Genetic Resources and Crop Evolution **52:** 41–52, 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

Jin Chen[1,2,*], Pingsheng Wang[3], Yongmei Xia[1], Mei Xu[3] and Shengji Pei[2]
*[1]Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan province 666303, China; [2]Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan province 650204, China; [3]Tea Research Institute, Yunnan Academy of Agricultural Sciences, Menghai, Yunnan province 666201, China; *Author for correspondence (e-mail: biotrans@bn.yn.cninfo.net)*

Received 12 June 2002; accepted in revised form 13 June 2003

Key words: Allozyme, Biochemistry, *Camellia*, Differentiation, Genetic diversity, Morphology

Abstract

We evaluated morphological, isozyme and biochemical diversity of a total of 87 accessions in the genus *Camellia* [*Camellia sinensis* var. *sinensis* (10), *C. talinensis* (7), *C. sinensis* var. *dehungensis* (3), C. *crassicolumna* (3) and *C. sinensis* var. *assamica* (64)]. Great variation of morphological characters was apparent within each taxa. Across the five taxa, all leaf and most flower quantitative characters showed significant differences while all fruit quantitative characters measured did not differ significantly, and, seven (i.e., life form, bud color, petal texture, pubescence on ovary, style number, stamen location and locule per fruit) of the 33 qualitative characters yield significant differences. As a whole, caffeine content had the highest variation with CV of 22.7%, water extract solid showed the least variation (13.4%) and content of polyphenols (20.0%) and free amino acids (18.8%) showed intermediate variations. *Camellia taliensis* and *C. sinensis* var. *assamica* had significantly higher content of polyphenols and water extract solid than in the other three taxa, while no significant differences were detected for the content of caffeine and free amino acids. For allozyme study, 14 loci presented good resolution, among which, nine loci (64%) were polymorphic in each taxon (AAT-3, FUM-1, 6PDG-1, G6PDH-1, G3PDH-1, ME-1, PGM-1, PGM-2 and SKD-1). The percentage of polymorphic loci (P) for each taxon was 21.4–50.0%. Mean heterozygosity per locus (H_e) varied 0.114–0.218. F_{ST} value indicated that only 4.6% of the variations could be ascribable to genetic differences among taxa. Genetic relationships among the five taxa revealed by allozymes, were also exposed by the result of clustering of the morphological and biochemical characters.

Introduction

The tea plant, which has been cultivated in China for more than 2000 years (Li 1983; Evans 1992; Weatherstone 1992), is a beverage crop with worldwide significance. Due to its high economic importance, extensive collections of tea plants have been made in China and several other countries (Wachira et al. 2001). Numerous studies to evaluate genetic diversity have also been conducted by incorporating various methods, including morphology (Wichremaratne 1981; Toyao and Takeda 1999), biochemistry (Takeda 1994; Magoma et al. 2000), and the use of genetic markers, e.g., RFLPs

(Matsumoto et al. 1994), RAPDs (Lee et al. 1995; Wachira et al. 1995, 1997; Kaundun et al. 2000; Kaundun and Park 2002), AFLPs (Paul et al. 1997), and ISSRs (Lai et al. 2001). However, most of the materials used in these studies were from non-indigenous countries such as Kenya, Japan, and the UK. The assessment of genetic diversity on materials from China is still lacking.

Camellia sinensis var. *sinensis* and *C. sinensis* var. *assamica* are the two main taxa for commercial cultivation. However, there are a number of species of *Camellia*, such as *C. taliensis*, *C. crassicolumna*, *C. sinensis* var. *dehungensis*, etc., that have been used as tea by local people in parts of Asia (Chang and Bartholomew 1984), in particular in Yunnan province (Ming 2000). Therefore, it would be obviously useful to both breeding programs and for the germplasm conservation of tea plants to understand the differentiation of morphology, biochemical, and genetic aspects among those taxa.

Yunnan province is one of the centers of genetic diversity of many crops benefiting from its particular geographical location, complicated landforms, climate conditions, and the settlement of numerous indigenous minority groups in this region (Zeng et al. 2001). Among these crops is tea (Chen 1994; Ming 2000). Yunnan province hosts most species of *Thea* section of *Camellia* (Chang and Bartholomew 1984; Ming 2000). Tea cultivation in Yunnan may date back to the Tang Dynasty (618–907 AD.) (Xiang 1962). Yunnan 'Pu'er' Tea, a famous tea that was exported to Tibet and many different places throughout history, played an important role in the local economy (Song and Li 1990; Sao and Sheng 1993). Even today, tea plantation in Yunnan occupies 162968 ha in area, with 75137 *T* in production in 1999 (Yunnan Statistic Bureau 2000). In brief, genetic resource of tea in the Yunnan province is undoubtedly one of the most important parts of the germplasm resource.

The evaluation of genetic diversity in tea is necessary from the view of both morphological traits and genetic markers. In this paper, by combining the assessment of morphology with biochemistry and enzymatic techniques, we aim to reveal the diversity and differentiation of *Camellia sinensis* and its several wild relatives in Yunnan province. The implications of our results hold promise for assessing genetic diversity and for strategies of germplasm conservation.

Table 1. List of evaluated *Camellia* species and the sample number. The taxonomic classification for the plants was based on Ming (2000).

