SEQUENCING RELATIONSHIPS BETWEEN OF GENUS CAMELLIA (GREEN TEA) BASED ON THE CHARACTERISATION OF NUCLEAR ITS REGION

Myong Jung
College of Future Multidisciplinary Studies, Tongmyong University, Busan 48520
KOREA

ABSTRACT

The objective of this study was to verify genetic diversity of genus *Camellia*. This study used Blast data (NCBI) of ITS region (partial internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial 26S ribosomal DNA gene) for further analysis of studies. The mean nucleotide frequencies for thirty-one of genus *Camellia* are A = 18.2%, C = 35.0%, G = 32.3%, and T = 14.5%. There were a total of 695 positions in the final dataset. Total alignment length of genus *Camellia* is 695 positions, of which 44 are parsimony-informative, 117 variables, 69 singletons, 549 conserved, and 528 coverage (100%). Substitution pattern and rates were estimated under the Tamura-Nei (1993) model. Each entry is the probability of substitution from one base (row) to another base (column). The probability of changing from C to T(U) was 30.6 percent, higher than other substitution probabilities. The estimated Transition/Transversion bias (R) is 2.58. Number of segregating sites of genus *Camellia* was 117 and nucleotide diversity (π) was 0.033. Under the neutral mutation hypothesis, the probability that D is negative (-1.490) is small than 0.5.

Keywords: *Camellia***,** ITS region**,** nucleotide frequencies**,** substitution pattern**.**

INTRODUCTION

Tea (*Camellia* L.) is one of oldest, non-alcoholic and caffeine-containing beverage or medicine. Tea has used for this purpose in China for nearly 3,000 years. Genus *Camellia* includes about 280 species and many species of this genus is economically important trees or shrubs in the family Theaceae (Vijayan et al., 2009; Yang et al., 2013). The genus *Camellia* is native to Southern, Eastern Asia and China, which possess more than 80% of the species and are the center of species diversity (Gao et al., 2005). It is an important component of the warm temperate and subtropical forests of that region.
Teas have been classified into *Camellia sinensis* var. *sinensis* O. Kuntze (China type), *C. sinensis* var. *assamica* (Masters) Wright (Assam type), *C. sinensis* subspecies *lasiocalyx* (Planch.) Wright (Combod type) (Roy and Chakraborty, 2009).

According to the Record of Gaya, cited in the Memorabilia of the Three Kingdoms, the legendary queen Heo Hwang-ok, a princess of Ayodhya (Ancient India Kingdom), brought the *Camellia sinensis* (var. *assamica*) tea plant from India to Korea (Hanjae and Venerable, 2010). In practice, however, Labrador tea and fruit teas, such as magnolia berry tea and goji berry tea, were more widely used in the Samhan Era (631-647) instead. Korean tea is a beverage consisting of boiled water infused with leaves (such as the tea plant *Camellia sinensis*).

Genetic diversity or variation and its measurement have vital importance in interpretation, understanding and management of species or populations as well as phylogenetic relationships within very related intraspecies. Population geneticsis essentially a study of the causes and

effects of genetic variation within and between populations, and in the past, isozymes have been amongst the most widely used molecular markersfor this purpose (Hamrick and Godt, 1989; Wendel and Weeden, 1989).

Many researchers have developed the molecular methods based on DNA for identification of related intraspecies. A *molecular marker* is a molecular contained within a sample taken from an orgaism or other matter. Many DNA sequences used to mark or track a particular location (locus) on a particular chromosome, i.e. marker gene. DNA sequences often give higher resolution than do other molecular markers (Kass and Wink, 1997).

Simple sequncing molecular markers, such as RAPD, AFLP and ISSR, have conducted to the evaluation of genetic diversity and genetic relationship of tea germplasms (Wachira et al., 1995; Chen et al. 1998; Liang et al. 2000; Luo et al. 2002; Huang et al. 2004; Yao et al. 2005). The systematically evaluated tea germplasms provide elite materials both for individual selection and parental selection for hybridization as well as for mutation breeding.

