Cladistic analyses of a few members of *Cucurbitaceae* using *rbc*L nucleotide and amino acid sequences

Uma Reddy B.
Department of Bioinformatics, Gulbarga University, Gulbarga, India, E. mail: drumareddy11@yahoo.co.in, Ph: +91-9449329266

Abstract-The *rbc*L gene and protein sequences have been used in addressing systematic questions among the few selected members of the family *Cucurbitaceae* has been investigated. In order to elucidate the systematic positions, a set of chloroplast-*rbc*L nucleotide sequences (from 42 taxa of 7 genera) and aminoacid sequences (from 52 taxa of 10 genera) were withdrawn from GenBank and GenPept databases, respectively. The evolutionary distance was inferred from these sequences by employing Bootstrap method of UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and MP (Maximum Parsimony) using MEGA (Molecular Evolutionary Genetic Analysis) software. From the separate analysis produced almost similar although not identical results, no strongly supported incongruent results. The members of the genus *Austrobryonia* showed strictly monophyletic, *Trichosanthes, Luffa, Momordica* and *Coccinia* are found to be paraphyletic. But the members of the genus *Cucumis* are distributed throughout these hiraeoid clades, confirming the polyphyly of this large genus observed in both the family trees. From the results, it is also clear that, the chloroplast-*rbc*L gene and aminoacid sequences resolved the relationships, as well as provided a good indication of major supra-generic groupings among the selected members of the family *Cucurbitaceae*. These results provide the necessary frame work and explicit phylogenetic hypotheses from which further reversionary and other systematic studies can proceed.
Key Words: *rbc*L-nucleotide, *rbc*L-aminoacid, tribe, polyphyletic, *Cucurbitaceae*, Phylogenetic relationships, MEGA

INTRODUCTION

The use of nucleotide sequence comparisons for estimating phylogenetic relationships has attracted wide spread attention among botanists studying a variety of plant groups [1]. As pointed out by Palmer *et al.,* (1988) [2], nucleotide sequences have the advantages of being rapidly produced and easily assessed for homology. This is particularly true for *rbc*L, in which there is virtually no length variation and it has proven utility in reconstructing phylogenetic relationships at the family level [3]. *Cucurbitaceae* is a plant family, comprises 960 species of 125 extant genera. Commonly known as melons, gourds or cucurbits and includes crops like Cucumbers, Squashes, Luffas, Melons and Water melons. The family is predominantly distributed around the tropics. Most of the plants in this family are annual vines. Many species have large yellow or white flowers. The stems are hairy and pentangular; tendrils are present at 90° to the leaf petioles. Leaves are ex-stipulate, alternate, simple, palmately lobed or palmately compound. The flowers are unisexual with male and female flowers on different plants (dioecious) or on the same plant (Monoecious). The female flowers have inferior ovaries. The fruit is often a kind of berry called a pepo [4-5]. The present study was aimed to address the inter-generic, tribal and subfamilial relationships among the few selected members of *Cucurbitaceae*, a set of *rbc*L nucleotide and amino acid sequences were chosen. The current studies also stressed to examine the degree of congruence between these chloroplast-*rbc*L sequences and compare their patterns of molecular evolution and assess their utility in estimating phylogeny.

MATERIALS AND METHODS

The large subunit of chloroplast-*rbc*L nucleotide sequences (from 42 taxa) and amino acid sequences (from 52 taxa) of *Cucurbitaceae* were withdrawn from the GenBank and GenPept databanks respectively, are presented in the table-1 and 2. The multiple sequence alignment was performed by GeneBee-ClustalW service to findout the variability in different regions of the sequences in case of nucleotide sequences and also the conserved / motif regions in amino acid sequences. The phylogenetic tree was constructed on the basis of *rbc*L nucleotide sequences ranging 1191 to 1423 from 42 taxa of 7 genera and amino acid sequences ranging 421 to 476 from 52 taxa of 10 genera, based on UPGMA and MP methods using MEGA software [6]. Tree statistics included the consistency index [7], retention index (RI) [8] and rescaled consistency index (RCI) for all sites. Branch lengths and at the level of support for branches of the phylogenetic tree was evaluated with bootstrap analysis [9] to verify the strength of the branches based on 100 replications using branch and bound search. Bootstrap percentages are also described as high (85 – 100%), moderate (75 -84%) or low (50 – 74%) [10]. The number of nucleotide substitutions per site was estimated by Kimura's two parameter method [11].