Taxa	No. of sample	Clue
Camellia sinensis var. *sinensis* (L.) O. Kuntze	10	SIN
C. taliensis (W. W. Smith) Melchior	7	TAL
C. sinensis var. *dehungensis* (H. T. Chang et Chen) Ming	3	DEH
C. crassicolumna H. T. Chang	3	CRA
C. sinensis var. *assamica* (Masters) Kitamura	64	ASS
Total	87	

Materials and methods

Plant materials

During 1981–1985, the second author (P.S. Wang) and his colleagues in Tea Research Institute of Yunnan Academy of Agricultural Sciences conducted an extensive survey on tea plant varieties consumed by local people in Yunnan province. Seeds or seedlings of those local varieties, a total of about 700 accessions, were collected and then cultivated at the 'National Tea Plant Germplasm Collection (Menghai)' located within the Institute (21° 28′ N, 99° 56′ S, 1300 m in altitude). All of the plants were maintained under similar agricultural management practices. Plants were fertilized with roughly 250 kg N ha^{-1}, supplied as carbamide (46% N) in two split application and one time organic fertilizer. Irrigation, and pest and disease control were according to guidelines for normal production. We chose a total of 87 samples (see Table 1 for the sample size of each taxa) from the National Tea Plant Germplasm Collection. We understood that the sample size for *C. taliensis*, *C. sinensis* var. *dehungensis* and *C. crassicolumna* was not sufficient for a logical comparison for the heterozygosity and percentage of polymorphisms of allozymes among species. However, they were all the samples for the three species available in the National Tea Plant Germplasm Collection. For each accession, we collected data and materials from 10 individual plants. We aimed to collect the representative sample of commonly cultivated tea

in the province (especially for *C. sinensis* var. *assamica*) based on geographical distribution.

Morphology

Fifty-five characters were measured which include 22 quantitative characters and 33 qualitative characters (Appendix 1). We collected mature leaves with full size for measurement from the summer shoots during August in 2000. We have 10 duplicates for those length measurements and three duplicates for the weight data. For temporal measurement, we collected data once a week and averaged the data of 1990–2000.

Biochemistry

Fresh plant materials (two leaves and a bud) were collected during February–June in 2001, then they were dried under 80 °C. Samples were analyzed for the content of caffeine, water extract solid, polyphenols and free amino acids at the Biochemical Analysis Center of Yunnan Academy of Agricultural Sciences in Kunming. The Measurements followed the Sate Standard of China for tea content determination recorded as GB 8305-87, 8312-87, 8313-87 and 8314-87 (TLPMB and ITP 1989).

Isozyme electrophoresis

Leaf samples (1–2 cm^2) or bud tissue were collected in the field and placed in microfuge tubes, and stored in an ice-filled cooler. Samples were kept in cold storage until processed, for up to 3 days. Enzymes were extracted in a buffer consisting of 0.1 M Tris–HCl (pH 7.5), 6.6% PVPP and β-mercaptoethanol (1 drop/ml). Enzyme electrophoresis employed cellulose acetate gels in either a Tris–glycine (TG) or citrate–morpholine (CM) electrode buffer (Hebert and Beaton 1993). Ten screened enzymes, coding for 14 loci were successfully resolved: one locus each for aldolase (ALD, E.C. [Enzyme Commission] 4.1.2.13), fumerase (FUM, E.C.4.2.1.2), isocitrate dehydrogenase (IDH, E.C. 1.1.1.42), glucose-6-phosphate dehydrogenase (G6PDH, E.C.1.1.1.49), 6-phosphogluconate dehydrogenase (6PGDH, E.C.1.1.1.44), shikimate dehydrogenase (SKD, E.C. 1.1.1.25), and malic enzyme (ME, E.C. 1.1.1.40); two loci each for glyceraldehydes-3-phosphate dehydrogenase (G3PDH, E.C.1.2.1.12), phosphoglucomutase (PGM, E.C.5.4.2.2); and three loci for aspartate amino transferase (AAT, E.C.2.6.1.1). For other 14 enzymes we could not score bands due to poor staining or resolution. Enzymes resolved on the TG buffer were FUM and PGM and the enzymes resolved on the CM buffer were AAT, ALD, IDH, G3PDH, ME, SKD, 6-PGDH and G-6-PDH. TG gels were run for 15–20 min at 200 mV and CM gels from 20–30 min at 150 mV.

Interpretation of the genetic basis of the stain patterns was based on the number of isozymes reported for diploid plants (Weeden and Wendel 1989). The most anodal isozyme was designated as '1', likewise, the most anodal allozyme was designated 'a'.

Data analysis

After homogeneity of sample variances was verified using a Levene's test, an ANOVA/MANOVA model (in which taxa was defined as the fixed independent variable) was used to compare means among different taxa for all quantitative characters. A Kruskal–Wallis test was used to determine the differences of qualitative characters among the five taxa (STATICA for Windows. StatSoft Inc. 1995, Tulsa, OK). Comparisons of means among taxa for the biochemical traits and the leaf characters were done using T′, T-K, GT_2 method (BIOM Stat: Statistical Software for Biologists, Version 3.2, 1996. Applied Biostatistics, Inc., Setanket, NY, USA). Mean value of the morphological and biochemical characters of each taxa were used to produce a cluster analysis with STATICA statistical package (STATICA for Windows. StatSoft). Cluster analyses were created on the dissimilarity matrix of Euclidean distances with UPGMA as the clustering algorithm.

For allozyme, we treated each taxa as a population and used the Wright's (1978) F statistics to calculate each polymorphic locus in each taxa. Genotype frequencies for each locus and genetic structure in each taxa were analyzed using BIOSYS 1.7 (Swofford and Selander 1989). The percentage of polymorphic loci (P), expected (H_e) and observed (H_o) heterozygosity and mean number of alleles (A) in each taxa were estimated by Nei's method (1978). With the Wright's F value, we try to understand to what extend the genetic

Table 2. Variation of 22 quantitative morphological characters of *C. sinensis* var. *sinensis* (SIN), *C. taliensis* (TAL) and *C. sinensis* var. *assamica* (ASS).