Moreover, comparative studies of nucleotide sequences provide a means for analying phylogenetic relationships over a wide rage of taxonomic levels (White et al., 1990). Non-coding sequences (ITS) of nuclear genes are being used to investigate phylogenetic relationships in plants (Hsiao et al., 1994). Nuclear ribosomal DNA internal transcribed spacer sequences (ITS) is eukaryotic ribosomal RNA genes (known as ribosomal DNA or rDNA) are found as parts of repeat units that are arranged in tandem arrays, located at the chromosomal sites known as nuclear organizing regions. Each repeat unit consists of a transcribed region (having genes for 18S, 5.8S and 26S rRNAs and the external transcribed spacers i.e. ETS1 and ETS2) and a non-transcribed spacer (NTS) region. In the transcribed region, internal transcribed spacers (ITS) are found on either side of 5.8S rRNA gene and are described as ITS1 and ITS2. The two internal transcribed spacers (ITS1 and ITS2) have been shown to be relatively valuable targets for defining markers in systematic studies (White et al., 1990).

The development and application of molecular markers in tea breeding are recent, dating back only to the mid-1990s (Gunasekare, 2007). This review focusses on the different molecular markers used in the genetic improvement of tea, both locally and internationally. The majority of molecular marker studies in tea have been confined to genetic diversity analysis. The present work was deigned to review the phylogenetic relationships of the sixteen taxa of the genus *Camellia* by comparing cloned sequences of the ITS regions.

METHODOLOGY

DNA Extraction

Korean cultivar, *Camellia sinensis* var. *sinensis* was collected from Sacheon-ci, Gonmyeong-meon, Gyeongsangnam-do province in Korea. Total genomic DNA was extracted from a fresh young leaves using the plant DNA Zol Kit (Life Technologies Inc., Grand Island, New York, U.S.A.) according to the manufacturer's protocol. DNA was checked for shearing and concentration by DyNA 200 fluorometer (Amersham Pharmacia Biotech, USA).

ITS analysis

Primers 1-4 were used to amplify the ITS region, which consisted of ITS1, 5.8S rDNA, and ITS2 regions (White et al., 1990). PCR materials (50 ul volume) included 50 ng of genomic DNA, 100 uM of each dNTP, 0.2 uM of each primer, 1 x enzyme buffer, and 2 unit of Taq polymerase. The amplification profile was 28 cycles of 94℃ for 30 sec, 42℃ for 60 sec, 72℃ for 60 sec, preceded by an initial denaturation at 94℃ foe 90 sec and followed by a final extension at 72℃ for 5 min.

PCR products were separated on 1.5% agarose gels and purified using the QIAquick Gel Extraction Kit (QIAGEN). The amplified fragments were cloned into a bluescript vector and sequenced using ABI Prism 377 Sequencer (Applied Biosystem, USA). At least ten individuals' clones of each taxon were analyzed.

Data analysis

Earlier study of ITS sequences in Camellia obtained from GenBank (Table 1). In this expanded study, More than 95% of similar sequences were compared using only one data with higher similarity. Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. For each model, AICc value (Akaike Information Criterion, corrected), Maximum Likelihood value (*lnL*), and the number of parameters (including branch lengths) are also presented (Tamura and Nei, 1993). Non-uniformity of evolutionary rates among sites may be modeled by using a discrete Gamma distribution (+*G*) with 5 rate categories and by assuming that a certain fraction of sites is evolutionarily invariable (+*I*). Whenever applicable, estimates of gamma shape parameter and/or the estimated fraction of invariant sites are shown. Assumed or estimated values of transition/transversion bias (*R*) are shown for each model, as well. 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 (Disparity Index test) (Kumar and Gadagkar, 2001).

Codon-based tests of neutrality for analysis between ITS gene sequences were conducted using the Nei-Gojobori method and evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Gene trees were constructed wuth the use of maximum parsimony (MP) analyses in PAUP* version 4.0b8 (Swofford, 2001). MP was inferred using heuristic search, branch-swapping options and tree bisection-reconnection using MEGA X (Kumar et al., 2018). Confidence values for individual branches were determined by a bootstrap analysis with 100 repeated sampling of the data.

