

Development of a Molecular Marker to Identify a Candidate Line of Turmeric (*Curcuma longa* **L.)** with a High Curcumin Content

---Development of Molecular Marker of Turmeric

Hiroshi Hayakawa^{1,2}, Tetsuya Kobayashi², Yukio Minaniya^{2,3}, Katsura Ito⁴, Akira Miyazaki², Tatsuya Fukuda^{2*}, Yoshinori Yamamoto²

¹United Graduate School of Agricultural Sciences, Ehime University, Monobe, Nankoku, Japan; ²Faculty of Agriculture, Kochi University, Monobe, Nankoku, Japan; ³Graduate School of Integrated Arts and Sciences, Kochi University, Monobe, Nankoku, Japan; ⁴JST Innovation Satellite Kochi, Laboratory of Applied Entomology, Faculty of Agriculture, Kochi University, Monobe, Nankoku, Japan.

Email: tfukuda@kochi-u.ac.jp

Received October 13th, 2010; revised November 15th, 2010; accepted November 29th, 2010.

ABSTRACT

Dried and fresh rhizomes of the spice turmeric (Curcuma longa L.) are well known in traditional medicine, and curcumin is widely used in various geographic regions. Although there are differences in the amount of curcumin within this species, identification of the candidate line by rhizome is difficult because of the relative simplicity of its morphological characteristics. To accurately identify lines of C. longa with a high content of curcumin, we analysed several sequences of chloroplast DNA. First, to determine the appropriate outgroup taxa in which to conduct infraspecific analyses of C. longa, we reconstructed the molecular phylogenetic tree of C. longa and its allied species. The results showed that C. aromatica and C. zedoaria are closely related to C. longa. Next, to develop a molecular marker for identifying lines of C. longa with a high content of curcumin, a network analysis using chloroplast microsatellite regions was performed. Results showed that a unique haplotype within C. longa corresponds to the high curcumin content line. Therefore, the chloroplast microsatellite regions used for the analysis allowed us to determine the lines of this species with high curcumin content.

Keywords: Chloroplast DNA, Curcuma, Curcuma longa, Curcumin, Molecular Analysis

1. Introduction

Traditional medicine is known to be fertile ground for the source of modern medicines [1]. One medicine in that category is curcumin, a yellow coloring agent present in the spice turmeric (*Curcuma longa* L.) that belongs to the ginger family (Zingiberaceae). Besides its use in cooking to add color and as a preservative, turmeric is used in Indian traditional medicine to treat various common ailments including stomach upset, flatulence, dysentery, ulcers, jaundice, arthritis, sprains, wounds, acne, and skin and eye infections [2]. In this way, curcumin is widely used in various geographic regions in which the morphologies of the dried or fresh rhizomes of *Curcuma* plants are highly similar, and it is therefore difficult to

correctly identify individual Curcuma species.

Recent molecular phylogenetic studies have revealed new aspects of the relationships among *Curcuma* species. For example, to identify the rhizomes of *C. longa*, Sasaki *et al.* [3] developed a molecular identification method for *Curcuma* species using an amplification-refractory mutation system analysis of 18S ribosomal ribonucleic acid (rRNA) and the *trnK* gene. Although this method can identify species in a line of samples with high accuracy, the process is complicated. This marker is sensitive to experimental conditions of the polymerase chain reaction (PCR) because judgment is based on the presence or absence of PCR products. In addition, Minami *et al.* [4] reported that *Curcuma* species can be identified using an intergenic spacer between *trnS* and *trnfM* of chloroplast deoxyribonucleic acid (cpDNA). Although this method is useful for identifying *C. longa* from various *Curcuma* rhizomes, the study by Minami *et al.* [4] may require re-evaluation since the number of sample species used were few, and it did not employ appropriate outgroup taxa. The same can be said of the study by Sasaki *et al.* [3]. In our view, out-groups are important for assessing plesiomorphic or apomorphic states for each characteristic and for interpreting the phylogenetic relationships within *C. longa*. Alternatively, Aoi *et al.* [5] reported a difference in the amount of curcumin found among individuals of *C. longa*. Therefore, it is necessary to use infraspecific genetic markers to determine the amount of curcumin in individuals of this species in order to acquire it efficiently.