	SIN				TAL				ASS			
Characters*	*n*	Range	Mean ± SD	C.V.%	*n*	Range	Mean ± SD	C.V.%	*n*	Range	Mean ± SD	C.V.%
Leave and shoot												
Date of bud first flushing (4)	10	2.2–3.3	2.7 ± 0.36	13.3	7	1.5–2.1	1.8 ± 0.27	15.0	64	1.8–3.8	2.6 ± 0.53	20.3
Length of plucking shoot (7)	10	5.8–8.2	6.9 ± 0.76	11.0	7	5.6–10.9	8.3 ± 1.80	21.7	64	6.8–13.0	9.3 ± 1.29	13.9
Weight of plucking shoot (8)	10	41–108	68 ± 17.7	25.9	7	73–162	109 ± 29	26.9	64	64–244	118 ± 35	29.7
Leaf length (19)	10	7.5–14.3	10.4 ± 2.80	26.9	7	8.5–15.1	12.4 ± 2.25	18.1	64	9.2–17.6	13.5 ± 1.99	14.7
Leaf breath (20)	10	3.2–6.4	4.4 ± 1.15	26.1	7	3.2–6.1	5.2 ± 1.01	19.4	64	3.7–8.0	5.4 ± 0.78	14.4
Leaf area (21)	10	24.0–91.1	48.2 ± 25.6	53.1	7	27.5–92.1	65.7 ± 22.17	33.7	64	35.4–116.8	74.5 ± 20.1	27.0
Leaf shape index (22)	10	2.1–2.8	2.4 ± 0.20	8.3	7	2.2–2.6	2.4 ± 0.16	6.7	64	2.2–3.0	2.5 ± 0.16	6.4
Number of leaf vine (23)	10	8.8–13.6	10.5 ± 1.7	16.2	7	7.5–14.5	11.4 ± 2.37	20.8	64	9.2–14.7	11.8 ± 1.37	11.6
Flower												
Date of blooming (26)	10	6.9–9.5	8.3 ± 0.71	8.6	7	8.1–10.2	9.0 ± 0.67	7.4	64	7.2–11.5	9.1 ± 0.86	9.5
Period of flowering (27)	10	1.0–3.0	2.0 ± 0.47	23.5	7	1.0–3.0	2.0 ± 0.82	41.0	64	1.0–4.0	1.9 ± 0.82	43.2
Pedicel length (28)	10	0.8–1.4	1.1 ± 0.20	18.2	7	0.9–1.5	1.2 ± 0.29	24.2	64	0.5–1.5	0.9 ± 0.19	21.1
Sepal diameter (32)	10	3.5–4.3	3.8 ± 0.25	6.8	7	3.4–6.1	4.8 ± 1.12	23.3	64	2.3–5.7	3.8 ± 0.65	17.1
Flower diameter (33)	10	32.3–41.0	36.4 ± 2.81	7.7	7	35.3–52.2	42.9 ± 5.01	11.7	64	25.1–48.1	37.2 ± 4.7	12.6
Petal length (34)	10	16.3–23.6	19.6 ± 2.49	12.7	7	18.4–25.5	22.1 ± 2.89	13.1	64	15.3–28.0	20.3 ± 2.7	13.3
Length of column (39)	10	8.1–15.1	11.7 ± 2.12	18.1	7	11.6–17.9	13.9 ± 2.14	15.4	64	9.1–15.5	11.7 ± 1.47	12.7
Fruit and seed												
Fruit size (45)	10	23.1–32.0	26.4 ± 2.89	10.9	6	22.4–35.7	26.6 ± 4.64	17.4	64	19.0–36.70	27.7 ± 4.40	15.9
Pericarp thickness (48)	10	0.7–1.9	1.0 ± 0.35	35.0	6	0.9–1.7	1.1 ± 0.31	28.2	64	0–2.2	1.1 ± 0.36	32.7
Peduncle length (49)	10	8.8–14.1	11.6 ± 1.99	17.2	6	3.3–15.8	10.6 ± 4.16	39.2	64	6.8–15.0	10.3 ± 1.74	16.9
Peduncle diameter (50)	10	1.6–3.6	2.6 ± 0.62	23.8	6	2.3–4.0	3.2 ± 0.75	23.4	64	1.4–4.2	2.7 ± 0.59	21.9
Seed ripe date (51)	10	9.5–11.3	10.3 ± 0.58	5.6	6	9.6–12.3	10.5 ± 1.12	10.7	64	8.3–12.3	10.4 ± 0.62	6.0
Seed size (53)	10	11.8–17.1	15.2 ± 1.53	10.1	6	15.2–18.5	15.6 ± 1.27	8.3	64	10.5–19.3	15.1 ± 1.78	11.8
Seed weight (55)	10	102–216	147 ± 34.1	23.2	6	86–235	144 ± 50.7	35.2	64	43–390	146 ± 63.3	43.4

*Number in parentheses is the character number in Appendix 1.

variation could be ascribable to genetic difference of different taxa. Genetic distance between taxa was estimated by the Modified Rogers' Distance (*D*) (Wright 1978).

Results

Morphological characters

Due to the small sample size of *C. crassicolumna* and *C. sinensis* var. *dehungensis*, we presented the variation of morphological characters only in the *C. sinensis* var. *sinensis*, *C. taliensis* and *C. sinensis* var. *assamica*. All three taxa had great morphological variation within a taxon (Table 2). Of 22 quantitative characters, five characters exhibited a coefficient of variation >20% across all three taxa, i.e., leaf area, weight of plucking shoot, period of flowering, pericarp thickness, and seed weight (Table 2).

A significant differentiation occurred for the quantitative characters among the taxa. All of the measured characters for leaf and shoot showed significant differences. Most of the flower characters differed significantly among the taxa, except the characters of period of flowering and petal length. None of fruit characters showed significant differences (Table 3).

Seven of the 33 evaluated qualitative characters yielded significant differences across taxa. These characters included life form ($n = 87$, $P = 0.0002$), bud color ($n = 87$, $P = 0.014$), petal texture ($n = 87$, $P = 0.043$), pubescence on ovary ($n = 86$, $P = 0.011$), style number ($n = 86$, $P = 0.0005$), stamen location ($n = 79$, $P = 0.005$) and locule per fruit ($n = 79$, $P = 0.0049$). Other qualitative characters did not show significant difference among the five taxa ($P > 0.05$).

Biochemistry

Comparing the four biochemical characteristics, caffeine content had the highest variation with a coefficient of variation (=mean/standard deviation) of 22.7%, and water extract solid showed the least variation with a coefficient of variation (CV) of 13.4%. The content of polyphenols and free amino acids showed intermediate variations with CV of 20.0 and 18.8%, respectively (Table 4).