RESULTS

Sequences of ITS region (partial internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial 26S ribosomal RNA gene) for sixteen taxa of genus *Camellia* were successful in all of the species. Aligned nucleotidesof ITS RNA were varied within genus *Camellia* varying from 559 bp in *C. sinensis* cultivar De Hua She 10 to 657p in *C. taliensis* and *C. euphlebia* with a mean of 639.7 bp (Table 2). The base composition did not show a significant difference among total taxa. The mean nucleotide frequencies for thirty-one of genus *Camellia* are A = 18.2%, C = 35.0%, G = 32.3%, and T = 14.5%. There were a total of 695 positions in the final dataset. Total alignment length of genus *Camellia* is 695 positions, of which 44 are parsimony-informative, 117 variables, 69 singletons, 549 conserved, and 528 coverage (100%).

Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. For each model, AICc value (Akaike Information Criterion, corrected), Maximum Likelihood value (lnL), and the number of parameters (including branch lengths) are also presented (Table 3).

Substitution pattern and rates were estimated under the Tamura-Nei model (Table 4). Each entry is the probability of substitution from one base (row) to another base (column). The

probability of changing from C to T(U) was 30.6 percent, higher than other substitution probabilities. The estimated Transition/Transversion bias (R) is 2.58.
Number of segregating sites of genus *Camellia* was 117 and nucleotide diversity (π) was 0.033. Under the neutral mutation hypothesis, the probability that the Tajima test statistic (D) is negative (-1.490) is small than 0.5 (Table 5).

Disparity Index per site is shown for all sequence pairs (Table 6). 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. 105 (87.5%) values of the them are greater than zero. Values greater than 0 indicate the larger differences in base composition biases than expected based on evolutionary divergence between sequences and by chance alone.

A pairwise distance (PD) based on the proportion of shared sequences was used to evaluate relatedness among nine strains (Table 7). The estimate of PD ranged from 0.002 between *C. semiserrata* for. *albiflora* and *C. phellocapsa* to 0.080 between *C. sinensis* wild Mt. Giri and *C. semiserrata.*

The evolutionary history based on sequences of genus *Camellia* was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-1178.23) is shown. ITS gene of genus *Camellia* used with success to reconstruct phylogenetic relationships. Initial tree(s) for the heuristic search. Neighbor-Join and BioNJ algorithms was used initial tree(s) for the heuristic search. Maximum Composite Likelihood (MCL) was also used to estimate a matrix of pairwise distances and then selecting the topology with superior log likelihood value. All ITS rRNA of genus *Camellia* generated in this study exhibited well solved topology with high bootstrap support irrespective of the methods (parsimony) and the setting used (Fig. 1). Sixteen taxa formed two or three distinct clades with moderate bootstrap support (for convenience, only bootstrap values under parsimony criterion are presented here). *C. sinensis* var. *sinensis* and *C. sinensis* isolate T1236 were grouped together.

Table 1. 16 taxa of genus *Camellia* and their information of internal transcribed spacer (ITS) from GenBank.

Taxa	No. of GenBank	Authors
Camellia sinensis var. *sinensis*	This study	
C. sinensis isolate T1236	HM061493.1	Vijayan and Tsou, 2007
C. sinensis cultivar De Hua She	FJ00483.1	Lee et al., 2008
C. sinensis var. *assamica*	MH270493.1	Zeng et al., 2018
C. sinensis for. *formosensis*	EF544713.1	Vijayan and Tsou, 2007
C. semiserrata	EF649688.1	Tian and Li, 2007
C. semiserrata for. *albiflora*	EF649691.1	Tian and Li, 2007
C. phellocapsa	EF649689.1	Tian and Li, 2007
C. liberistamina	EF649692.1	Tian and Li, 2007
C. tachangensis var. *remotisserrata*	EU579724.1	Tsou et al., 2009
C. leptphylla	EU579732.1	Tsou et al., 2009
C. taliensis	EU579778.1	Tsou et al., 2009

C. waldeniae	EU579788.1	Tsou et al., 2009
C. angustifolia	FJ432097.1	Tsou et al., 2009
C. kwangsiensis	FJ432106.1	Tsou et al., 2009
C. euphlebia	FJ432147.1	Tsou et al., 2009

