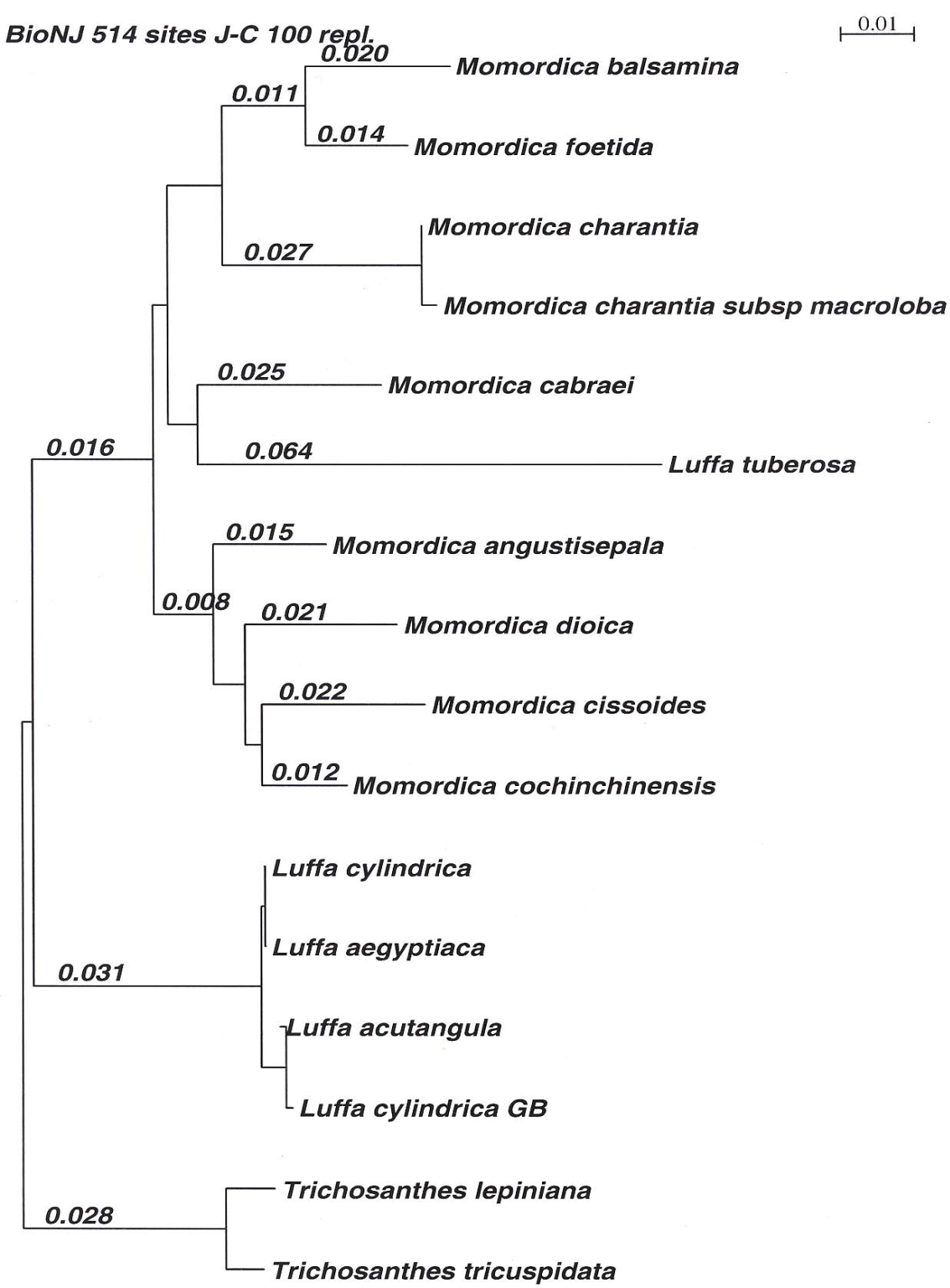

Fig. 4- BioNJ tree inferred from internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed using SeaView after multiple alignment in MUSCLE. The number above the line indicates branch length