Table 3. Manova test for the differences of 20 morphological characters among the five taxa. (df of difference between taxa = 4).

Characters*	df of error	*F*	*p*-level
Leave and shoot			
Length of plucking shoot (7)*	82	11.112	<0.0001
Weight of plucking shoot (8)	82	5.525	0.0005
Leaf length (19)	82	5.925	0.0003
Leaf breath (20)	82	4.295	0.0033
Leaf area (21)	82	4.418	0.0028
Leaf shape index (22)	82	3.819	0.0068
Number of leaf vine (23)	82	2.510	0.0480
Flower			
Date of blooming (26)	80	3.390	0.0130
Period of flowering (27)	80	0.198	0.9386
Pedicel length (28)	80	5.629	0.0005
Flower diameter (33)	80	2.693	0.0367
Petal length (34)	80	1.247	0.2980
Length of column (39)	80	2.779	0.0323
Fruit and seed			
Fruit size (45)	73	0.321	0.8630
Pericarp thickness (48)	73	1.613	0.1802
Peduncle diameter (50)	73	1.324	0.2690
Seed ripe date (51)	72	1.122	0.3530
Seed size (53)	72	0.558	0.6938
Seed weight (55)	72	0.036	0.9970

Number in parentheses is the character number in Appendix 1.

Camellia taliensis and *C. sinensis* var. *assamica* have significantly higher content of polyphenols and water extract solid than the other three species. There are no significant differences among the five taxa for the content of caffeine and free amino acids (Table 4).

Allozyme

We resolved 10 enzyme systems coding for 14 loci, of which nine (64%), AAT-3, FUM-1, 6PGD-1, G6PDH-1, G3PDH-1, ME-1, PGM-1, PGM-2 and SKD-1 were polymorphic at the species level (Table 5). Two loci had three alleles, seven loci had two alleles, and five loci were monomorphic (Table 5). The mean number of alleles per locus is only 1.52. The percentage polymorphic loci varied from 21.4% (*C. sinensis* var. *dehungensis*) to 50.0% (*C. sinensis* var. *sinensis*). As the sample size of the five taxa was extremely different, the comparison of value of the percentage polymorphic loci may not be reliable. In three of the five taxa, the observed heterozygosity was lower than the expected heterozygosity (Table 6), though chi-square tests

Table 4. The amount and variation of biochemical characteristics of five taxa. Means with a common letter do not differ from other means ($P \leq 0.05$). See Table 1 for the abbreviations of the taxa.

		Caffeine (%)				Polyphenols (%)				Free amino acids (%)				Water extract solid (%)			
Taxa	*N*	Max	Min	Mean ± SD	CV	Max	Min	Mean ± SD	CV	Max	Min	Mean ± SD	CV	Max	Min	Mean ± SD	CV
SIN	10	5.4	2.3	3.67 ± 0.86 A	23.4	33.2	19.5	23.48 ± 4.34 A	18.5	5.1	3.3	4.02 ± 0.47 A	11.7	42.5	29.4	33.50 ± 4.28 A	12.8
TAL	7	3.7	2.1	2.98 ± 0.58 A	19.5	34.8	20.5	29.37 ± 5.67 B	19.3	5.6	3.6	4.61 ± 0.80 A	17.4	44.0	29.7	38.50 ± 5.38 B	14.0
DEH	3	4.5	2.4	3.43 ± 1.07 A	31.2	30.1	20.8	24.80 ± 4.78 A	19.3	5.2	4.1	4.47 ± 0.63 A	14.1	37.6	31.5	34.90 ± 3.11 A	8.9
CRA	3	3.7	2.3	3.22 ± 0.84 A	26.1	29.6	20.1	24.60 ± 4.77 A	19.4	6.0	3.9	4.72 ± 0.65 A	13.8	36.4	28.8	33.30 ± 3.99 A	12.0
ASS	65	5.3	2.0	3.43 ± 0.81 A	23.6	35.8	18.5	26.64 ± 5.28 B	19.8	7.5	3.2	4.36 ± 0.85 A	19.5	45.6	27.1	36.36 ± 4.80 B	13.2
Total	88	5.4	2.0	3.43 ± 0.78	22.7	35.8	18.5	26.3 ± 5.26	20.0	7.5	3.2	4.36 ± 0.82	18.8	45.6	27.1	35.99 ± 4.83	13.4

Table 5. Isozyme frequencies at 14 loci scored for 5 taxa of *Camellia*.

	Taxa				
Locus	ASS	CRA	DEH	SIN	TAL
ATT-3					
(*n*)*	67	3	3	14	8
A**	0.261	0.333	0.000	0.464	0.375
B	0.739	0.667	1.000	0.536	0.625
FUM-1					
(*n*)	67	3	3	14	8
A	0.993	1.000	1.000	1.000	1.000
B	0.007	0.000	0.000	0.000	0.000
G3PDH-1					
(*n*)	51	1	3	5	6
A	0.902	1.000	1.000	0.700	1.000
B	0.010	0.000	0.000	0.000	0.000
C	0.088	0.000	0.000	0.300	0.000
G6PDH-1					
(*n*)	59	2	3	12	8
A	0.517	0.750	0.500	0.542	0.500
B	0.102	0.250	0.167	0.083	0.188
C	0.381	0.000	0.333	0.375	0.313
ME-1					
(*n*)	67	3	3	14	8
A	0.015	1.000	0.000	0.143	0.125
B	0.985	0.000	1.000	0.857	0.875
6PGDH-1					
(*n*)	67	3	3	14	8
A	0.978	1.000	1.000	0.929	1.000
B	0.022	0.000	0.000	0.071	0.000
PGM-1					
(*n*)	67	3	3	14	8
A	0.993	1.000	1.000	1.000	1.000
B	0.007	0.000	0.000	0.000	0.000
PGM-2					
(*n*)	67	3	3	14	8
A	0.478	0.167	0.667	0.429	0.375
B	0.425	0.833	0.333	0.500	0.563
C	0.097	0.000	0.000	0.071	0.063
SKD-1					
(*n*)	66	3	3	13	7
A	0.644	0.500	0.833	0.577	0.571
B	0.356	0.500	0.167	0.423	0.429

**n* indicates the sample size for each locus in each taxon.
***A, B C and D represent faster, fast, medium, and slow electromorphs, respectively.
Following five loci were monomorphic, AAT-1, AAT-2, ALD-1, G3PDH-2, and IDH-1.

indicated these excesses are not significant. For both P and H_e, the highest values were obtained in *C. sinensis* var. *sinensis* (Table 6).