Table 2. Base frequencies and total length across 16 taxa of genus *Camellia* using internal transcribed spacer (ITS)

Taxa	Base (%)				Total (bp)
	T(U)	C	A	G	
Camellia sinensis var. *sinensis*	14.1	35.6	17.6	32.7	637
C. sinensis isolate T1236	14.2	35.7	17.6	32.5	636
C. sinensis cultivar De Hua She	17.9	30.4	21.5	30.2	559
C. sinensis var. *assamica*	14.0	35.7	18.0	32.4	645
C. sinensis for. *formosensis*	14.1	34.6	18.2	33.1	638
C. semiserrata	14.3	35.7	18.2	31.8	650
C. semiserrata for. *albiflora*	14.4	35.2	18.6	31.8	651
C. phellocapsa	14.4	35.3	18.6	31.6	651
C. liberistamina	14.3	35.7	17.3	32.7	636
C. tachangensis var. *remotisserrata*	14.2	35.1	17.8	32.8	646
C. leptphylla	14.1	35.0	18.1	32.8	652
C. taliensis	14.8	35.9	17.4	32.0	657
C. waldeniae	14.0	35.2	17.8	33.0	630
C. angustifolia	14.2	34.9	18.2	32.6	653
C. kwangsiensis	14.8	34.7	17.7	32.8	637
C. euphlebia	14.2	35.3	18.7	31.8	657
Mean	14.5	35.0	18.2	32.3	639.7

Table 3. Maximum likelihood estimate of gamma parameter for site rates

Model	Param	BIC	AICc	lnL	Invariant	Gamma	R
T92+G	32	3819.68	3588.42	-1762.1	n/a	0.19	3.03
T92+G+I	33	3827.05	3588.56	-1761.17	0.50	0.67	3.06
HKY+G	34	3832.70	3586.99	-1759.38	n/a	0.18	3.04
TN93+G	35	3840.17	3587.24	-1758.50	n/a	0.19	3.03
HKY+G+I	35	3840.32	3587.39	-1758.57	0.50	0.67	3.09
TN93+G+I	36	3847.57	3587.42	-1757.58	0.49	0.66	3.08
T92+I	32	3848.02	3616.76	-1776.27	0.42	n/a	2.71

HKY+I	34	3861.25	3615.54	-1773.65	0.42	n/a	2.71
GTR+G	38	3862.29	3587.71	-1755.71	n/a	0.18	2.99
TN93+I	35	3867.85	3614.92	-1772.34	0.42	n/a	2.71
GTR+G+I	39	3870.31	3588.51	-1755.1	0.49	0.66	3.03
T92	31	3874.11	3650.07	-1793.94	n/a	n/a	2.58
HKY	33	3887.22	3648.73	-1791.25	n/a	n/a	2.57
GTR+I	38	3891.45	3616.86	-1770.28	0.42	n/a	2.69
TN93	34	3893.84	3648.13	-1789.95	n/a	n/a	2.57
K2+G	31	3894.35	3670.30	-1804.05	n/a	0.21	2.82
K2+G+I	32	3900.96	3669.70	-1802.75	0.47	0.65	2.86
GTR	37	3918.65	3651.28	-1788.5	n/a	n/a	2.58
K2+I	31	3919.99	3695.95	-1816.88	0.42	n/a	2.63
K2	30	3943.15	3726.32	-1833.07	n/a	n/a	2.58
JC+G	30	3979.92	3763.10	-1851.46	n/a	0.21	0.50
JC+G+I	31	3986.92	3762.87	-1850.34	0.45	0.63	0.50
JC+I	30	4004.24	3787.42	-1863.62	0.42	n/a	0.50
JC	29	4026.26	3816.66	-1879.24	n/a	n/a	0.50

Table 4. Maximum likelihood estimate of substitution matrix

	A	T/U	C	G
A	-	1.860	4.505	**19.855**
T/U	2.336	-	**30.623**	4.154
C	2.336	**12.644**	-	4.48
G	**11.167**	1.860	4.505	-

Table 5. Results from Tajima's neutrality test for internal transcribed spacer (ITS) sequences of genus *Camellia*

M	S	ps	Θ	$\prod$	D
16	117	0.168	0.051	0.033	-1.490

M = number of sites, S = Number of segregating sites, ps = S/M, Θ = ps/a1, and π = nucleotide diversity. D is the Tajima test statistic.