RESULTS AND DISCUSSION

From the separate analysis it was noticed that, almost similar although not identical results, no strongly supported incongruent results. Based on the nucleotide family tree it is clear that, the members of the genus *Austrobryonia* are showed

strictly monophyletic. Perhaps, the members of genera *Trichosanthes, Luffa, Momordica* and *Coccinia* are paraphyletic, but the members of the genus *Cucumis* in particularly, are distributed throughout these hiraeoid clades, confirming the polyphyly of this large genus observed in both the family trees. It is evidenced from the Figure-1 and Figure -2. Phylogenetic tree was designed to identify the ideal regions that could be used for defining inter and intra-generic relationships. Phylogenetic tree from the nucleotide sequences of 42 taxa of 7 genera belongs to family *Cucurbitaceae* were inferred using UPGMA and MP methods. The total length of the tree was found to be 11884. The consistency index is 0.292, the retention index is 0.655 and the rescaled consistency index for all sites is 0.191. All positions containing gaps and missing data were eliminated from the dataset. A total of 1165 residues were used to construct a phylogenetic tree. The family tree obtained from the UPGMA method was split into eight major clades. Where as in case of phylogenetic analysis of the *rbc*L protein sequences from 52 taxa of 10 genera belonging to family *Cucurbitaceae* was also inferred by UPGMA method and was split into six well supported major clades. The total length of the tree was found to be 5349. The consistency index is 0.753, the retention index is 0.870 and the rescaled consistency index for all sites is 0.655. All positions containing gaps and missing data were eliminated from the dataset. A total of 381 residues were used to construct a phylogenetic tree. The results of the Cladistic analysis of the *rbc*L nucleotide sequences based on UPGMA method was clearly displayed that, the clade-II and clade-VIII consists of mainly the members of the genus *Cucumis* and sharing the common features like, tendrils slender, simple, flowers monoecious, male flowers fascicled, calyx and corolla are companulate, stamens-3, free, reticulated pollens, inserted on corolla tube, anthers oblong or ovate, female flowers fascicled or solitary, ovary cylindric, style short, stigmas 3-5, ovules numerous, horizontal, fruits polymorphic, fleshy, indehiscent, smooth or verrucose, seeds numerous, compressed, emarginated. Subclade-I of clade-I consists of dominatingly the members of the genus *Trichosanthes* showing highest level of sequence identity, made up of strictly 1356 nucleotide residues and sharing the common morphological features like, pollen striate, smooth or verrucose, flowers with larger stamens, united with triplicate thecae. The members of the clade –II and III comprises the members of tribes, such as Melothrieae and Trichosantheae sharing common characteristics like, receptacle tube relatively long, alike in male and female flowers. The members of the Clade-VI strictly belongs to tribe Benincaseae and sharing common features like, ovules many, horizontal, pollen reticulate triporate, petals fringed or with ventral scales [12]. The clades –VII mainly consists of the members of the genus *Austrobryonia,* which exhibiting highest level of sequence similarity and strictly made up of 1355 nucleotide residues. In addition to this, they also sharing the common morphological characteristics like, monoecious scabrid herbs, stems annual from a perrennating root stock, tendrils simple, leaves petiolate, flowers pedicillate, small, 3 – 15mm diameter, corolla yellow – green or yellow, seeds ovate, compressed margins, pollen tricolpate – oblate. The key differences among the members of the genus *Austrobryonia* such as female flowers in fascicles of 2 – 5, mostly co-axillary with 1- 3 males, fruit ellipsoidal, 10 -14 mm long to *A. micrantha*. Where as, female flowers solitary, rarely 2 per fascicle, sometimes co-axillary with males, stigmatic lobes linear, fruit ellipsoidal, 18 - 35 mm long in case of *A. centralis.* Stigmatic lobes capitate, fruit sub-globose, fruit 18-25mm diameter, endemic in inland Northern Australia to *A. argillicola.* However, fruit 12 -18 mm diameter, endemic to the Pilbara region of Western Australia in *A. pilbarensis* [13].The results of the Cladistic analysis of the *rbc*L amino acid sequences based on UPGMA method strongly suggests that, the clade-I consists of mainly the members of the genus *Cucumis* belongs tribe Melothtrieae and *Coccinia adoensis* belongs to tribe Benincaseae showing the highest level of sequence similarity and strictly made up of 476-475 amino acid residues. Similarly, the members of the subclade-II of clade-III made up of strictly 452 amino acids, subclade-III of clade-III possess 461 amino acids, clade-V consists of 465, clade-VI showing strictly 457 amino acids and exhibiting the highest level of sequence similarity among the members and are clustered based on the total number of amino acids and sequence identity within these clades. In assessing the results presented here, are compared with the most recent system of classification of the family *Cucurbitaceae* as given by Jeffrey (2005). This new technology affords the opportunity to apply methods of molecular analysis, especially DNA sequencing to Plant systematics [5, 12].