F_{ST} value ranged from 0.002 for FUM to 0.449 for ME. The mean F_{ST} over all loci was 0.046 (Table 7).

Table 6. Genetic variability at 14 loci in the five taxa of *Camellia*.

Taxa	*n*	*A*	*P*	H_o	H_e
ASS	63.9 (1.7)	1.9 (0.2)	35.7	0.176 (0.076)	0.164 (0.062)
CRA	2.6 (0.2)	1.3 (0.1)	28.6	0.083 (0.045)	0.140 (0.063)
DEH	3.0 (0.0)	1.3 (0.2)	21.4	0.143 (0.084)	0.114 (0.064)
SIN	12.5 (0.9)	1.6 (0 .2)	50.0	0.191 (0.079)	0.218 (0.068)
TAL	7.6 (0.2)	1.5 (0 .2)	35.7	0.154 (0.091)	0.178 (0.070)

See Table 1 for the abbreviations of the taxa. n = mean sample size; A = mean number of alleles per locus; P = proportion of polymorphic loci; H_o = observed mean heterozygosity per locus; H_e = expected mean heterozygosity per locus (Nei 1978; unbiased estimate). Standard deviations in parentheses.

Table 7. Summary of *F* statistics at all polymorphic loci for five taxa of *Camellia*.

Locus	F_{IS}	F_{ST}	F_{IT}
AAT-3	0.4196	0.0402	0.4430
FUM-1	−0.0075	0.0022	−0.0053
G3PDH-1	−0.1599	0.0510	−0.1007
G6PDH-1	−0.6159	0.0109	−0.5982
ME-1	1.0000	0.4488	1.0000
6PGDH-1	−0.0438	0.0161	−0.0270
PGM-1	−0.0075	0.0022	−0.0053
PGM-2	0.0994	0.0218	0.1191
SKD-1	0.0088	0.0116	0.0203
Mean	−0.0350	0.0463	0.0129

Similarity among taxa

The similarities among taxa were examined by both morphology and allozymes (Figure 1). Modified Rogers' distance (D) based on the allozyme data ranged from 0.983 to 0.996 among *C. sinensis* var. *dehungensis*, *C. taliensis*, *C. sinensis* var. *assamica* and *C. sinensis* var. *sinensis*. *Camellia crassicolumna* was consistently far from other taxa and *C. sinensis* var. *assamica* was most tight with *C. taliensis*. These can be seen in the UPGMA dendrogram derived from the allozyme data (Figure 1b). The pattern for *C. crassicolumna* to other four taxa and the high similarity of *C. sinensis* var. *assamica* and *C. taliensis* was also exposed by morphological and biochemical traits (Figure 1a).

Discussion

Previous studies have shown that allozyme polymorphism was limited to a few enzyme loci in tea,

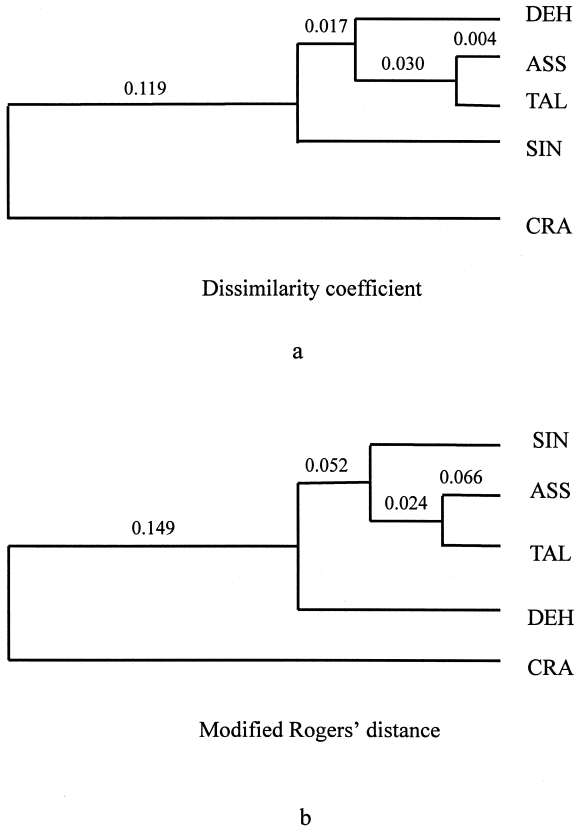

Figure 1. (a). Dendrogram obtained for the five taxa after cluster analysis of 55 morphological characters and four biochemical data. (b). The relationship of the five taxa based on protein electrophoresis. Dendrogram was constructed by the UPGMA method of cluster analysis. Cophenetic correlation = 0.963. See Table 1 for the abbreviations of the taxa. Branch lengths are shown above each branch.

hence it could not be used to efficiently analyze diversity (Ikeda et al. 1991; Lu et al. 1992). We tested a total of 30 different enzymatic systems and obtained 10 systems with 14 loci that presented clear resolution in this study. Among the 14 loci detected in this study, nine loci were found to be polymorphic (Table 5), and is rather informative to reveal genetic diversity. Some patterns of the genetic identity of the five taxa indicated by allozyme were consistent with the results based on the clustering of morphological and biochemical traits (Figure 1). To our knowledge, this is the first report to show the promising results of using allozymes for the investigation of genetic diversity in tea plants. With our results, we may conclude that enzymatic analysis is also an effective technique for the tea investigation of diversity.