Table 6. Disparity index among 16 taxa of genus *Camellia*

Taxa								
EF544713.1								
EF649688.1	1.0000							
EF649689.1	1.0000	0.2226						
EF649691.1	1.0000	0.2444	1.000					
EF649692.1	1.0000	1.0000	0.176	0.252				
EU579724.1	1.0000	1.0000	1.000	1.000	1.000			
EU579732.1	1.0000	0.3688	0.178	0.254	1.000	1.000		
EU579778.1	0.3244	1.0000	1.000	1.000	0.170	0.134	0.212	
EU579788.1	0.300	1.000	1.000	1.000	1.000	1.000	1.000	0.15

	4	0	0	0	0	0	0	2							
FJ004883	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000						
FJ432097.1	1.000	1.000	0.328	1.000	1.000	1.000	1.000	0.304	1.000	0.000					
FJ432106.1	0.256	0.316	1.000	1.000	0.216	0.310	0.110	1.000	0.172	0.000	0.254				
FJ432147.1	0.108	1.000	1.000	1.000	0.078	0.154	0.056	1.000	0.096	0.000	0.092	0.406			
HM061493.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	0.258		
KOREA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.342	1.000	0.000	1.000	1.000	0.214	1.000	
MH270493.1	1.000	1.000	1.000	1.000	0.226	0.234	0.038	1.000	0.016	0.000	0.032	0.124	0.080	1.000	1.000

Table 7. Pairwise distance among 16 taxa of genus *Camellia*

Taxa								
EF544713.1								
EF649688.1	0.041							
EF649689.1	0.039	0.005						
EF649691.1	0.041	0.006	0.002					
EF649692.1	0.006	0.037	0.036	0.037				
EU579724.1	0.026	0.051	0.049	0.051	0.023			
EU579732.1	0.014	0.044	0.042	0.044	0.008	0.024		
EU579778.1	0.038	0.046	0.044	0.046	0.031	0.049	0.046	
EU579788.1	0.011	0.03	0.03	0.03	0.00	0.01	0.00	0.04

		8	6	8	7	4	8	2							
FJ004883	0.0882	0.0885	0.0887	0.0889	0.0879	0.0891	0.0891	0.0884	0.0887						
FJ432097.1	0.0144	0.0043	0.0042	0.0044	0.0008	0.0024	0.0003	0.0046	0.0008	0.0089					
FJ432106.1	0.0030	0.0038	0.0036	0.0038	0.0028	0.0041	0.0038	0.0014	0.0034	0.0077	0.0038				
FJ432147.1	0.0063	0.0066	0.0068	0.0069	0.0063	0.0070	0.0070	0.0061	0.0063	0.0097	0.0070	0.0061			
HM061493.1	0.0016	0.0036	0.0034	0.0036	0.0014	0.0034	0.0024	0.0046	0.0021	0.0076	0.0024	0.0038	0.0064		
KOREA	0.0014	0.0031	0.0029	0.0031	0.0010	0.0027	0.0021	0.0042	0.0015	0.0080	0.0021	0.0036	0.0063	0.0006	
MH270493.1	0.0011	0.0036	0.0034	0.0036	0.0005	0.0027	0.0014	0.0034	0.0011	0.0074	0.0014	0.0024	0.0068	0.0016	0.0013

