Journal of Food and Drug Analysis

Volume 31 | Issue 3 Article 5

SSR-based molecular diagnosis for Taiwan tea cultivars and its application in identifying cultivar composition of the processed tea

ORIGINAL ARTICLE

SSR-based molecular diagnosis for Taiwan tea cultivars and its application in identifying cultivar composition of the processed tea

Chih-Yi Hu [a], Hsien-Tsung Tsai [a], Chui-Feng Chiu [a], Tsung-Chen Su [a], Nguyen Hoang Khoi Le [b], Shu-Dan Yeh [b,*]

[a] Tea Research and Extension Station, Yangmei, Taoyuan City, 326011, Taiwan
[b] Department of Life Sciences, National Central University, Taoyuan City, 320317, Taiwan

Abstract

Taiwan specialty teas are produced with distinct manufacturing processes from specific cultivars of tea plants in *Camellia*. Due to the widespread transplantation of Taiwan tea cultivars and active international trading of tea materials, an accurate and reliable method to identify tea cultivars at the border is vital to protect the image of premium Taiwan specialty teas. In this study, we introduced the Taiwan Tea Variety Identification (TTVID) kit, a capillary electrophoresis-based multiplex PCR assay consisting of 12 simple sequence repeat (SSR) markers. A database composing these 12 SSR loci genotypes in 144 cultivars was established for marker assessment and molecular diagnosis. The power of discrimination on a locus ranged from 0.7894 to 0.966 and the combined match probability of 12 SSR loci was 5.34e-14. Cultivar pairwise comparison among 144 accessions showed that over 90.6% of the pairs had differential genotypes on at least 10 of 12 SSR loci. Further assessment showed that the TTVID kit could unambiguously recognize the cultivars mixed in the loose-leaf teas processed with various degrees of fermentation and roasting. Our results suggested that this TTVID kit effectively identified cultivar composition in loose-leaf tea and is helpful for border control in preventing adulteration and fraud in the Taiwan tea market.

Keywords: *Camellia*, DNA fingerprinting, Food security, Microsatellite repeats

1. Introduction

Tea, made of *Camellia sinensis*, is the second most popular nonalcoholic beverage, after water, in the world. The loose-leaf tea produced in Taiwan is known for its exceptional quality attributed to tea cultivars and tea processing, and was the largest export commodity in Taiwan [1,2]. Various tea cultivars with superior traits that are suitable for manufacturing specialty teas were selected and developed by the Tea Research and Extension Station (TRES) in Taiwan and local tea growers since tea plants were introduced to Taiwan in the 1800s [3]. These so-called Taiwanese tea cultivars are classified into three taxonomies, *C. sinensis* var. *sinenesis*, *C. sinensis* var. *assamica*, and *Camellia formosensis*, which is native to Taiwan. Some elite cultivars, including V019, V020, V022, and V025 (Table 1), are protected by the "Plant Variety and Plant Seed Act" enacted in 2004 in Taiwan. Nonetheless, transplantation in other countries has occurred in several elite cultivars developed before 2004, such as TTES No. 12 (V012 in Table 1), commonly used to produce several Taiwan specialty teas.

Rapid changes in consumption habits and culture resulted in an increasing demand for tea products in Taiwan and a negative trade balance of loose-leaf teas since 1991 [2]. More than 30,000 metric tons of processed tea were imported to Taiwan each year in the past five years from neighboring producers to manufacture commodity teas for local consumption, such as bottled tea, ready-to-drink iced tea, and hand-shaking tea [4]. In contrast to commodity teas, Taiwan

Received 2 February 2023; accepted 17 May 2023.
Available online 31 August 2023

* Corresponding author at: No. 300, Zhongda Rd., Zhongli District, Taoyuan City, 320317, Taiwan. Fax: +886 3 4228482.
E-mail address: shudanyeh@ncu.edu.tw (S.-D. Yeh).

https://doi.org/10.38212/2224-6614.3465

Table 1. List of 144 cultivars used in this research.

Cultivar code	Cultivar Name	Classified group[a]	Origin	Sampling plantation[b]
V001	TTES No.1	Sinensis	Hybrid	Wenshan
V002	TTES No.2	Sinensis	Hybrid	Wenshan
V003	TTES No.3	Sinensis	Hybrid	Wenshan
V004	TTES No.4	Sinensis	Hybrid	Wenshan
V005	TTES No.5	Sinensis	Hybrid	Wenshan
V006	TTES No.6	Sinensis	Hybrid	Wenshan
V007	TTES No.7	Assamica	Hybrid	Yuchi
V008	TTES No.8	Assamica	Hybrid	Wenshan
V009	TTES No.9	Sinensis	Hybrid	Wenshan
V010	TTES No.10	Sinensis	Hybrid	Wenshan
V011	TTES No.11	Sinensis	Hybrid	Wenshan
V012	TTES No.12	Sinensis	Hybrid	Wenshan
V013	TTES No.13	Sinensis	Hybrid	Wenshan
V014	TTES No.14	Sinensis	Hybrid	Wenshan
V015	TTES No.15	Sinensis	Hybrid	Yangmei
V016	TTES No.16	Sinensis	Hybrid	Yangmei
V017	TTES No.17	Sinensis	Hybrid	Wenshan
V018	TTES No.18	Assamica	Hybrid	Yuchi
V019[c]	TTES No.19[c]	Sinensis	Hybrid	Wenshan
V020[c]	TTES No.20[c]	Sinensis	Hybrid	Wenshan
V021	TTES No.21	Sinensis	Hybrid	Wenshan
V022[c]	TTES No.22[c]	Sinensis	Hybrid	Yangmei
V023	TTES No.23	Sinensis	Hybrid	Yuchi
V024	TTES No.24	Formosana	Hybrid	Taitung
V025[c]	TTES No.25[c]	Assamica	Hybrid	Yuchi
H001	FKK No.1	Assamica	Selected	Yangmei
H002	Tainon No.2028	Sinensis	Selected	Yangmei
H003	Tainon No.2022	Sinensis	Selected	Yangmei
H004	Tainon No.8	Sinensis	Selected	Yangmei
H005	Tainon No.20	Sinensis	Selected	Yangmei
H006	Tainon No.335	Sinensis	Selected	Yangmei
H007	Tainon No.983	Sinensis	Selected	Yangmei
H008	Tainon No.1958	Sinensis	Selected	Yangmei
H010	Tainon No.399	Sinensis	Selected	Yangmei
H011	Tainon No.483	Sinensis	Selected	Yangmei
H012	Tainon No.1511	Sinensis	Selected	Yangmei
H013	Tainon No.2044	Assamica	Selected	Yangmei
H014	64–5312	Sinensis	Selected	Yangmei
H016	72–209	Sinensis	Selected	Private plantation
F001	Yin-Long	Sinensis	Landrace	Private plantation
L001	Chin-Shin-Oolong	Sinensis	Landrace	Wenshan
L002	Shy-Jih-Chuen	Sinensis	Landrace	Wenshan
L003	Chin-Shin-Dahpan	Sinensis	Landrace	Yangmei
L003T	Chin-Shin-Dahpan-2	Sinensis	Landrace	Yangmei
L004	Chin-Shin-Gantzy	Sinensis	Landrace	Yangmei
L005	Ying-Jy-Horng-Shin	Sinensis	Landrace	Wenshan
L006	Horng-Shin-Gantzy	Sinensis	Landrace	Wenshan
L006F	Dah-Yeh-Oolong	Sinensis	Landrace	Yangmei
L007	Hwang-Gan	Sinensis	Landrace	Yangmei
L008	Bair-Mau-Hour	Sinensis	Landrace	Wenshan
L009	Horng-Shin-Dahpan	Sinensis	Landrace	Yangmei
L010	Horng-Shin-Wuu-Yi	Sinensis	Landrace	Yangmei
L012	Han-Xiao	Sinensis	Landrace	Yangmei
L013	Ying-Zhi-Zao-Zhong	Sinensis	Landrace	Yangmei
L014	Bu-Zhi-Chun	Sinensis	Landrace	Yangmei
L015	Tao-Ren-Zhong	Sinensis	Landrace	Yangmei
L016A	Tian-Gong	Sinensis	Landrace	Yangmei
L016B	Xiao-Ye-Zhu-Ye	Sinensis	Landrace	Yangmei
L017	Zao-Zhong	Sinensis	Landrace	Yangmei
L018	Mao-Zi-1	Sinensis	Landrace	Yangmei
L018W	Mao-Zi-2	Sinensis	Landrace	Wenshan