CONCLUSION

In conclusion, the chloroplast-*rbc*L sequence analyses clarified the phylogenetic relationships among the selected genera of *Cucurbitaceae*. The monophyly of *Austrobryonia* is well supported because of its highest level of sequence similarity. The polyphyletic condition of the genus *Cucumis* is supported by the higher rate of the sequence divergence that is significant variation in the total number of nucleotide and aminoacid residues among the different members of the *Cucumis* from 1289 – 1416 nucleotides and 421 – 476 aminoacids. These results provide the necessary frame work and explicit

phylogenetic hypotheses from which further reversionary and other systematic studies can proceed.

References

[1] Clegg M. and Zurawski. (1992) Chloroplast DNA and the study of Plant Phylogeny present status and Future prospects. *In Soltis, P.; Soltis D. and Doyle J. (eds.), Molecular Systematics of Plants.* Chapman and Hall, New York, NY.

[2] Palmer J., Jansen R., Michaels H., Chase M. and Manhart J. (1988) *Ann Mo Bot Gard,* 75, 1180-1206.

[3] Cameron K.M., Chase M.W., Whitten W.M., Kores P.J., Jarrell D.C., Albert V.A., Yukawa T., Hillis H.G. and Goldman D.H. (1999). *Am J Bot,* 86, 208 – 224.

[4] Susanne Renner S., Hanno Schaefer and Alexander Kocyan (2007) *BMC Evol Biol,* 7, 58.

[5] Jeffrey C. (2005) *Bot Zhurn,* 90, 332 – 335.

[14]

[6] Sudhir Kumar., Masatoshi Nei., Joel Dudley and Koichiro Tamura. (2008) *Brief in Bioinf*, 9(4), 299-308.

[7] Kluge A.G. and Farris J.S. (1969) *Syst Bot*, 18, 1-22.

[8] Farris J.S. (1989) *Cladistics*, 5, 417 – 419.

[9] Felsenstein J. (1985) *Evolution*, 39, 783 - 791.

[10] Kenneth Cameron M., Mark Chase W., Mark Whitten W., Paul Kores J., David Jarrell C., Victor Albert A., Tomosia Yukawa., Harold Hillis G. and Douglas Goldman H. (1999) *Am J Bot*, 86(2), 208 – 224.

[11] Kimura M. (1980) *J Mol Evol*, 16, 111-120.

[12] Vasishta P.C. (2001) *Taxonomy of Angiosperms*. Published by *Sharma, G.S. Proprietor, R. Chand and Company*, 1 – Ansari Road, Daryaganj, New Delhi, 32.

[13] Hanno Schaefer., Ian R.H., Telford and Susanne Renner S. (2008) *Syst Bot*, 33(1), 125-132.

Table 1: List of the chloroplast- *rbc*L nucleotide sequences used to infer the phylogenetic relationships within the family *Cucurbitaceae*