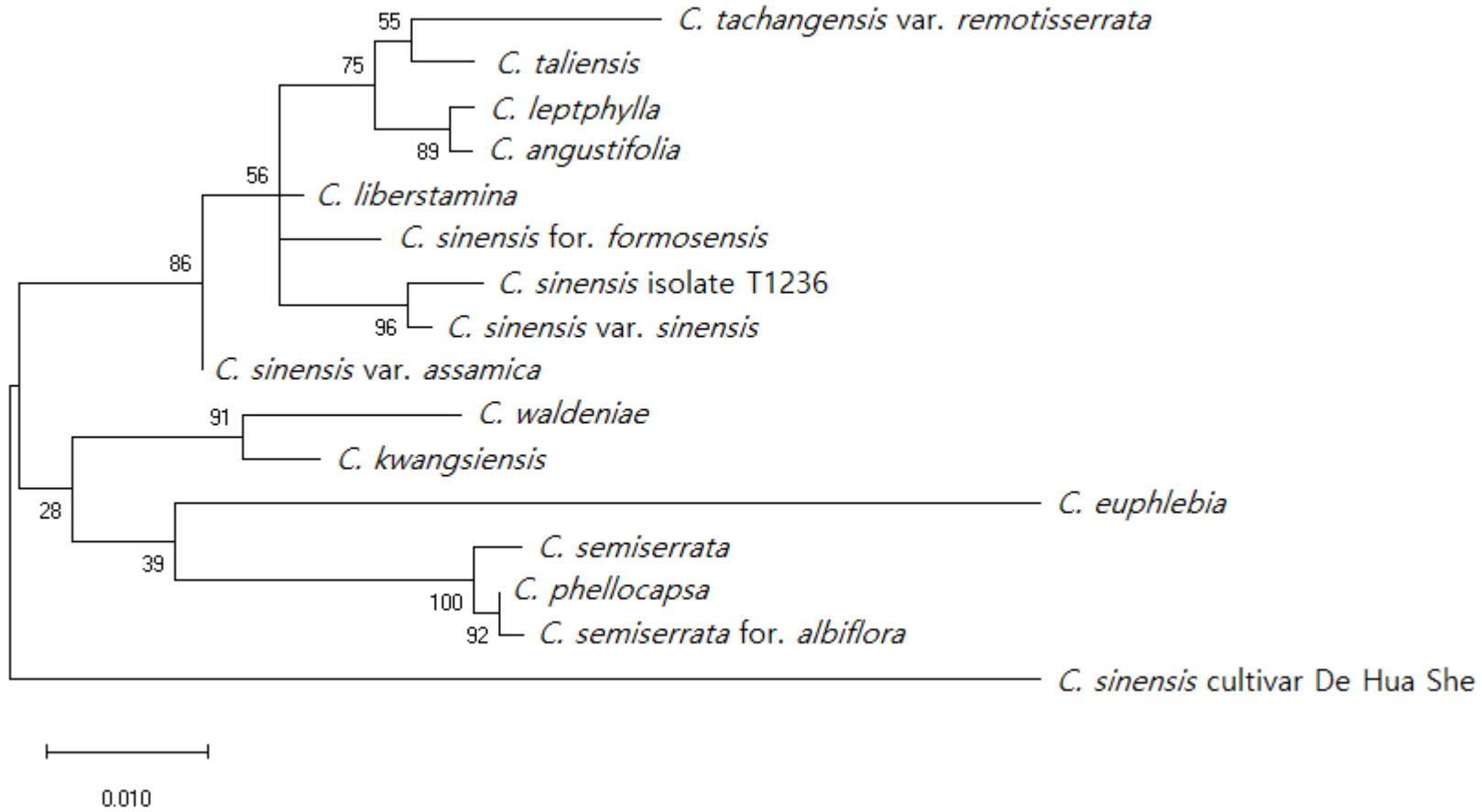

Figure 1. The maximum parsimonious tree for sixteen taxa of genus *Camellia* based on internal transcribed spacer (ITS) sequences using MEGA 10.1. The values of bootstrap were shown in side of vertical lines.