(*continued on next page*)

Table 1. (continued)

Cultivar code	Cultivar Name	Classified group[a]	Origin	Sampling plantation[b]
L018Y	Mao-Zi-3	Sinensis	Landrace	Yangmei
L019	Qing-Xin-Zao-Zhong	Sinensis	Landrace	Yangmei
L020	Wan-Zhong	Sinensis	Landrace	Yangmei
L021	Niu-Pu	Sinensis	Landrace	Wenshan
L022	San-Cha-Zhi-Lan	Sinensis	Landrace	Yangmei
L023	Ji-Long-Jin-Gui	Sinensis	Landrace	Yangmei
L024	Wu-Gu-Zi	Sinensis	Landrace	Yangmei
L026	Hei-Mao-Hou	Sinensis	Landrace	Wenshan
L028	Tu-Zi-Keng-Bai-Mao-Hou	Sinensis	Landrace	Yangmei
L029	Feng-Lin-Zi	Sinensis	Landrace	Yangmei
L030	Zhi-Lan	Sinensis	Landrace	Yangmei
L031	Xiao-Cu-Keng	Sinensis	Landrace	Yangmei
L032	Jin-Gui	Sinensis	Landrace	Yangmei
L033W	Wuu-Yi	Sinensis	Landrace	Wenshan
L034	Hong-Xin-Oolong	Sinensis	Landrace	Yangmei
L035	Bai-Xin-Oolong	Sinensis	Landrace	Yangmei
L037	Da-Hu-Wei	Sinensis	Landrace	Yangmei
L039	Tao-Ren-Wu	Sinensis	Landrace	Yangmei
L040	Niu-Shi-Wu	Sinensis	Landrace	Yangmei
L042	Zhu-Ye	Sinensis	Landrace	Yangmei
L043	Bo-Ye	Sinensis	Landrace	Yangmei
L044	Shen-Man-Zhong	Sinensis	Landrace	Yangmei
L046	Huang-Zhi	Sinensis	Landrace	Yangmei
L047	Xiao-Ye-Tie-Guan-Yin	Sinensis	Landrace	Yangmei
L048	Dan-Shui-Qing-Xin	Sinensis	Landrace	Yangmei
L049	Wen-Shan-Zhi-Lan	Sinensis	Landrace	Wenshan
L050	Bai-Ye	Sinensis	Landrace	Wenshan
L051	Qi-Du-Bai-Zhong	Sinensis	Landrace	Wenshan
L052	Zhu-Ye-Wen-Shan	Sinensis	Landrace	Wenshan
L053	Ji-Long-Bai-Zhong	Sinensis	Landrace	Wenshan
L054	Zi-Ya-Cha No.113	Assamica	Selected	Yangmei
L055	Zi-Ya-Cha No.117	Assamica	Selected	Yangmei
L056	TTES No.12 Zi-Ya No.1	Sinensis	Selected	Wenshan
L057	55-23-6	Assamica	Selected	Yuchi
I001	Kimen	Sinensis	Imported	Yuchi
I002	Kimen-2	Sinensis	Imported	Yuchi
I003	Kimen-3	Sinensis	Imported	Yuchi
I004	Kimen-4	Sinensis	Imported	Yuchi
I005	Shiang-Yuan	Sinensis	Imported	Wenshan
I006	Burma	Assamica	Imported	Yangmei
I006F	Burma No.761	Assamica	Imported	Yuchi
I007	Assam	Assamica	Imported	Yangmei
I008	Kyang	Assamica	Imported	Yangmei
I009	Jaipuri	Assamica	Imported	Yangmei
I010	Manipuri	Assamica	Imported	Yangmei
I011	Shan	Assamica	Imported	Yangmei
I012	Hu-Nan	Sinensis	Imported	Yangmei
I013	Pin-Shui	Sinensis	Imported	Yangmei
I014	Fwu-Jou	Sinensis	Imported	Yangmei
I015T	Yabukita-2	Sinensis	Imported	Yangmei
I016T	Suruga	Sinensis	Imported	Yangmei
I017T	Uji	Sinensis	Imported	Yangmei
I018	Oto	Sinensis	Imported	Yangmei
I019	Da-Jin-Zao-Sheng	Sinensis	Imported	Yangmei
I020	Da-Ji-Ling	Sinensis	Imported	Yangmei
I021	Hann-Koou-1	Sinensis	Imported	Yangmei
I021W	Hann-Koou-2	Sinensis	Imported	Wenshan
I022	Rou-Gui	Sinensis	Imported	Yangmei
I023	Qi-Lan	Sinensis	Imported	Yangmei
I024	Ba-Xian	Sinensis	Imported	Yangmei
I025	Shoei-Shian-1	Sinensis	Imported	Yangmei

(*continued on next page*)

Table 1. (continued)

Cultivar code	Cultivar Name	Classified group[a]	Origin	Sampling plantation[b]
I025W	Shoei-Shian-2	Sinensis	Imported	Wenshan
I026	Gao-Lu	Sinensis	Imported	Yangmei
I027	Huang-Jin-Gui	Sinensis	Imported	Yangmei
I028	Mei-Zhan	Sinensis	Imported	Yangmei
I029	Yan-Chuan	Sinensis	Imported	Yangmei
I030	Fu-Yun No.7	Sinensis	Imported	Yangmei
I031	Tie-Guan-Yin	Sinensis	Imported	Wenshan
I032	Qian-Jin-Hong	Assamica	Imported	Wenshan
I033	Yabukita	Sinensis	Imported	Yangmei
I034	Asanoka	Sinensis	Imported	Yangmei
I035	Okumidori	Sinensis	Imported	Yangmei
W001	De-Hua-She wild tea	Formosana	Native	Yuchi
W002	Fong-Huang wild tea	Formosana	Native	Yuchi
W003	Mei-Yuan wild tea	Formosana	Native	Yuchi
W004	Le-Ye wild tea	Formosana	Native	Yuchi
W006	Nan-Fong wild tea	Formosana	Native	Yuchi
W008	Yong-Kang wild tea No.2	Formosana	Native	Taitung
W009	Yong-Kang wild tea No.5	Formosana	Native	Taitung
W010	Yong-Kang wild tea No.13	Formosana	Native	Taitung
W011	Yong-Kang wild tea No.20	Formosana	Native	Taitung
W012	Ming-Hai wild tea	Formosana	Native	Taitung
W013	Shui-Li wild tea	Formosana	Native	Taitung

[a] Sinensis: the cultivars belonging to *C. sinensis* var. *sinensis*. Assamica: the cultivars belonging to *C. sinensis* var. *assamica*. Formosana: the cultivars belonging to *C. formosensis*.