Most of the variation was found within taxa rather than among taxa in our study. The fixation index (F_{ST}) was 0.046 (Table 7), which was extremely low, indicating that only 4.6% of variation is ascribable to genetic difference among taxa. This is much lower than results from other studies based on AFLPs and RAPDs marker. Wachira et al. (2001) reported that 72% of variation resided among individuals within populations of *Camellia sinensis* and its wild *Camellia* relatives based on the RAPD and AFLP markers. Kaundun and Park (2002) stated that 16% of the total diversity of RAPD-PCR markers was observed among populations of Korean tea [*Camellia sinensis* (L.) O. Kuntze].

Camellia taliensis was considered as a separate species based on the morphological character of styles by most taxonomists, which are divided into five distinct horizontal arms at the distal end (Sealy 1958; Chang and Bartholomew 1984; Banerjee 1992; Ming 2000). In contrast to other studies (Wachira et al. 1997), *C. taliensis* showed very high similarity to *C. sinensis* var. *assamica* in this study (Figure 1, Table 4). The reason for the high similarity of *C. taliensis* to *C. sinensis* var. *assamica* may be interpreted by a possibly genetic exchange between the two taxa in this area, and the *C. taliensis* in this study may not be the archetypal species. A relatively far distance between *C. crassicolumna* and the other four taxa (Figure 1) found in this study is in agreement with the results revealed by biochemical numerical analysis (Du et al. 1990).

Camellia sinensis var. *assamica* is the taxa historically cultivated in the Yunnan province. Other taxa, such as *C. taliensis*, *C. sinensis* var. *dehungensis* and *C. crassicolumna*, named by local people as 'wild tea', were often used as substitutes of *C. sinensis* var. *assamica* (authors' observation). Cultivated plants often lose genetic diversity due to human selection, which may be revealed by allozymes as a decrease of heterozygosity (H) in comparison to their nearest wild relatives (Doebley 1989). In this study, H value of *C. sinensis* var. *assamica* was not significantly lower than the other wild tea species (Table 6). The reason for that requires further study. Nonetheless, human selection appeared to influence several morphological

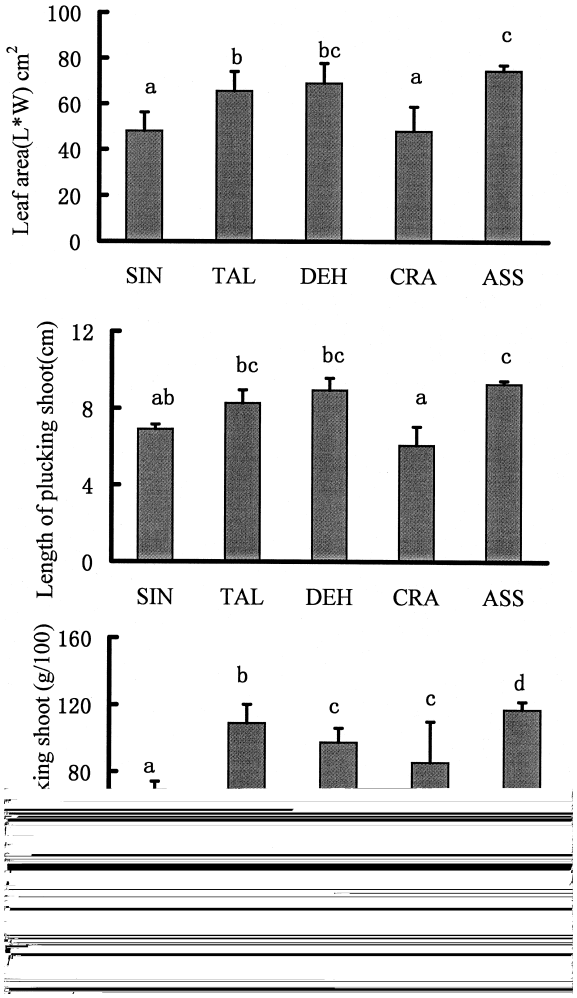

Figure 2. Comparison of leaf area, length of plucking shoot and weight of plucking shoot of five taxa. SIN = *C. sinensis* var. *sinensis*, TAL = *C. taliensis*, DEH = *C. sinensis* var. *dehungensis*, CRA = *C. crassicolumna*, ASS = *C. sinensis* var. *assamica*. Bars indicate ± 1 SE of the mean. Means with a common letter do not differ from other means ($P \leq 0.05$).

characters of *C. sinensis* var. *assamica*. The leaf area of *C. sinensis* var. *assamica* was significantly larger than the other taxa except *C. sinensis* var. *dehungensis*. The plucking-shoot weight of *C. sinensis* var. *assamica* reached significantly larger value than the other four taxa (Figure 2).

Different types of tea (for example, *C. sinensis* var. *assamica* versus *C. sinensis* var. *sinensis*) may have different biochemical contents and those are important factors that contribute to tea quality (Ukers 1935; Wright and Gilchrist 1961; Takeda 1994). *Camellia sinensis* var. *assamica* showed a high content of polyphenols when compared to *C. sinensis* var. *sinensis*, as in other studies (Wright and Gilchrist 1961; Takeda 1994; Magoma et al. 2000). Certain amounts of Caffeine and free amino acids observed in *C. crassicolumna* contrasted with the results by Du et al. (1990), which showed only trace amounts of the two chemicals in *C. crassicolumna.*

Although a lot of effort has been invested in the tea germplasm conservation (Wachira et al. 2001; Zeng et al. 2001), proper strategies still need to be implemented to meet the need of conservation in this area. Due to the high genetic diversity of populations, *ex situ* conservation should have a sufficient sample size for each variety as good representatives of the genetic base. Collections of tea germplasm should aim to include the complete range of geographic distribution. Meanwhile, *in situ* conservation plans need to be initiated in the Yunnan province. The target areas should also include those 'primitive places' (Zeng et al. 2001), where some hundred years old 'low productive' tea gardens in forest are being replaced by commercial tea plantations. The *In situ* conservation area, either focusing on different taxa of *Camellia* or the primitive tea gardens, should make strict exclusion of commercial tea plants. Hybridization among different taxa of *Camellia* is probable (Takeda 1990), and the possible hybridization may cause genetic introgressions to those archetypal genetic resources in the primitive tea gardens as genetic introgressions to local varieties of other crops have been noticed (Levin 2002).