DISCUSSION

Anatomical and morphological characters are very valuable for green tea identification, so many analyzed a number of botanical characters for many *Camellia* species (Lum et al., 2007). Chen et al (2005) reported genetic diversity and differentiation of *C. sinensis* (cultivated tea) and its wild relatives in Yunnan province of China based on allozyme studies. As the result, the percentage of polymorphic loci for each taxon was 21.4–50.0%. Mean heterozygosity per locus (He) varied 0.114–0.218. Genetic divergence between one population indicated that only 4.6% of the variations could be ascribable to genetic differences among taxa. Molecularly, high diversity exists among all the tea genotypes and extensively different from each other on the basis of genetic polymorphism using SCoT (Start Codon Targeted Polymorphism) and RAPD markers (Chen and Yamaguchi, 2002; Collard Band Mackill, 2009). Vijayan et al (2009) reported molecular taxanomy of *Camellia* (Theaceae) inferred from nrITS sequences. According to their results, the 5.8S region was 164 bp for most species, but in numbers belonging to section *Oleifera* and several allied species had an insertion of A or AA after the fist four (TAAA) conserved basis of 5.8S. The unusal length variablity in the 5.8S region revealed the evolutionary dynamism of nrDNA in *Camellia*. In this study, the fist four of 5.8S was conserved (TAAA). There was 69 singletons , 549 conserved among 695 bp nrITS sequences in *Camellia*. The number of segregating sites of genus *Camellia* was 117 based on ITS rDNA sequeces and nucleotide diversity (π) was 0.033 (Table 5).

The genus *Camellia* is endemic to southeastern Asia, and contains a large number of species. More than 400 species have been named and published, but the number has been reduced by combination during taxonomic revisions. There is much disagreement concerning the status of many *Camellia* species. Three different taxonomic authorities place the number of valid species from 80 to 280. Nevertheless, a large majority of the species is native to China.

For many years cultivated teas have been classified into three varieties derived from the species, *Camellia sinensis* L. (Wachira et al., 1995). There varieties are indigenous to different geographic regions of Southeast Asia. Namely, varieties indigenous to Assam in India, Indochina, and China have been named var. *assamica* (Assam tea), var. *assamica* ssp. *lasiocalyx* (Southern fom or Cambod tea). Roy and Chakraborty (2009) reported genetic diversity was greast within China type clones, followed by Assam and Combod type. More than two hundred improvedcultivars have been registered in the country sofar, these cultivars have made considerablecontribution in the whole tea industry in China (Chen et al., 2007). It is noteworth that China type tea clones should be conserved properly along with other varital genotypes for the improvement of commerical tea through breeding, individual selection to hybridization and genetic technical approaches in the world.

Cloning and sequencing of commerical varieties showed high similarity to one of the *C. sinensis* sequences based on ITS region (Lum et al., 2007). An anaysis of a 347-nucleotide region of the *mat*K gene revealed also most similar to the *C. sinensis mat*K sequences in GneBank (Lum er al., 2007).

REFERENCES

Chen, J., Wang, P., Xia, Y., Xu, M., & Pei, S. (2005) Genetic diversity and differentiation of *Camellia sinensis* L. (cultivated tea) and its wild relatives in Yunnan province of China, revealed by morphology, biochemistry and allozyme studies. *Genetic Resources Crop and Evolution,* 52, 41–52.

Chen L, & Yamaguchi S. (2002) Genetic diversity and phylogeny of tea plant (*Camellia sinensis*) and its related species and varieties in the section Thea genus *Camellia* determined by randomly amplified polymorphic DNA analysis. *The Journal of Horticultural Science and Biotechnology*, 77, 729-32.

Chen, L., Zhou, Z.X., & Yang, Y.J. (2007) Genetic improvement and breeding of tea plant (*Camellia sinensis*) in China: from individual selection to hybridization and molecular breeding. *Euphytica*, 154, 239-48.

Collard, B.C.Y., & Mackill, D.J. (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reporter*, 27, 86-93.

Chen L, Yang YJ, Yu FL, Gao QK, Chen DM (1998) A study on genetic diversity of 15 tea cultivars (*Camellia sinensis* (L.) O. Kuntze) using RAPD markers. *Journal of Tea Science*, 18, 21–7.

Gao, J.Y., Parks, C.R., & Du, Y.Q. (2005) *Collected species of the genus Camellia and illustrated outline*. Zhejiang: Zhejiang Science and Technology Press.

Gunasekare, M.T.K. (2007) Applications of molecular markers to the genetic improvement of *Camellia sinensis* L. (tea) – A review. *The Journal of Horticultural Science and Biotechnology*, 82, 161-169.