[b] The fresh leaves were collected either from private plantations or the experimental tea plantations of TRES in Wenshan, Yangmei, Yuchi, and Taitung.

[c] The cultivars were registered with plant variety rights based on The Plant Variety and Plant Seed Act, Council of Agriculture, R. O. C.

specialty teas are crafted from whole leaves through delicate fermentation (oxidation to be exact) and roasting processes. Taiwan specialty teas are generally categorized into eight types depending on the manufacturing processes and Taiwanese tea cultivars [2]. Due to the exceptional flavors and excellent qualities, Taiwan specialty teas are usually expensive compared to commodity teas. In recent years, tea fraud has raised concerns over the sustainability of the Taiwanese tea industry. Some Taiwanese tea cultivars, such as TTES No. 12, were cultivated in other countries for lower growing and production costs. The processed products of these tea materials are imported to Taiwan and then reprocessed and labeled as counterfeit Taiwan specialty teas. Such adulteration severely damages the domestic tea industry by affecting the market price and the reputation of premium Taiwan specialty tea. Although the Taiwan government has imposed specific laws and regulations to restrict the importation of tea materials derived from Taiwanese tea cultivars, an efficient approach is required to identify cultivar composition in the processed teas at the border.

Appearances, such as leaf morphology, color, and aroma, are traditionally used to identify cultivars in tea materials and processed teas. However, disguisable properties are insufficient as they fluctuate depending on the environments where the tea plants grow and tea processing. Molecular markers based on genetic information provide more effective and reproducible indicators since external factors do not affect them. Previously, numerous codominant DNA markers, such as cleaved amplified polymorphic sequences (CAPS), single nucleotide polymorphisms (SNP), and insertion-deletion (Indel), were developed to distinguish *Camellia* species and Taiwan tea cultivars [3,5]. These locus-specific markers offer high resolution and specificity but are mostly diallelic in the studied populations. Only using a high number of markers can achieve adequate discriminating power, but it also increases the time and costs to genotype [6]. To overcome these limitations, simple sequence repeats (SSR), also known as microsatellites, are notable for their applications in tea plant genetics due to their multi-allelic nature and high polymorphism [7–9]. SSR markers across the tea genome were frequently used in genetic mapping, marker-assisted selection, and cultivar fingerprinting [9–12]. Several previous studies also showed that SSR markers are helpful to identify the cultivars mixed in processed green teas, which is lightly fermented and non-roasted [13–15]. Nevertheless, an effective molecular diagnosis method that is suitable to identify Taiwanese cultivars in processed tea samples yet to be established, especially in those specialty teas processed with high level of fermentation and roasting.

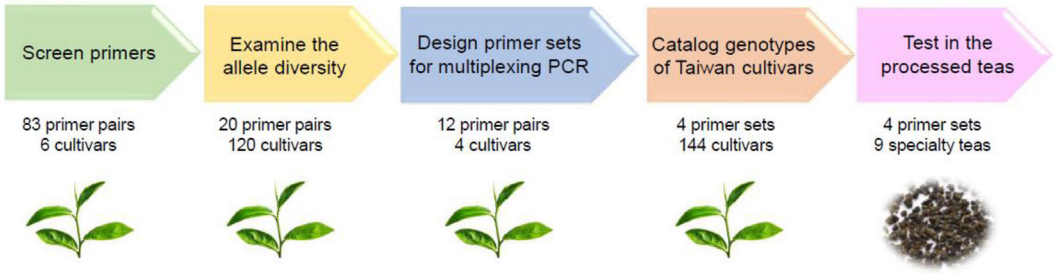

Fig. 1. The procedure and materials used to develop and test Taiwan Tea Variety Identification (TTVID) kit.

The more SSR loci with higher polymorphism offer more power in identifying tea cultivars, but more PCR reactions and allele diagnosing efforts accompany them. In this study, we developed Taiwan Tea Variety Identification (TTVID) kit, a capillary electrophoresis-based multiplex PCR assay consisting of 12 SSR markers screened from previously identified SSR loci (Fig. 1). To reduce the costs and time involved in the repeated single-amplicon-based PCR reactions, we optimized the primer sequence composition and amplicon length to simultaneously amplify three loci in each PCR reaction, known as multiplex PCR. Using the TTVID kit, we established a genotype database of 144 tea cultivars available in Taiwan. Also, we demonstrated the effectiveness of the TTVID kit by examining the cultivars in nine processed tea samples, including all eight types of Taiwan specialty teas. Our analysis suggests that the TTVID kit could effectively identify cultivars in the processed teas and, therefore, can be reliably applied to inspect the imported processed teas.

2. Materials and methods

2.1. Plant materials and DNA isolation

Fresh tea leaves of 144 tea cultivars, including 115 var. *sinensis* accessions, 17 var. *assamica* accessions, and 12 native cultivars of *C. formosensis* were collected from four experimental plantations (Wenshan, Yangmei, Yuchi, and Taitung) of the Tea Research and Extension Station in Taiwan. Detail information on these cultivars is listed in Table 1. The tea leaves were lyophilized and grinded with a 2010 Geno/Grinder (SPEX Sample Prep) before being stored at −80 °C.

A total of nine specialty teas, comprising all eight types of Taiwan specialty teas, were either provided by private manufacturers and the regional Farmers' Association or prepared by TRES between Summer of 2018 and Spring of 2021 (Table 2). The tea samples were made between Summer 2018 to Spring 2021. Single-granular and multi-granular sampling methods were applied to investigate the ability to distinguish cultivars in tea mixtures since some specialty tea samples might be made of more than one cultivar. Tea granules are usually processed from a hand-picked branch with one apical bud and two to three young leaves. For the single-granular sampling, ten singular granules were individually sampled in each specialty tea sample. Two replicates of 10 g tea granules were sampled in each specialty tea sample for multi-granular sampling method. All samples were stored at room temperature before they were ground and stored at −80 °C.

Genomic DNAs were extracted by following either the cetyltrimethylammonium bromide (CTAB) method [16] or the Quick-DNA plant/seed miniprep kit (Cat. No. D6020, Zymo Research, California, USA) for processed tea samples. Genomic DNAs were eluted/resuspended in 50 ul elution buffer and stored at −20 °C.