Sr. No	Accession Number (GenBank Databank)	Number of base pairs	Taxa used	References
01	EU155605	1356	*Trichosanthes schlechteri*	Schaefer *et al.,* (2008)
02	EU155604	1356	*Trichosanthes pentaphylla*	Schaefer *et al.,* (2008)
03	EU155603	1356	*Trichosanthes cucumerina*	Schaefer *et al.,* (2008)
04	EU155602	1356	*Trichosanthes bracteata*	Schaefer *et al.,* (2008)
05	EU037005	1393	*Trichosanthes villosa*	Schaefer *et al.,* (2008)
06	EU037000	1375	*Trichosanthes amara*	Schaefer *et al.,* (2008)
07	DQ535857	1423	*Trichosanthes reticulinervis*	Kocyan *et al.,* (2007)
08	DQ535855	1410	*Trichosanthes kirilowii*	Kocyan *et al.,* (2007)
09	DQ535819	1357	*Trichosanthes pubera*	Kocyan *et al.,* (2007)
10	AF206755	1402	*Cucumis sativus*	Soltis *et al.,* direct submission on (19-NOV-1999)
11	DQ535803	1381	*Cucumis zeyherii*	Kocyan *et al.,* (2007)
12	DQ535802	1406	*Cucumis sagittatus*	Kocyan *et al.,* (2007)
13	DQ535801	1416	*Cucumis metuliferus*	Kocyan *et al.,* (2007)
14	DQ535800	1289	*Cucumis melo*	Kocyan *et al.,* (2007)
15	DQ535799	1405	*Cucumis hirsutus*	Kocyan *et al.,* (2007)
16	DQ785838	1374	*Cucumis sacleuxii*	Renner *et al.,* (2007)
17	DQ785837	1374	*Cucumis prophetarum*	Renner *et al.,* (2007)
18	DQ785836	1394	*Cucumis myriocarpus*	Renner *et al.,* (2007)
19	DQ785832	1397	*Cucumis hystrix*	Renner *et al.,* (2007)
20	DQ785831	1374	*Cucumis humifructus*	Renner *et al.,* (2007)
21	DQ785836	1394	*Cucumis heptadactylis*	Renner *et al.,* (2007)
22	DQ785829	1374	*Cucumis ficifolius*	Renner *et al.,* (2007)
23	DQ785828	1392	*Cucumis dipsaceus*	Renner *et al.,* (2007)
24	DQ785827	1374	*Cucumis anguria*	Renner *et al.,* (2007)
25	DQ535826	1423	*Luffa acutangula*	Kocyan *et al.,* (2007)
26	DQ535827	1387	*Luffa aegyptiaca*	Kocyan *et al.,* (2007)
27	EU436385	1356	*Luffa graveolens*	Schaefer *et al.,*(2009)
28	L21941	1402	*Luffa quinquefida*	Swensen *et al.*,(1994)
29	DQ535798	1380	*Cucumella bryonifolia*	Kocyan *et al.,* (2007)
30	DQ785826	1407	*Cucumella aspera*	Renner *et al.,* (2007)
31	DQ535829	1391	*Momordica foetida*	Kocyan *et al.,* (2007)
32	EF487554	1355	*Momordica cochinensis*	Schaefer *et al.,* (2008)
33	DQ535960	1264	*Momordica charantia*	Irwin *et al.,* (2008)
34	DQ535759	1191	*Momordica calantha*	Kocyan *et al.,* (2007)
35	DQ535793	1391	*Coccinia rehmannii*	Kocyan *et al.,* (2007)
36	DQ535792	1452	*Coccinia grandis*	Kocyan *et al.,* (2007)
37	AY968520	1391	*Coccinia sessilifolia*	Zhang *et al.,* (2006)
39	EF487553	1355	*Austrobryonia pilbarensis*	Schaefer *et al.,* (2008)
40	EF487552	1355	*Austrobryonia micrantha*	Schaefer *et al.,* (2008)
41	EF487550	1355	*Austrobryonia centralis*	Schaefer *et al.,* (2008)
42	EF487549	1355	*Austrobryonia argillicola*	Schaefer *et al.,* (2008)

Table 2: List of the chloroplast-*rbc*L aminoacid sequences used to infer the phylogenetic relationships within the family *Cucurbitaceae*.