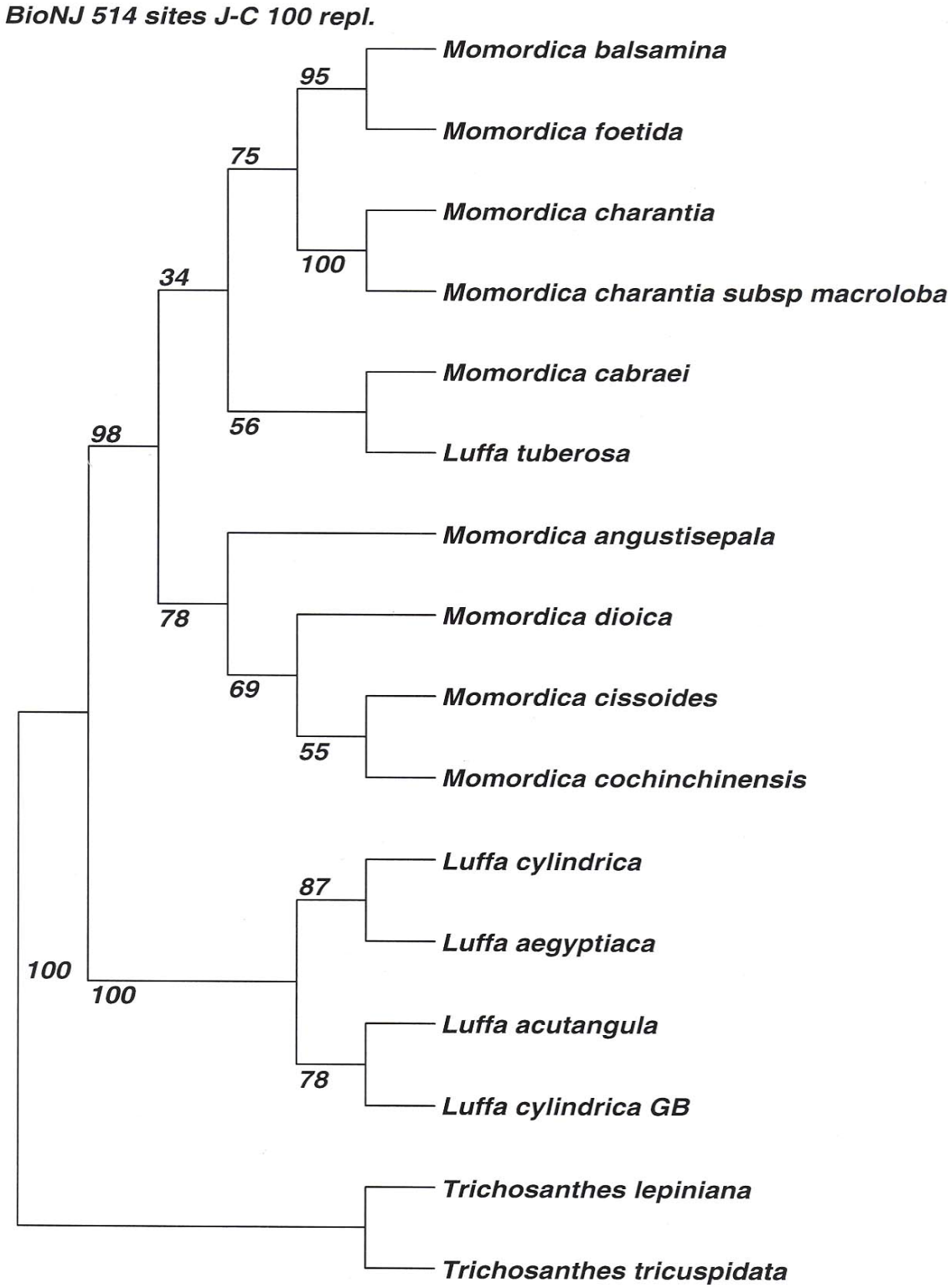

Fig. 5-BioNJ tree inferred from internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed using SeaView after multiple alignment in MUSCLE. The number at nodes indicates bootstrap support