2.2. SSR screening and multiplexing genotyping

We first tested eighty-three SSR primer pairs adopted from three previous studies on six tea cultivars (L001, L002, V012, V020, V008, and V018) which are commonly grown in Taiwan [7,17,18]. Forward primers were 5′ tagged with 5′ACGACGTTGTAAA3′, following multiplex-ready PCR method [19], and labeled with one of the four fluorescent dyes, 5-FAM, JOE, TAMRA, and ROX. Reverse primers were 5′ tagged with 5′GTTTAAGTTCCCATTA3' [19]. Primer pairs were initially screened on agarose gel electrophoresis and subsequently validated on PRISM 3730 XL DNA analyzer (Applied Biosystems) at Genomics company (New Taipei City, Taiwan) (data not shown).

We further investigated the allele number and polymorphism of the candidate 20 loci by using forward primers directly labeled with either 5-FAM or JOE at the 5' end. Reverse primers were tagged with PIG tailing GTTTCTT [20]. The resulting 12 primer pairs were grouped into four sets for multiplexing PCR (Table 3). All the forward primers were

Table 2. Nine Taiwan specialty tea samples used in this research.

Sample ID	Taiwan specialty tea name	Provider	Year	Appearance	Fermentation level[e]	Roasting level
ST01	Taiwan Green tea (Pi Luo Chun)	Sanxia FA[a]	Spring, 2019	Stripe	None	None
ST02	Wenshan Paochong tea	Pinglin FA[a]	Spring, 2020	Stripe	Light	Light
ST03	High-Mountain Oolong tea	Meishan, Chiayi[b]	Winger, 2020	Rounded[d]	Light to Medium	Light
ST04	Tongding Oolong tea	Lugu FA[a]	Spring, 2020	Rounded[d]	Medium	Medium
ST05	Tieh-Kuan-Yin tea	Mucha FA[a]	Spring, 2020	Rounded[d]	Medium	Heavy
ST06	Black Oolong tea	Luye, Taitung[b]	Spring, 2021	Rounded[d]	Heavy	Medium to Heavy
ST07	Oriental Beauty tea	Longtan, Taoyuan[b]	Summer, 2020	Curly	Heavy	None
ST08	Taiwan Black tea var. *sinensis*	TRES[c] -Yangmei	Summer, 2020	Stripe	Full	None
ST09	Taiwan Black tea var. *assamica*	TRES[c] -Yuchi	Summer, 2018–2019	Stripe	Full	None

[a] FA: Farmers' Association.
[b] Private tea manufacture.
[c] TRES: Tea Research and Extension Station, Taiwan Agriculture Council.
[d] These are so-called ball type.
[e] The term "fermentation level" is used in the communication to the public, tea growers, and tea manufacturers. In scientific speaking, this is an oxidation process in which no enzyme or no microorganism was involved.

labeled with 5-FAM. Multiplexing PCR tests were carried out using Taq polymerase (Fast-Run™ 2x Taq master mix, Protech) and underwent 40 cycles. Amplicons were inspected by agarose gel electrophoreses and a PRISM 3730 XL DNA analyzer (Applied Biosystems) in Genomics company (New Taipei City, Taiwan). Fragment sizes of alleles were estimated using GeneMapper 4.0 software (Applied Biosystems), and the amplicons from TTES No. 12 (V012) and Chin-Shin-Oolong (L001) were included as reference samples in each capillary electrophoresis analysis. The primer sequences, PCR conditions, and images of gel electrophoresis are available upon request.

2.3. Marker validation and data analysis

Genotyping of 144 tea cultivars was performed based on the difference in fragment sizes on each SSR locus for marker assessment. Match probability and power of discrimination were calculated as follows.

Match probability, $PM = \sum_{i=1}^{n} Gi^2$.where Gi is the allele frequency of allele i of a locus, n is the allele numbers of a locus.

Combined match probability, $CPM = \prod_{i=1}^{m} PMi$. where PMi is the match probability of a locus, m is the number of loci.

Power of discrimination, $PD = 1\text{-}PM$.

Combined power of discrimination, $CPD = 1\text{-}\prod_{i=1}^{m}(1 - PDi)$. where PDi is the PD of locus i and m is the number of loci.

All metrics were calculated using FORSTAT and STRAF software developed for forensic science [21,22]. We performed pairwise comparisons between cultivar genotypes to further evaluate the discriminating power of SSR markers. Cultivar identification of single and multi-granular samples

Table 3. The information of 12 SSR loci in TTVID kit, divided into the four sets, A, B, C, and D, for triplexing PCR.

Sets	SSR Locus	SSR motif	Min size (bp)	Max size (bp)	Genomic location	Source
A	A01	TTC	117	176	Chr9: 6,992,031 . 6,992,164	Hu et al., 2011
A	A02	ATTTTT	184	215	Chr11: 105,103,918 . 105,104,121	Ma et al., 2014
A	A03	CACCAT	238	279	Chr13: 42,973,238 . 42,973,514	Tal et al., 2013
B	B01	ATC	144	177	Chr9: 193,961,755 . 193,961,903	Tal et al., 2013
B	B02	TG	206	242	Chr4: 7,722,517 . 7,722,728	Tal et al., 2013
B	B03	ATACAC	263	289	Chr7: 88,946,853 . 88,947,127	Ma et al., 2014
C	C01	GA	112	143	Chr8: 41,514,686 . 41,514,803	Tal et al., 2013
C	C02	TTC	151	184	Chr6: 209,938,026 . 209,938,190	Tal et al., 2013
C	C03	TTC	209	264	Chr15: 138,974,946 . 138,975,164	Tal et al., 2013
D	D01	CCT	98	120	Chr4: 109,393,531 . 109,393,635	Tal et al., 2013
D	D02	AG	159	181	Chr7: 137,926,587 . 137,926,752	Tal et al., 2013
D	D03	GA	242	274	Chr3: 222,810,759 . 222,811,009	Tal et al., 2013

The reference genome for identifying genomic location of 12 SSR loci is *C. sinensis* var. *sinensis* cultivar Tieguanyin (Accession number in GenBank: JAFLEL000000000.1).

Table 4. Allele number, match probability and power of discrimination of 12 SSR loci in 144 tea cultivars. The best and worst numbers of each column are blue and red colored, respectively.

Locus	Allele #	Match probability[b]	Power of discrimination[c]			
			Total (144)	Sinensis (115)	Assamica (17)	Formosana (12)
A01	11	0.0751	0.9249	0.9075	0.8789	0.8611
A02	11	0.108	0.8920	0.8767	0.6713	0.7639
A03	11	0.0883	0.9167	0.8995	0.9135	0.7500
B01	10	0.0918	0.9082	0.8826	0.8581	0.7500
B02	17	0.0334	0.9666	0.9595	0.8858	0.4167
B03	7	0.2106	0.7894	0.7481	0.8028	0.6806
C01	13	0.0508	0.9492	0.9319	0.8927	0.8472
C02	13	0.0827	0.9173	0.8868	0.9204	0.7847
C03	19	0.0381	0.9619	0.9494	0.9273	0.8889
D01	7	0.0891	0.9109	0.8798	0.8843	0.7778
D02	12	0.1358	0.8642	0.8210	0.8754	0.7915
D03	15	0.0632	0.9468	0.9193	0.8304	0.8750
Average/ Combined[a]	12.2	5.34E-14	0.99999999999995	0.99999999997755339987	0.99999999192301602946	0.99999999999907185355

[a] The allele number is the average of 12 loci; the match probability and power of discrimination are multiplications of 12 loci as the combined values.