Sr. No	Accession Number (GenPept Databank)	Number of amino acids	Taxa used	References
01	ABW08086	464	*Trichosanthes villosa*	Schaefer *et al.,* (2008)
02	ABW08085	458	*Trichosanthes amara*	Schaefer *et al.,* (2008)
03	ABG24988	446	*Trichosanthes pubera*	Kocyan *et al* (2007)
04	ABX79799	452	*Trichosanthes schlechteri*	Schaefer *et al.,* (2008)
05	ABX79798	452	*Trichosanthes pentphylla*	Schaefer *et al.,* (2008)
06	ABX79797	452	*Trichosanthes cucumerina*	Schaefer *et al.,* (2008)
07	ABX79796	452	*Trichosanthes bracteata*	Schaefer *et al.,* (2008)
08	ABG25026	465	*Trichosanthes reticulinervis*	Kocyan *et al*., (2007)
09	ABG25025	457	*Trichosanthes ovigera*	Kocyan *et al*., (2007)
10	ABG25024	461	*Trichosanthes kirilowii*	Kocyan *et al*., (2007)
11	ABG24943	444	*Trichosanthes amara* (2)	Kocyan *et al*., (2007)
12	ABG24972	454	*Cucumis zeyherii*	Kocyan *et al*., (2007)
13	YP_247607	476	*Cucumis sativus*	Plader *et al.,* (2007)
14	ABG24970	465	*Cucumis metuliferus*	Kocyan *et al*., (2007)
15	ABG24971	466	*Cucumis sagittatus*	Kocyan *et al*., (2007)
16	ABG24969	421	*Cucumis melo*	A Kocyan *et al.,* (2007)
17	ABG24968	461	*Cucumis hirsutus*	Kocyan *et al*., (2007)
18	AAP88011	475	*Cucumis anguria*	Swensen and Clement, Direct Submission on (07-AUG-2002) Biology, Ithaca College, 953 Danby Rd, Ithaca, NY 14850, USA
19	ABH07710	457	*Cucumis heptadactylis*	Renner *et al.,* (2007)
20	ABH07718	457	*Cucumis sacleuxii*	Renner *et al.,* (2007)
21	ABH07717	457	*Cucumis prophetarum*	Renner *et al.,* (2007)
22	ABH07716	463	*Cucumis myriocarpus*	Renner *et al.,* (2007)
23	ABH07712	464	*Cucumis hystrix*	Renner *et al.,* (2007)
24	ABH07711	457	*Cucumis humifructus*	Renner *et al.,* (2007)
25	ABH07709	457	*Cucumis ficifolius*	Renner *et al.,* (2007)
26	ABH7708	463	*Cucumis dipsaceus*	Renner *et al.,* (2007)
27	ABS89014	452	*Austrobryonia pilbarensis*	Schaefer *et al*., (2008)
28	ABS89013	452	*Austrobryonia micrantha*	Schaefer *et al*., (2008)
29	ABS89011	452	*Austrobryonia centralis*	Schaefer *et al*., (2008)
30	ABS89010	452	*Austrobryonia argillicola*	Schaefer *et al*., (2008)
31	ABG24998	457	*Momordica foetida*	Kocyan *et al*.,(2007)
32	ABS89015	452	*Momordica cochinchinensis*	Schaefer *et al*., (2008)
33	ABG24929	421	*Momordica charantia*	Kocyan *et al*. ,(2007)
34	ABG24928	397	*Momordica calantha*	Kocyan *et al*.,(2007)
35	ABG24960	455	*Citrullus colocynthis*	Kocyan *et al*.,(2007)
36	ABG24914	439	*Citrullus lanatus*	Kocyan *et al*.,(2007)
37	ABG24967	457	*Cucumella bryoniifolia*	Kocyan *et al*.,(2007)
38	AAW56427	475	*Cucumella bryoniifolia* (2)	Renner *et al.,* (2007)
39	ABH07706	466	*Cucumella aspera*	Renner *et al.,* (2007)
40	AAA84184	466	*Cucurbita pepo*	Swensen *et al.,*(1994)
41	ABG24973	448	*Cucurbita ficifolia*	Kocyan *et al*.,(2007)
42	ABG24962	457	*Coccinia rehmannii*	Kocyan *et al*.,(2007)
43	ABG24961	461	*Coccinia grandis*	Kocyan *et al*.,(2007)
44	AAY44162	464	*Coccinia sessilifolia*	Zhang *et al*., (2006)
45	AAP88013	475	*Coccinia adoensis*	Swensen and Clement, Direct submission on (07-AUG-2002) Biology, Ithaca College, 953 Danby Rd, Ithaca, NY 14850, USA
46	ABG24995	465	*Luffa acutangula*	Kocyan *et al*.,(2007)
47	ACB58283	452	*Luffa graveolens*	Schaefer *et al.,* (2009)
48	AAA84358	466	*Luffa quinquefida*	Swensen *et al*.,(1994)
49	ABG24996	457	*Luffa aegyptiaca*	Kocyan *et al*.,(2007)
50	ACB58292	452	*Sicyos hillebrandii*	Schaefer *et al*.,(2009)
51	ABG25017	457	*Sicyos baderoa*	Kocyan *et al*.,(2007)
52	ABG25016	452	*Sicyos angulatus*	Kocyan *et al*.,(2007)