Acknowledgements

The authors gratefully acknowledge Liqiong Yin, Ling Zhang, Zhiling Bai and Hui Xiang for their assistance with field work; Shungao Zhang, Zaifu Xu and Yunqui Wang for stimulating discussion or comments; Ken Marr for technical advisor on allozyme study; and Venetia Briggs for improving the language. Jin Chen thank Theodore H Fleming for arranging his visit to the University of Miami that make he can use the university library for preparing the manuscript. Finally, the authors thank two anonymous referees for their criticisms and helpful comments on the manuscript.

Appendix 1. Fifty-five morphological characters evaluated in this study.

	Score				
Characters	1	2	3	4	5
Habitat					
1. Life form	Tree	Small tree	Shrub		
2. Tree shape	Pubescent	Semi-pubescent	Densely villous		
Leave and shoot					
3. Plucking point density	Dense	Mediate	Sparse		
4. Date for bud first flush: month + (date/days of the month)					
5. Bud color	Green	Yellowish/light green	Green pigmented with anthocyanin		
6. Bud hairiness	Absent	Few	Mediate	A lot	Plenty
7. Length of plucking shoot: Length of third leaf from the apical bud of growing shoot (cm)					
8. Weight of plucking shoot: Weight of flush shoot of three leaf and bud (g/100).					
9. Leaf shape:	Elliptic	Oblong	Obovate	Oblanceate	
10. Leaf color	Dark green	Green	Yellowdish green	Green pigmented with anthocyanin	
11. Leaf texture	Brittle and rigid	Rigid	Less rigid	Pliable	
12. Leaf apex	Obtuse	Pointless	Acute	Straight	Acumen obtuse
13. Leaf margin serration	Deep	Mediate	Flat		
14. Leaf margin serration	Dense	Mediate	Sparse		
15. Acutance of Leaf margin	Acute	Mediate	Obtusely.		
16. Petiole hairiness	Absent	Few	Some	A lot	Plenty
17. Main vine hairiness	Absent	Few	Some	A lot	Plenty
18. Vestiture on lower surface	Absent	Few	Some	A lot	Plenty
19. Leaf length (L) (cm)					
20. Leaf breath (B) (cm)					
21. Leaf area (L*B)(cm^2)					
22. Leaf shape index (L/B)					
23. Number of leaf vine					
24. Pubescence on flush shoot	Absent	Few	Some	A lot	Plenty
25. Pubescence on squama	Absent	Few	Some	A lot	Plenty
Flower					
26. Date of blooming: month + (date/days of the month)					
27. Duration of blooming: Total days/30					
28. Pedicel length (mm)					
29. Pubescence on pedicel	Absent	Few	Some	A lot	Plenty
30. Sepal color	Green	Light green	Green pigmented with anthocyanin		
31. Pubescence on sepal	Absent	Few	Some	A lot	Plenty
32. Sepal diameter (mm)					
33. Flower diameter (mm)					
34. Petal length (mm)					
35. Petal color	White	White mixed with green	Pink		
36. Petal texture	Thin	Mediate	Thick		
37. Pubescence on petal	Absent	Few	Some	A lot	Plenty

Appendix 1. (Continued)

Characters	Score				
	1	2	3	4	5
38. Pubescence on ovary	Absent	Few	Some	A lot	Plenty
39. Length of column (mm)					
40. Styles number	2	3	4–5	6	
41. Position of styles divided on column	Very shallow	Shallow	Middle	Bottom	
42. Height of column compare to stamen	Column > stamen	Equal	Column < stamen		
43. Stamen location style	Connected	Scattered			
Fruit and seed					
44. Fruit set	Absent	Few	Mediate	Many	
45. Fruit size in diameter (mm)					
46. Locule per fruit					
47. Fruit color	Green	Greenish brown	Slightly purple	Red	Dark red
48. Pericarp thickness (mm)					
49. Fruit peduncle length (mm)					
50. Fruit peduncle diameter (mm)					
51. Seed ripe time: month + (date/days of the month)					
52. Seed shape	Obovoid	Globular	Asymmetrical		
53. Seed size in diameter (mm)					
54. Seed color	Brown	Dark brown			
55. Seed weight (g/1000)					

References

Banerjee B. 1992. Botanical classification of tea. In: Wilson K.C. and Clifford M.N. (eds), Tea: Cultivation to Consumption, Chapman and Hall, London, pp. 25–51.

Chang H.T. and Bartholomew B. 1984. Camellia. B.T. Batsford, London.

Chen X.Y. 1994. The Original Locality of Tea Plant. Yunnan People's Press, Kunming Yunnan, China (in Chinese).

Doebley J. 1989. Isozymic evidence and the evolution of crop plants. In: Soltis D.E. and Soltis P.S. (eds), Isozymes in Plant Biology, Dioscorides, Portland, OR, pp. 165–191.

Du Q.Z., Li M.J. and Liu W.H. 1990. Chemical and numerical taxonomies of plants *Thea* section plants. J. Tea Sci. 10: 1–12 (in Chinese with English abstract).

Evans J.C. 1992. Tea in China. Greenwood Press, 88 Post Road West, Westport, pp. 1–22.

Hebert P.D.N. and Beaton M.J. 1993. Methodologies for Allozyme Analysis using Cellulose Acetate Electrophoresis: A practical Handbook. (An educational service of Helena Laboratories).

Ikeda N., Kawada M. and Takeda Y. 1991. Isozymic analysis of *Camellia sinensis* and its interspecific hybrids. In: Proceedings of the International Symposium of Tea Science, held at Shizouka, Japan, 26–28 August 1991, pp. 98 (Abstract).