[b] Match probability (*PM*) equals to $\sum_{i=1}^{n} Gi^2$, where Gi is the allele frequency of allele *i* on a locus, *n* is the allele numbers of a locus. Combined match probability (*CPM*) equals to $\prod_{i=1}^{m} PMi$, where *PMi* is the match probability of a locus, *m* is the number of the markers (loci).

[c] Power of discrimination (*PD*) equals to 1-*PM*. Combined power of discrimination (*CPD*) equals to 1- $\prod_{i=1}^{m}(1 - PDi)$, where *PDi* is the *PD* of locus *i* and *m* is the number of markers (loci).

was performed based on the genotype database of 144 tea cultivars.

3. Results

3.1. TTVID kit and genotyping of tea cultivars

To select the SSR markers, we followed the procedure in Fig. 1. From our screening of 83 primer pairs (targeting to 83 SSR loci) in six cultivars, including four var. *sinensis* and two var. *assamica*, a total of 20 primer pairs met the criteria of amplicons concentration, polymorphism information content (PIC), and motif length. These primers were further investigated for their allele numbers and frequencies in 120 Taiwan cultivars (data not shown). Finally, 12 SSR loci situating 9 out of 15 chromosomes in *C. sinensis* genome were selected for TTVID kit (Table 3).

To assess TTVID kit's effectiveness in distinguishing cultivars, we calculated genetic parameters of allele number, allele frequency, PM, CPM, PD, and CPD based on the genotype profile of 12 SSR loci in 144 tea cultivars (Table 4). 146 alleles were detected, ranging from 7 (B03 and D01) to 19 (C03) per SSR locus, with an average number of 12.2. The lowest PM value, 0.0344, was observed in the B02 locus, which exhibited the largest allele number. B02 locus offers as high as 96.66% of confidence to assign a cultivar identity among 144 tea cultivars. The CPM of 12 SSR loci was 5.34e-14, meaning that the CPD of 12 SSR loci is higher than 0.99999999999995. To compare the power to identify cultivars with high genetic similarity, we further classified 144 tea accessions into three groups: Sinensis, Assamica, and Formosana, representing the cultivars of *C. sinensis* var. *sinensis*, *C. sinensis* var. *assamica* and *C. formosensis*, respectively. While CPDs of the three groups were all higher than 0.99999999, Assamica group had the lowest CPD.

We also evaluated the effectiveness of the TTVID kit by comparing genotypes of cultivars pairwise. In this analysis, a differential locus refers a locus showing different genotypes between two compared cultivars. All cultivars showed at least 1 differential locus to another cultivar in a total of 10,296 pairwise comparisons and 41.3% of cultivar pairs showed 12 differential loci (Fig. 2). At least 10 differential loci were found in 90.6% of cultivar pairs, suggesting that genotype combinations of these 12 SSR loci were sufficient to distinguish tea cultivars in our database.

3.2. DNA extraction from processed tea samples

Although genomic DNAs might degrade during tea processing, successful DNA extraction from loose and bagged teas has been reported previously [5,15,23]. In our attempts, extracted DNA concentration differed among different Taiwan specialty tea samples, mainly depending on the manufacturing processes (Table 5). Low quantities

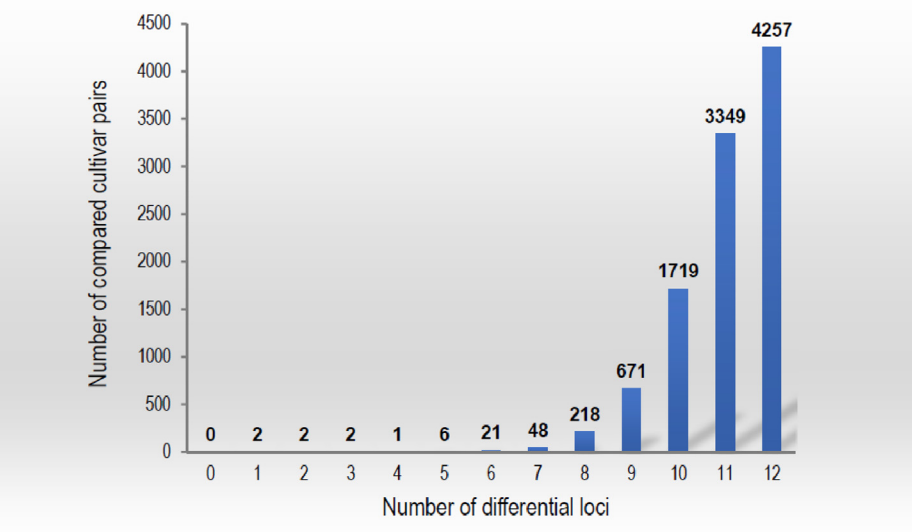

Fig. 2. Frequency distribution of pairwise comparisons among 144 cultivars, based on the genotypes of 12 SSR loci. The possible pairwise comparisons among 144 cultivars are 10,296. The number of cultivar pairs showing the corresponding number of differential loci is labeled on the top of each bar.

of DNAs were extracted from the tea samples manufactured with more fermentation (oxidation), such as Black Oolong and Taiwan Black teas. The quantity and quality of DNAs extracted from processed tea samples were adequate to perform multiplexing PCR using our TTVID kit. However, the strength of chromatographic signals in analyses of capillary electrophoresis varied across SSR loci and samples (Fig. 3). We found that C03 locus showed weak chromatographic signals in all tested samples. Meanwhile, A02 and D01 exhibited the strongest chromatographic signals. ST01, ST02, and ST03 showed strong chromatographic signals across 12 SSR loci, whereas the chromatographic signals of

Table 5. DNA amounts extracted from multigranular and monogranular samplings using Quick-DNA plant/seed miniprep kit. The information of fermentation and roasting level can be found in Table 2. The averages of each measurements and standard errors are listed.