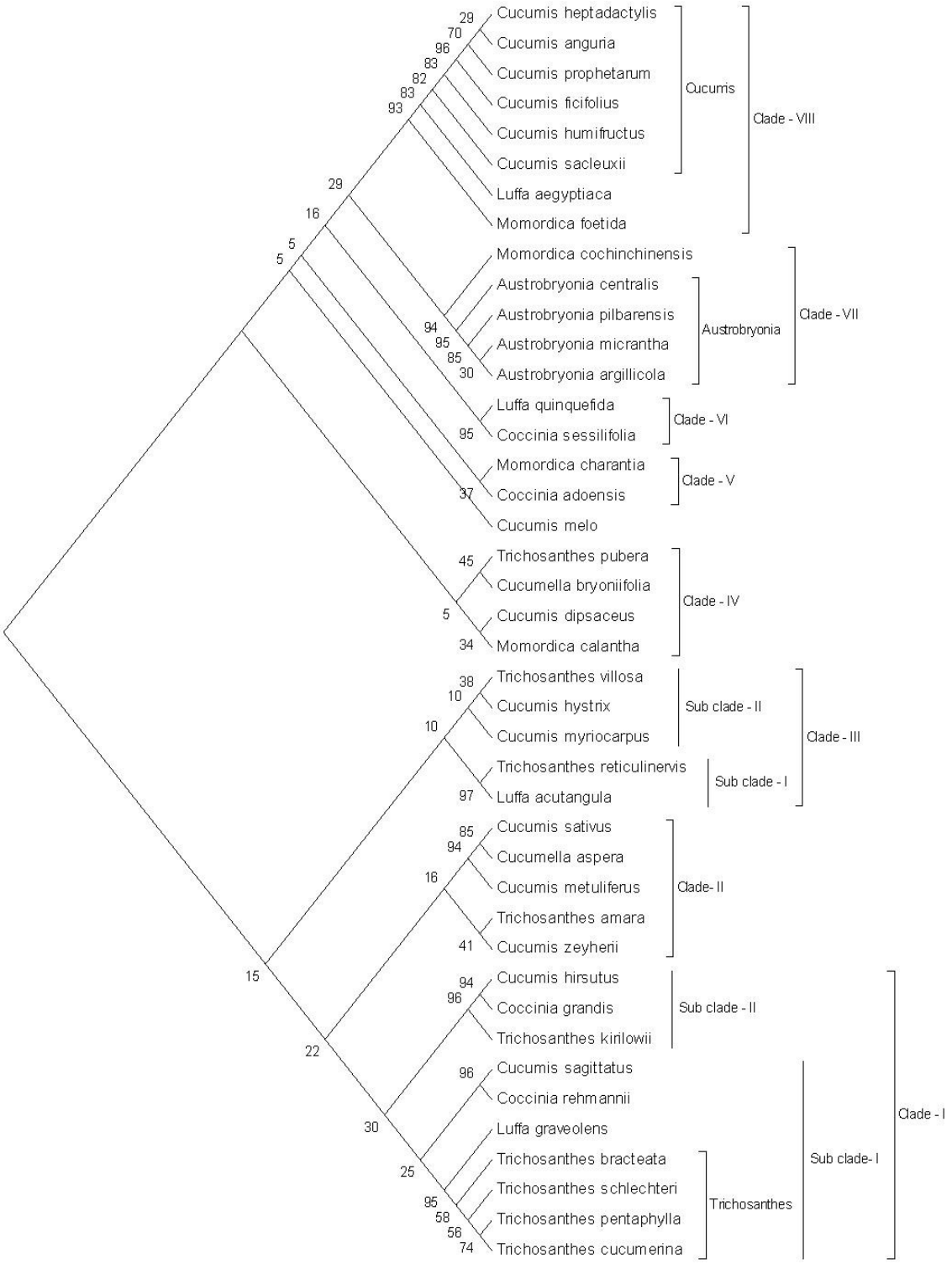

Fig 1: Phylogenetic tree from chloroplast-*rbc*L nucleotide sequences representing 42 taxa of family *Cucurbitaceae* was inferred using UPGMA method.

Fig 2: Phylogenetic tree from chloroplast- *rbc*L aminoacid sequences representing 52 taxa of family *Cucurbitaceae*, was inferred using UPGMA method