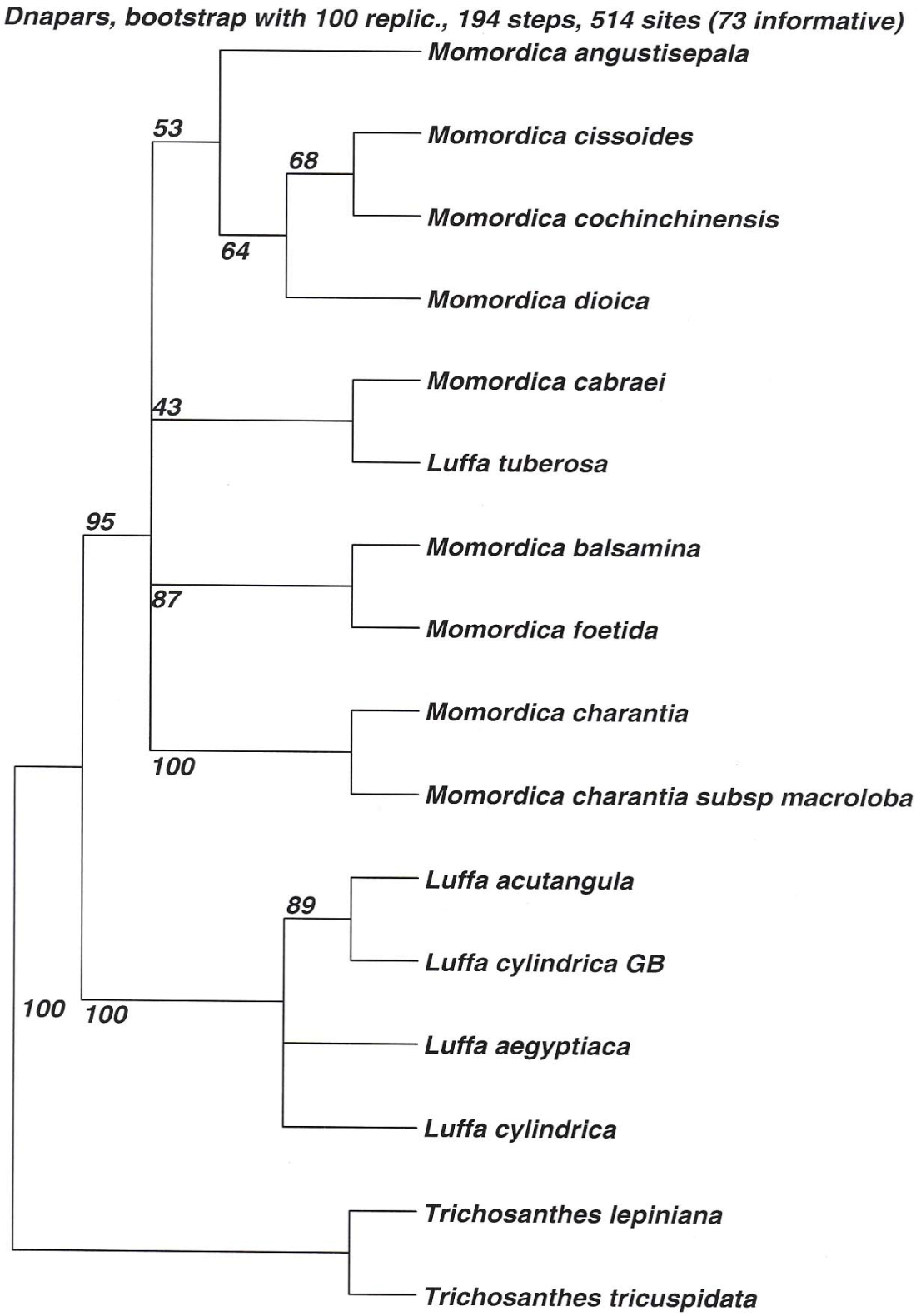

Fig. 6- Bootstrap strict consensus tree based on internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed using Maximum Parsimony method in SeaView after multiple alignment in MUSCLE. The number at nodes indicates bootstrap support in 100 bootstrap replicates