Sample ID	Sample name	Single granule[a]				Multiple granule[b]		
		Mean weight[1] (g)	Purity 260/280	Eluted Conc. (ng/ul)	Yield ratio[c] (ug/g)	Purity 260/280	Eluted Conc. (ng/ul)	Yield ratio[c] (ug/g)
ST01	Taiwan Green tea (Pi Luo Chun)	0.098 ± 0.011	1.81 ± 0.01	258.9 ± 39.6	132.4 ± 16.3	1.87 ± 0.00	309.3 ± 6.2	154.7 ± 3.1
ST02	Wenshan Paochong tea	0.088 ± 0.008	1.86 ± 0.05	77.7 ± 11.8	44.9 ± 6.7	1.90 ± 0.03	191.5 ± 2.8	95.7 ± 1.4
ST03	High-Mountain Oolong tea	0.163 ± 0.007	1.81 ± 0.01	126.2 ± 18.0	38.6 ± 4.9	1.84 ± 0.01	122.5 ± 14.5	61.3 ± 7.3
ST04	Tongding Oolong tea	0.132 ± 0.004	1.80 ± 0.01	100.6 ± 16.4	38.9 ± 6.8	1.84 ± 0.01	143.9 ± 1.5	72.0 ± 0.7
ST05	Tieh-Kuan-Yin tea	0.097 ± 0.004	1.84 ± 0.02	82.8 ± 21.6	42.1 ± 10.8	1.83 ± 0.01	104.3 ± 12.5	52.2 ± 6.3
ST06	Black Oolong tea	0.155 ± 0.004	**1.94 ± 0.11**	**8.8 ± 2.9**	**2.9 ± 1.0**	**2.02 ± 0.15**	**5.1 ± 0.5**	**2.5 ± 0.2**
ST07	Oriental Beauty tea	0.082 ± 0.003	1.86 ± 0.01	76.4 ± 7.6	47.5 ± 5.1	1.83 ± 0.00	57.8 ± 2.0	28.9 ± 1.0
ST08	Taiwan Black tea (var. *sinensis*)	0.073 ± 0.004	1.96 ± 0.05	33.8 ± 8.4	22.8 ± 5.0	1.85 ± 0.00	27.9 ± 1.4	14.0 ± 0.7
ST09	Taiwan Black tea (var. *assamica*)	0.098 ± 0.010	2.10 ± 0.08	20.9 ± 8.7	12.4 ± 5.5	1.78 ± 0.09	**8.6 ± 1.1**	**4.3 ± 0.6**

[a] Monogranular samplings. Each tea granule consists of 1 bud and 2–3 leaves. Ten granules from each tea sample were grinded individually. Various amounts of grinding tea power in 10 replicates were used in DNA isolation, thus the average weight of 10 replicates is listed.

[b] Multigranular samplings. Two replicates of DNA extraction were carried out per tea sample. In each replicate, 10 g of tea granules were grinded into tea power and 0.1 g of tea power was used in DNA isolation.

[c] The yield ratio was calculated from a multiplication of the eluted DNA concentration and 50 ul of final elution volume divided by the weight of tea powder used in each DNA extraction.

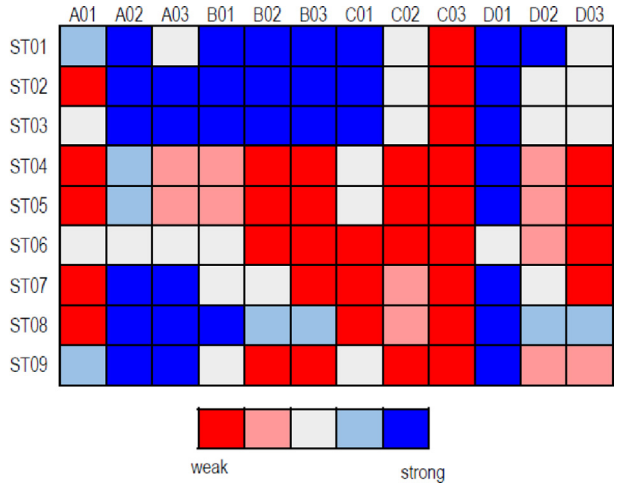

Fig. 3. Genotyping capability of 12 SSR loci in the tested Taiwan specialty teas according to the height and area of the peaks in the capillary electrophoresis chromatogram. High and sharp peaks offered higher certainty of allele identification, while low and vague peaks led to difficult allele callings. Ten monogranular samplings and two multi-granular samplings in each processed tea sample were visually inspected to categorize the genotyping certainty into five scales, ranging from weak to strong, as indicated in red to blue colors.

ST01 more resemble to those of fresh-leaf samples. Overall, chromatographic signals tend to be weak in the tea samples that were processed with higher degrees of o fermentation or roasting.

3.3. Identifying cultivars in Taiwan specialty teas using the TTVID kit

We further tested whether tea cultivars made of Taiwan specialty tea samples can be recognized using our TTVID kit. Since each granule of specialty teas originated from a tea plant, the identified genotype across 12 SSR loci in a single-granular sample was expected to match to a single cultivar perfectly. However, mixed genotype combinations and more than two alleles on one SSR locus might be detected in a multi-granular sample if a processed tea sample made of more than one cultivar. According to the genotyping results of multi-granular samplings, six out of nine tea samples were identified as mono-cultivar products (Fig. 4). The genotypes of 3 specialty teas with none or low fermentation and roasting, ST01, ST02, and ST03, across 12 SSR loci in the multi-granular sampling were unequivocally identified and fully supported by the data of the single-granular sampling. In heavily or fully fermented teas such as ST07, ST08, and ST09, the genotypes of some loci were ambiguous. In ST07, we could not identify the genotypes on the C03 locus in all samples and the genotypes on the C02 locus in four single-granular samples due to the weak chromatographic signals. The same issue also happened on the C01 and C02 loci of some ST08 single-granular samples and the B03, C01, C02, and C03 loci of ST09 samples. Even so, missing genotypes of one to three of these four loci did not affect cultivar identification in these tea samples.

To our surprise, signatures of cultivar mixtures were detected in ST04, ST05, and ST06 (Fig. 4). In single-granular data of ST04, we detected at least four unique genotype combinations and classified 8 samples into either L001 or V012 accession. In contrast, genotypes of the last two samples were not present in our 144-cultivar database. This result matched the multi-granular data, except on C03 and D03 loci. In ST05, the genotyping data of single-granular samples suggested that ST05 might originate from the mixtures of V012 and I031 accessions. However, V012 was dominantly detected in the multi-granular samples of ST05. In addition, two distinct alleles on the A03 locus were identified in the 10th single-granular sample while the rest matched to I031 genotype, indicating possible hidden genetic variation in I031 or a cultivar genetically close to I031. The genotyping quality of ST06 samples exhibited the lowest quality among specialty teas as alleles on C01 and C02 loci were hardly recognizable. Still, we unequivocally assigned eight single-granular samples to be V012 while the other two were undetermined.

4. Discussions

SSRs are one of the most common DNA markers in genetic studies and forensic science due to their high polymorphism, multiallelic nature, and abundance across the genome [8,9]. Although multiple SSR loci of *C. sinensis* were isolated previously, an effective SSR-based DNA fingerprinting method to identify tea cultivars has yet to be well-established [7,17,18]. In this study, we successfully developed a TTVID kit, consisting of four fluorescent-labeled primer sets to detect accurately and unambiguously 12 highly polymorphic SSR loci through multiplex PCR and capillary electrophoresis. Furthermore, a genotype database of these 12 SSR loci in was established in 144 tea cultivars.