Hamrick, J.L., & God, M.J.W. (1989) Allozyme diversity in plant species. In A.H.D. Brown, M.T. Clegg, A.L. Kahler & B.S. Weir (Eds.), *Plant population genetics, breeding and genetic resources*. Sinauer Sunderland, MA.

Hsiao, C, Chatterton, N.J., Asay, K.H., & Jensen, K.B. (1995) Molecular phylogeny of the Pooideae (Poaceae) based on nuclear rDNA (ITS) sequences. *Theoretical and Applied Genetics*, 90, 389-98.

Huang, F.P., Liang, Y.R., Lu, J.L., Chen, R.B., & Mamati, G. (2004). Evaluation of genetic diversity in Oolong tea germplasms by AFLP fingerprinting. *Journal of Tea Science*, 24, 183–189.

Hanjae, Y.M., & Venerable, C. (2010) *Korean tea classics*: by translated by Anthony, Brother Anthony of Taizé; Hong, Kyeong-hee; Owyoung, Steven D. Seoul: Seoul Selection.

Kumar, S., & Gadagkar, S.R. (2001) Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics*, 158, 1321-7.

Kumar, S., Stecher. G., Li. M., Knyaz. C., & Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547-9.

Liang, Y.R., Tanaka, J., & Takeda, Y. (2000) Study on diversity of tea germplasm by RAPD marker. *Journal of Zhejiang Forest College*, 17, 215–8.

Loffler, F., Sun, Q., Li, J., & Tiedje, J.M. (2000) 16S rRNA gene-based detection of tetrachloroethene-dechlorinating *Desulfuromonas* and *Dehalococcoides* species. *Applied Environmental Microbiology*, 66, 1369-74.

Luo, J.W., Shi, Z.P., Shen, C.W., Liu, C.L., Gong, Z.H., & Huang, Y.H. (2002) Studies on genetic relationships of tea cultivars [*Camellia sinensis* (L.) O Kuntze] by RAPD analysis. *Journal of Tea Science*, 22, 140-6.

Lum, M.R., Baycher, A., Prigge, B.A., Hardy, M., Heber, D., & Hirsch, A.M. (2006) Identification of green tea (*Camellia sinensis* L.) and tea oil (*Camellia oleifera* Abel.) by molecular biological, and anatomical methods. In: S.E. Ebeler, G.R. Takeoka & P. Winterhalter (Eds.), *Authentication of food and wine.* Oxford: Oxford University Press.

Roy, S.C., & Chakraborty, B.N. (2009) Genetic diversity and relationships among tea (*Camellia sinensis*) cultivars as revealed by RAPD and ISSR based fingerprinting. *Indian Journal of Biotechology*, 8, 370-6.

Swofford, D.L. (2001) *PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b10*. Massachusetts, USA: Sinauer, Sunderland.

Tamura, K., & Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10, 512-26.

Vijayan, K., Zhang, W., & Tsou, C. (2009) Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences. *American Journal of Botany, 96*, 1348-60.

Wachira, F.N. Waugh, R., Hackett, C.A., & Powell, W. (1995) Detection of genetic diversity in tea (*Camellia sinensis*) using RAPD markers. *Genome*, 38, 201-10.

Wendel, J.F., & Weeden, N.F. (1989) Visualization and interpretation of plant isozymes. In D.E. Soltis & P.S. Soltis (Eds.), *Isozymes in plant biology*. London: Chapman and Hall.

White, T.J., Bruns, T., Lee, S., & Taylor, J. (1999) Amplification and direct sequencing of fungal ribosomal genes for phylogenetics. In M.A. Innis, D.H. Gelfand, J.J. Sninsky & T.J. White (Eds.), *PCR protocols: a guide to methods and applications*, New York: Academic Press.

Yang, J.B., Yang, S.X., Li, H.T., & Yang, J., & Li, D.Z. (2013) Comparative chloroplast genomes of *Camellia* species. PLoS One. 8(8):e73053.

Yao, M.Z., Huang, H.T., Yu, J.Z., & Chen, L. (2005) Analysis of applicability of ISSR on molecular identification and relationship investigation of tea cultivars. *Journal of Tea Science*, 25, 153–7.