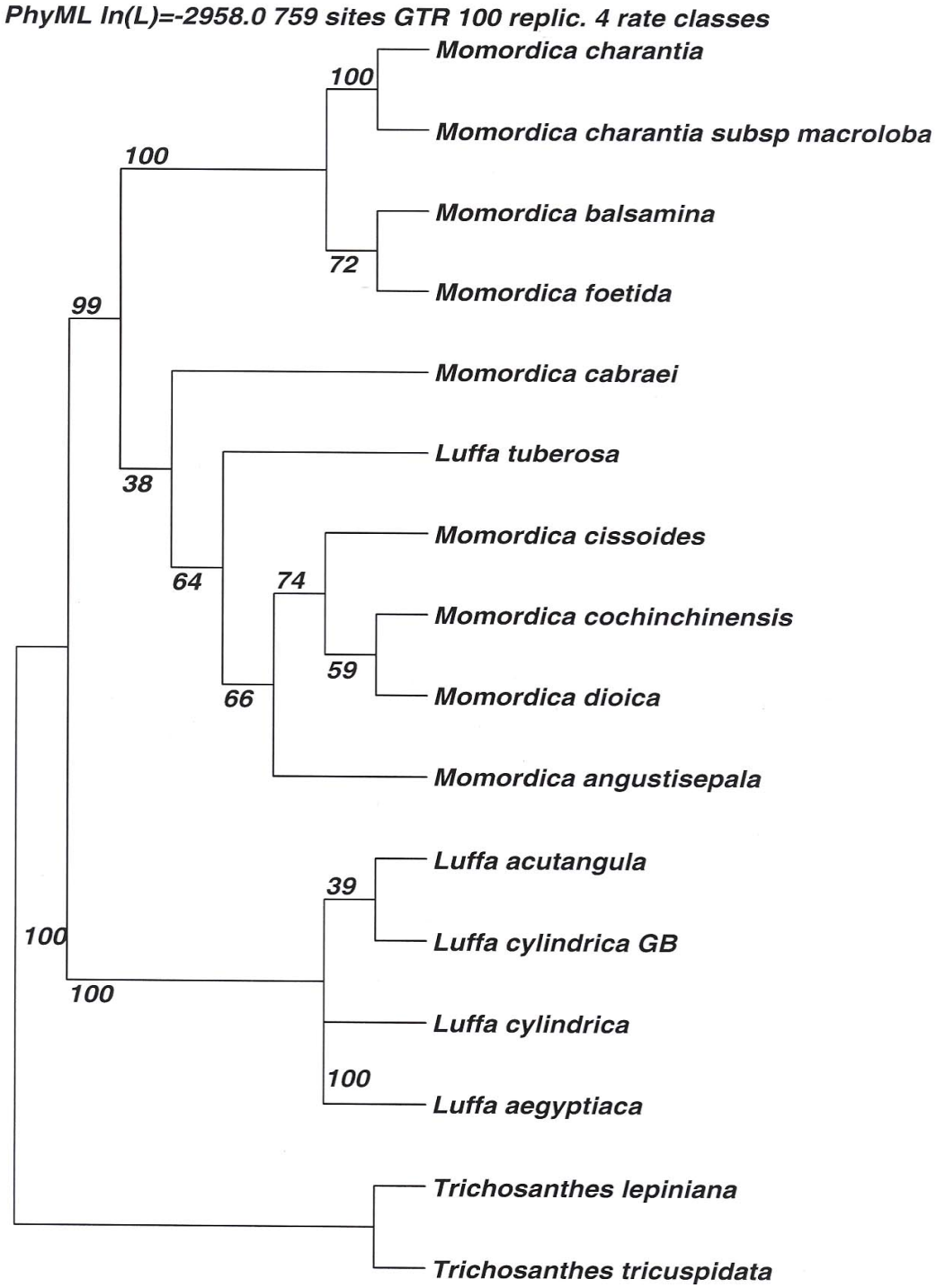

Fig. 7- Bootstrap strict consensus tree based on internal transcribed spacer region of nuclear ribosomal DNA. The tree constructed using Maximum Likelihood method in SeaView after multiple alignment in MUSCLE. Number at node indicates Bootstrap values 100 bootstrap replicates