Successful amplification and accurate allele calling of a specific SSR locus in each tea sample depend on several factors, such as template DNA integrity, allele frequency, primer specificity to each template DNA, sequence variation in priming sites, and the primer combination in a multiplexing PCR [24]. Each primer pair in our TTVID kit worked well in the separate PCR reaction using the DNA of fresh tea leaves. But PCR amplification efficiency of some primer pairs seemed compromised in a multiplexing PCR,

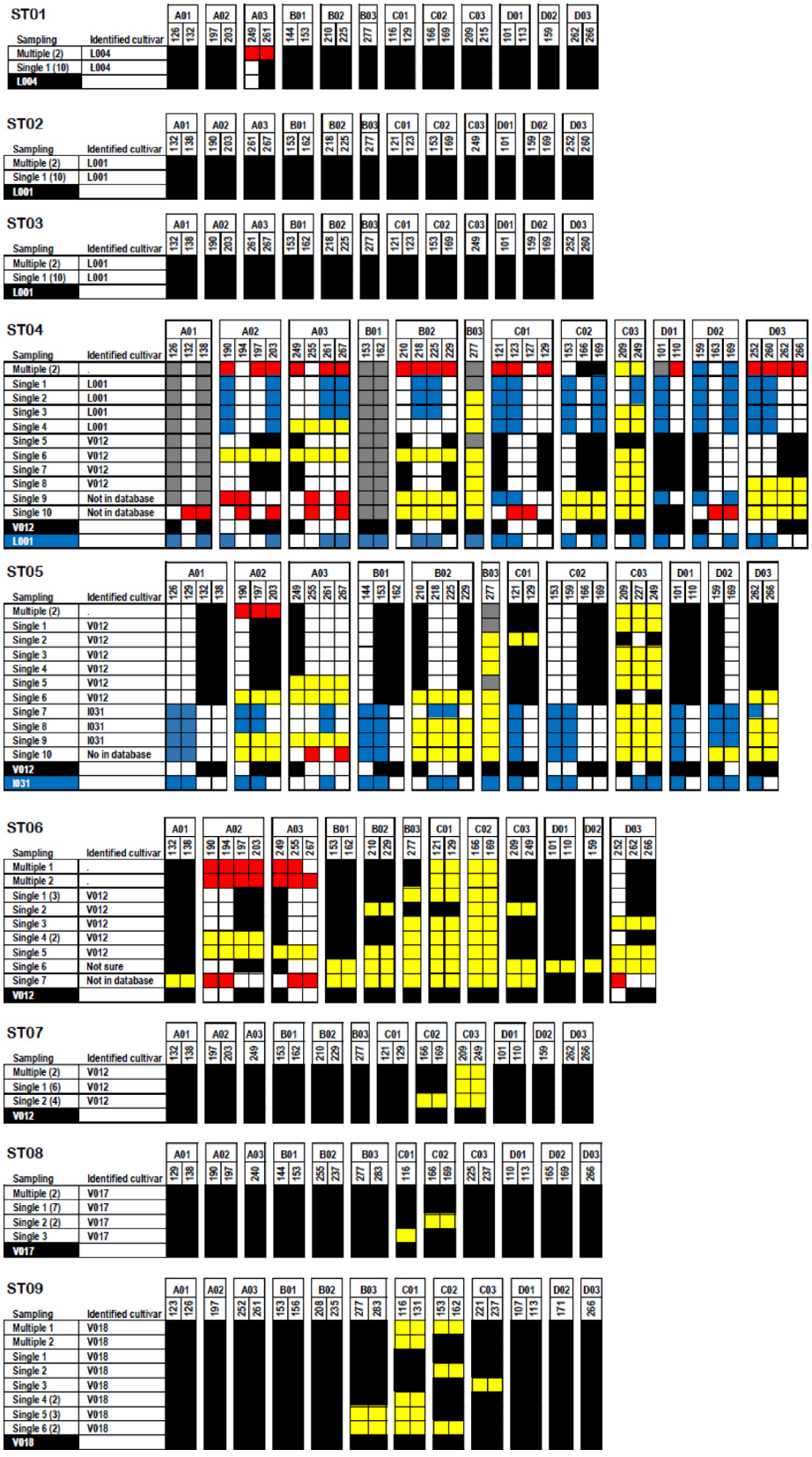

Fig. 4. Genotypes of 12 SSR loci in the multigranular and monogranular samplings of nine processed tea samples. "Multiple" and "Single" denote the multigranular samplings and monogranular samplings, respectively. Amplicon lengths of the alleles are shown right below each locus. The number of samplings with the same genotype combinations is provided in the parenthesis, otherwise is one. The alleles found in sampling are

indicated in gray, red, blue, and black colors, otherwise in white. All possible alleles are in yellow if the genotype of a locus was not recognizable. Black and blue colors indicate the genotype of a SSR locus in a sampling match to the identified cultivars provided in the bottom one or two rows. Gray color indicates that the genotype of two matching cultivars is identical. Red color indicates that the genotype is different from the identified cultivar shown in the bottom one or two rows.

probably due to primer competition within a primer set [24,25]. Moreover, DNA degradation in processed teas might result in low amplification and vague chromatographic peaks. Despite these caveats, our data supported that the multi-granular sampling method enables recognizing mixed-cultivar samples even after the specialty teas are manufactured by several steps known to break DNAs. In the comparisons of single and multi-granular samples, we also found that allele number and frequency might restrict the specificity of allele recognition and identification. Occurrence of random variation resulting from PCR artifacts may also result in heterozygote peak imbalance and high shutter [26]. Two strategies have been proposed to increase the liability of data analysis in forensic science. One is to increase the concentration and quality of DNA by improving DNA extraction method. The other is to carry out technical replicates. To rapidly extract DNA, we used a binding affinity-based DNA extraction kit. The concentration and quality of DNA isolated by this method were adequate for using the TTVID kit in most of the processed tea samples, except the Black Oolong tea (ST06) and Taiwan Black tea var. *assamica* (ST09). We also tried Exgene Plant SV kit (Cat. No. 117−101, GeneAll, Seoul, Korea), which was used in the previous study [23], but the DNA concentration was not increased (data not shown). Although the CTAB method is time-consuming, DNA concentration yielded from this method was indeed higher in our hand (data not shown). Thus, we recommended to extract two Taiwan specialty teas, Black Oolong tea and Taiwan Black tea, by using the CTAB method. In addition, including two to three technical replicates per sampling were suggested to eliminate the random errors in PCR and increase the liability of allele callings [8].

We have shown that the TTVID kit developed in this research adequately identified the cultivars used in processed teas. Several measures may be incorporated to efficiently identify tea cultivars from the inspected processed teas at the Taiwan border. First, multi-granular sampling can be used in the preliminary examination to rapid screen suspicious tea products. If signatures of multiple cultivars or problematic samples were recognized, single-granular sampling could be used in the follow-up examination to identify the specific cultivars. Such a strategy will facilitate the inspection of suspicious tea products by reducing processing time and examination costs. Second, single granules should be picked by a sample divider or cluster sampling to reduce sampling bias. Third, the concentration of template DNA and the relative concentration of primers in each primer set could be adjusted according to the type of specialty tea tested. The more time-consuming CTAB method might be used to extract DNA of the tea products manufactured by a heavily fermented (oxidative) process. Fourth, technical replicates in PCR can be implemented in the sample examination to reduce the false positive and negative identification of alleles. Fifth, the DNA concentration and the number of PCR cycle should be optimized to increase the liability of allele callings. More PCR artifacts are introduced by a higher number of PCR cycles, resulting in heterozygote peak imbalance and high shutter [8]. We performed PCR with 40 cycles in our experiments. Based on the studies in forensic science, allele drop-out or locus drop-out (false negative allele calling) are more likely to occur when only 28 cycles of PCR were done with the template DNA concentration lower than 100 pg. On the other hand, allele drop-in or allele imbalance are likely to occur in PCR with higher than 34 cycles. Further tests are required to optimize the number of PCR cycles and the concentration of DNAs extracted from different types of processed teas. Finally, our database of 12 SSR loci includes Taiwan cultivars only. Expanding our database by collecting more processed tea via the research institutes in other tea-producing regions is vital.