Kaundun S.S. and Park Y.G. 2002. Genetic structure of six Korean tea populations as revealed by RAPD-PCR markers. Crop Sci. 42: 594–601.

Kaundun S.S., Zhyvoloup A. and Park Y.G. 2000. Evaluation of the genetic diversity among elite tea (*Camellia sinensis* var. *sinensis*) accessions using RAPD markers. Euphytica 115: 7–16.

Lai J.A., Yang W.C. and Hsiao J.Y. 2001. An assessment of genetic relationships in cultivated tea clones and native wild tea in Taiwan using RAPD and ISSR maskers. Bot. Bull. Acad. Sinica 42: 93–100.

Lee S.H., Choi H.S., Kim R.S., Lee H.Y. and Nou I.S. 1995. Identification of Korean wild tea plants and Japanese green tea cultivars using RAPD markers. J. Korean Tea Sci. 1: 129–148.

Levin D.A. 2002. Hybridization and extinction. Am. Sci. 90: 254–261.

Li H.L. 1983. The domestication of plants in China: ecogeographical considerations. In: Keightley D.N. (ed.), The Origins of Chinese Civilization, University of California Press.

Lu C.Y., Liu W.H. and Li M.J. 1992. Relationship between the evolutionary relatives and the variation of esterase isozymes in tea plant. J. Tea Sci. 12: 15–20 (in Chinese with English abstract).

Magoma G.N., Wachira F.N., Obanda M., Imbuga M. and Agong S.G. 2000. The use of catechins as biochemical markers in diversity studies of tea (*Camellia sinensis*). Genet. Resour. Crop Evol. 47: 107–114.

Matsumoto S., Takeuchi A., Hayatsu M. and Kondo S. 1994. Molecular cloning of phenylalanine ammonia lyase cDNA and classification of varieties and cultivars of tea plants

(*Camellia sinensis*) using the tea PAL cDNA probe. Theor. Appl. Genet. 89: 671–675.

Ming T.L. 2000. Monograph of the Genus *Camellia*. Yunnan Science and Technology Press, Kunming, Yunnan, China, pp. 110–135 (in Chinese with English abstract).

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of indivisuals. Genetics 89: 583–590.

Paul S., Wachira F.N., Powell W. and Waugh R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. Theor. Appl. Genet. 94: 255–263.

Sao W.F. and Sheng P.H. 1993. The brief history and characteristic of Pu'er Tea in Yunnan province. Chinese Agr. Archaeol. (supplement issue for Chinese tea culture), 50–54 (in Chinese).

Sealy J. 1958. A Revision of the Genus *Camellia*. Royal Horticultural Society, London.

Song W.X. and Li D.P. 1990. Din Hai Yu Heng Zhi Jiao Zhu. Yunnan People's Press, Kunming, Yunnan, China (in Chinese).

Swofford D.L. and Selander R.B. 1989. BIOSYS-1: a computer program for the analysis of allelic variation in population genetics and biochemical systematics. Illinois Natural History Survey, Champaign, IL, USA.

Takeda Y. 1990. Cross compatibility of tea (*Camellia sinensis*) and its allied species in the genus *Camellia*. Jpn. Agrric. Res. Q. 24: 111–116.

Takeda Y. 1994. Differences in caffeine and tannin contents between tea cultivars and application to tea breeding. Jpn. Agric. Res. Q. 28: 117–123.

Toyao T. and Takeda Y. 1999. Studies on geographical diversity of floral morphology of tea plant (*Camellia sinensis* (L.) O. Kuntze) using the method of numerical taxonomy. Tea Res. J. 87: 39–57.

TLPMBand ITP (Tea and Livestock Products Management Bureau and Institute of Tea Processing, the Ministry of Commerce), 1989. The Physiochemical Analysis of Tea. Shanghai Science and Technology Press, Shanghai, China (in Chinese).

Ukers W.H. 1935. All About Tea. Tea and Coffee Trade Journals, New York.

Wachira F., Tanaka J. and Takeda Y. 2001. Genetic variation and differentiation in tea (*Camellia sinensis*) germplasm revealed by RAPD and AFLP variation. J. Hortic. Sci. Bio. Tech. 76: 557–563.

Wachira F., Wauch R., Hachett C.A. and Powell W. 1995. Detection of genetic diversity in tea (*Camellia sinensis*) using RAPD markers. Genome 38: 201–210.

Wachira F.N., Powell W. and Waugh R. 1997. An assessment of genetic diversity among *Camellia sinensis* L. (cultivated tea) and its wild relatives based on randomly amplified polymorphic DNA and organelle specific STS. Heredity 78: 603–611.

Weatherstone J. 1992. Historical introduction. In: Willson K.C. and Clifford M.N. (eds), Tea: Cultivation to Consumption, Chapman and Hall, London.

Weeden N.F. and Wendel J.F. 1989. Genetics of plant isozymes. In: Soltis D.E. and Soltis P.S. (eds), Isozymes in Plant Biology, Dioscorides Press, Portland, OR, pp. 46–72.

Wichremaratne M.R. 1981. Variation in some leaf characteristics in tea (*Camellia sinensis* L.) and their use in the identification of clones. Tea Q. 50: 183–189.

Wright S. 1978. Evolution and the Genetics of Populations, vol. 4, Variability Within and Among Natural Populations. University of Chicago Press, Chicago, IL.

Wright W. and Gilchrist R.C.J.H. 1961. The concept of kind of tea. Nature 191: 14–16.

Xiang D. 1962. Man Shu Jiao Zhu. Zhong Hua Shu Ju, Beijing, China (in Chinese).

Yunnan Statistic Bureau, 2000. Yunnan Annual Statistic Report. China Statistic Press, Beijing, China, 222 pp. (in Chinese).

Zeng Y.W., Wang J.J., Yang Z.Y., Shen S.Q., Wu L.H., Chen X.Y. and Meng J.G. 2001. The diversity and sustainable development of crop genetic resources in the Lancang River Valley. Genet. Resour. Crop Evol. 48: 297–306.