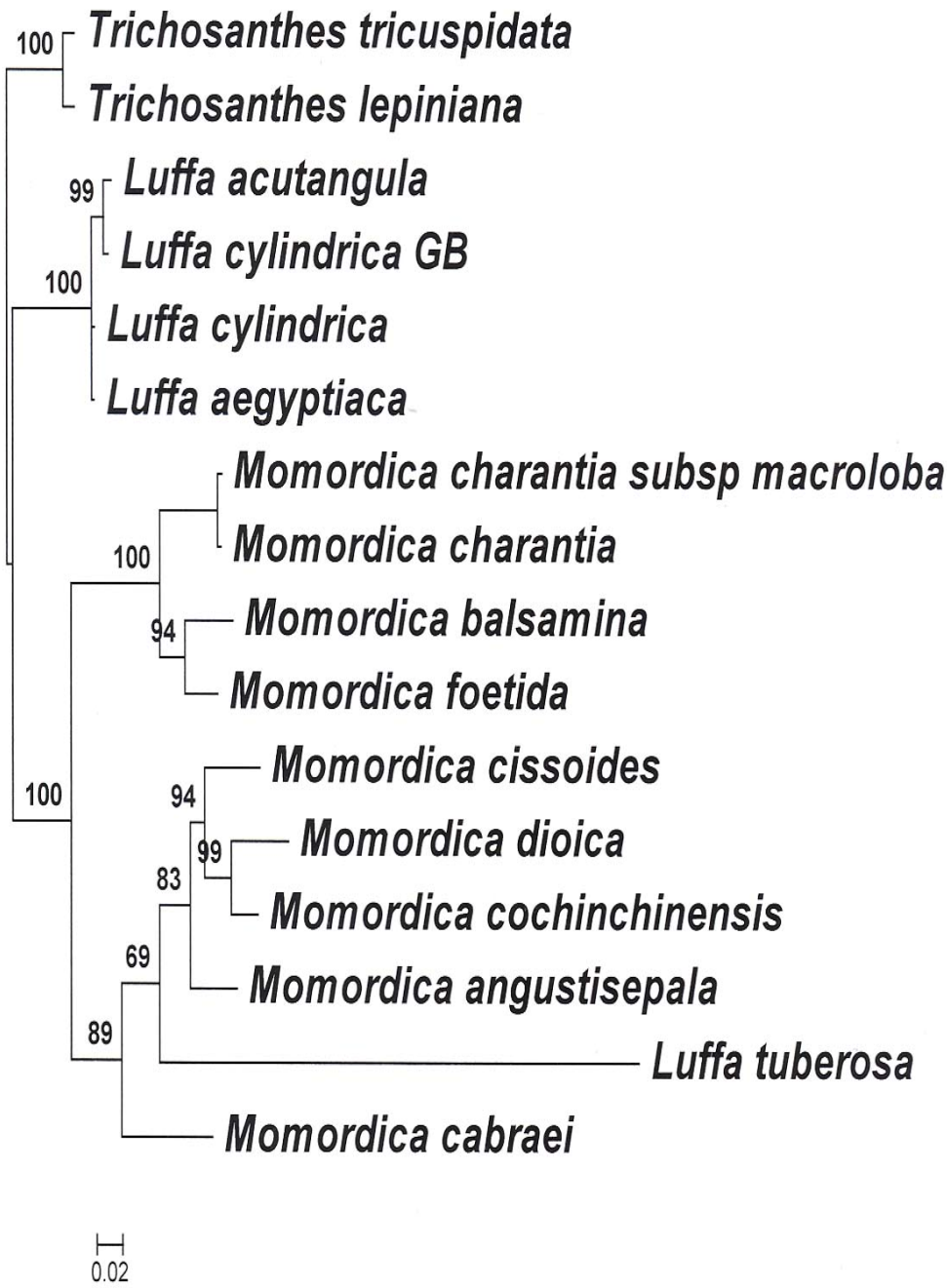

Fig. 8-Bayesian phylogeny with bootstrap support based on analysis of ITS sequences of nrDNA.

Annexure II

*** Abbreviation of Taxon**

Taxon	Abbreviation
Momordica cabraei	**Cabr**
Momordica angustisepala	**Angu**
Momordica cissoides	**Ciss**
Momordica balsamina	**Bals**
Momordica foetida	**Foet**
Momordica charantia	**Char**
Momordica charantia subsp. *macroloba*	**Chma**
Momordica cochinchinensis	**Coch**
Momordica dioica	**Dioi**
Luffa acutangula	**Acut**
Luffa cylindrica	**Cyli**
Luffa aegyptiaca	**Aegy**
Luffa cylindrica	**Cygb**
Luffa tuberosa	**Tube**
Trichosanthes lepiniana	**Lepi**
Trichosanthes tricuspidata	**Tric**