A few differential loci reflected a close genetic relationship between two compared cultivars. For instance, only one differential locus was found in V008−I009, echoing that TTES No. 08 (V008) is originated from Jaipuri (I009). Several cultivar pairs showing four or fewer differential loci are landrace cultivars established from single individuals by tea growers since 1810 in Taiwan. Overall, the genotype combinations of 12 SSR loci in our TTVID kit were effectively detected subtle genetic differences among the cultivars in our current database.

5. Conclusions and remarks

Accurately identifying cultivars in a short period is vital to prevent tea adulteration and minimize the negative impact on tea trading. More SSR markers can reduce the incidental mis-identification of tea cultivars, but the repeated PCR and amplicon

examination in gel electrophoresis for multiple SSR loci are time-consuming, especially in inspecting a large quantity of samples during tea importation. Our strategy of using multiplexing PCR helps facilitate the inspection process by reducing inspection time, reagent consumptions, and work load. Our analysis shows that the selected 12 SSR markers offer enough power to distinguish over 100 tea cultivars. Our success on genotyping 12 SSR and identifying cultivars from Tongding Oolong, Tieh-Kuan-Yin, and Oriental Beauty tea samples demonstrated that the SSR-based molecular diagnosis is also applicable to highly fermented or roasted tea. Therefore, the TTVID kit can provide a reliable molecular diagnosis for inspecting processed teas at the Taiwan border.

Authors' contributions

C.Y. Hu, H.T. Tsai, C.F. Chiu, and T.C. participated in the research design and conducted the experiments. C.Y. Hu, N.H.K. Le, and S.D. Yeh participated in data analysis. C.Y. Hu and S.D. Yeh wrote the manuscript.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank tea manufacturers and Farmers' Associations in Sanxia, Pinglin, Lugu, and Mucha for providing the specialty tea samples. Funding was provided by the Council of Agriculture, R.O.C. 104AS-6.2.1-TS-T3 to C.Y. Hu, 110AS-1.3.2-ST-a3 to S.D. Yeh, and the Ministry of Science and Technology (now known as the National Science and Technology Council) MOST110-2311-B-008-002 to S.D. Yeh.

References

[1] Lei L. Tea supply chain in East Asia. Institute of Developing Economies, Japan External Trade Organization(JETRO); 2016. Contract No.: 571.
[2] Su T-C, Yang M-J, Huang H-H, Kuo C-C, Chen L-Y. Using sensory wheels to characterize consumers' perception for authentication of Taiwan Specialty teas. Foods 2021;10:836.
[3] Hu C-Y, Tsai Y-Z, Lin S-F. Development of STS and CAPS markers for variety identification and genetic diversity analysis of tea germplasm in Taiwan. Bot Stud 2014;55:1–15.
[4] Institute ATR. The potential analysis report of market competition on Taiwan tea industy. Agricultural Technology Research Institute; 2019.
[5] Fang W, Meinhardt LW, Tan H, Zhou L, Mischke S, Wang X, et al. Identification of the varietal origin of processed loose-leaf tea based on analysis of a single leaf by SNP nanofluidic array. Crop J 2016;4:304–12.
[6] Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, et al. Current trends in microsatellite genotyping. Mol Ecol Resour 2011;11:591–611.
[7] Hu C-Y, Lin Y-C, Hsieh W-T, Tseng Y-H, Lin S-F, Tsai Y-Z. Using EST-SSR markers to identify tea (*Camellia sinensis*) cultivars in Taiwan. Taiwan Tea Res Bull 2011;30:9–12.
[8] Butler JM. Advanced topics in forensic DNA typing: methodology. Academic Press; 2011.
[9] Korir NK, Han J, Shangguan L, Wang C, Kayesh E, Zhang Y, et al. Plant variety and cultivar identification: advances and prospects. Crit Rev Biotechnol 2013;33:111–25.
[10] Wang RJ, Gao XF, Kong XR, Yang J. An efficient identification strategy of clonal tea cultivars using long-core motif SSR markers. SpringerPlus 2016;5 [In English].
[11] Kato F, Taniguchi F, Monobe M, Ema K, Hirono H, Maeda-Yamamoto M. Identification of Japanese tea (*Camellia sinensis*) cultivars using SSR marker. J Jpn Soc Food Sci 2008;55:49–55.
[12] Wang LB, Xun HS, Aktar S, Zhang R, Wu LY, Ni DJ, et al. Development of SNP markers for original analysis and germplasm identification in *Camellia sinensis*. Plants-Basel 2023;12.
[13] Ujihara T, Hayashi N, Tanaka J. Identification of material cultivar of green tea infusions by simple sequence repeat markers. Food Sci Technol Res 2012;18:209–17.
[14] Ujihara T, Matsumoto S, Hayashi N, Kohata K. Cultivar identification and analysis of the blended ratio of green tea production on the market using DNA markers. Food Sci Technol Res 2005;11:43–5.
[15] Ovesná J, Hrbek V, Svoboda P, Pianta V, Kučera L, Hajšlová J, et al. Microsatellite fingerprinting and metabolite profiling for the geographical authentication of commercial green teas. J Food Compos Anal 2021;101:103981.
[16] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 1987;19:11–5.
[17] Tan LQ, Wang LY, Wei K, Zhang CC, Wu LY, Qi GN, et al. Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant (*Camellia sinensis*). PLoS One 2013;8:e81611.
[18] Ma J-Q, Yao M-Z, Ma C-L, Wang X-C, Jin J-Q, Wang X-M, et al. Construction of a SSR-based genetic map and identification of QTLs for catechins content in tea plant (*Camellia sinensis*). PLoS One 2014;9:e93131.
[19] Hayden MJ, Nguyen TM, Waterman A, Chalmers KJ. Multiplex-ready PCR: a new method for multiplexed SSR and SNP genotyping. BMC Genom 2008;9:80.
[20] Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 1996;20:1004–10.
[21] Gouy A, Zieger M. STRAF-A convenient online tool for STR data evaluation in forensic genetics. Forensic Sci Int Gen 2017;30:148–51.
[22] Ristow PG, D'Amato ME. Forensic statistics analysis toolbox (FORSTAT): a streamlined workflow for forensic statistics. Forensic Sci Int Genet Suppl Ser 2017;6:e52–4.
[23] De Castro O, Comparone M, Di Maio A, Del Guacchio E, Menale B, Troisi J, et al. What is in your cup of tea? DNA verity test to characterize black and green commercial teas. PLoS One 2017;12:e0178262.
[24] Henegariu O, Heerema N, Dlouhy S, Vance G, Vogt P. Multiplex PCR: critical parameters and step-by-step protocol. Biotechniques 1997;23:504–11.
[25] Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE. Multiplex PCR: optimization and application in diagnostic virology. Clin Microbiol Rev 2000;13:559–70.
[26] Butler JM. Advanced topics in forensic DNA typing: interpretation. Academic Press; 2014.