Large-scale sequencing of a predefined region of approximately 1500 base pairs (bp) of the cpDNA matK has one main goal in monocotyledonous plants including C. longa: to clarify the phylogenetic position and closely related species in order to determine the outgroup taxa of unknown individuals [6]. Although the sequence of the matK gene plays an important role in identifying the outgroup taxa of C. longa, it cannot be used to identify lines in this species containing high amounts of curcumin because of its low substitution rate [7]. Thus, determining more reliable genetic markers within C. longa requires more informative DNA regions. In recent years, cpDNA sequence data have been used frequently to reconstruct the phylogenetic relationships of a wide range of land plants [8]. Some regions were reported to have nucleotide substitution rates that are sufficiently high to facilitate the elucidation of relationships among closely related species [7,9]. Moreover, the utility of non-coding cpDNA regions within species-level phylogeny has been demonstrated in many taxonomic groups of land plants [10]. However, phylogenetic analysis using regions with high substitution rates, such as microsatellites, is often biased by multiple substitutions at a single site. In addition, because algorithms of phylogenetic analysis can cause problems when data do not represent a tree-like structure [11], we also analysed the cpDNA sequences with the statistical parsimony network approach. This approach reflects the genealogical relationships of the sequences used; that is, single-mutation steps separate adjacent haplotypes in the network, and older haplotypes are placed at internal branching points, whereas younger haplotypes occur towards the tip position. In the current study, we show that this analysis of the infraspecific relationships within C. longa can result in identification of clear relationships not only among young and rapidly speciating groups but also among old lineages reaching deep into the divergence of this group.

2. Materials and Methods

2.1. Cultivation and Measurements

The experiments were carried out in the fields (sandy soil) of the Faculty of Agriculture, Kochi University, Japan for four years (2006-2009). Table 1 lists the sample lines used for the cultivation experiments in this study: 2 C. aromatica Salisb., 12 C. longa L., and 1 C. zedoaria Rosc. Rows were constructed with a width of 70 cm and height of 20 cm. Rhizomes were transplanted 8 cm below the soil surface in one-row ridges, on hills separated by a distance of 30 cm, in late May for four years. For fertileizer dressing, a total of 1.5 kg/a of N, 0.6 kg/a of P₂O₅, and 1.4 kg/a of K₂O was applied over four years. In addition, 200 kg/a of compost fertilizer, 15 kg/a of magnesia lime, and 30 kg/a of chicken droppings were applied (Chicken droppings were not used in 2006.). The experimental plots were arranged in a randomized complete design with two replicates, which formed three rows. Due to a lack of seed rhizomes, some lines of C. longa were examined without replicate or two rows. Rows were covered by silver mulch. Irrigation was performed using a sprinkler, with observation of the amount of precipitation and the condition of the plants. Four or six hills containing average-sized shoots per line were sampled in early December each year for four years.

For curcumin analysis, the curcuminoid content of curcumin 1, curcumin 2, and curcumin 3 (curcumin, demethoxycurcumin, and bis-demethoxycurcumin, respectively) of the primary branch rhizome was measured by high performance liquid chromatography (HPLC, HITACHI, L-2420), according to the method described by Sato *et al.* [12]. We extracted curcumin by 80% of ethanol from 0.3 g of the dried and ground samples. HPLC was performed on a column of YMC-Pack Pro C18. Mixed acetonitrile with 0.1% of phosphoric acid (50:50) was used for the mobile phase. Current flowed at a rate of 1 mL per min. The amount of injection was 10 μ L.

2.2. DNA Analysis

Total DNA was isolated from fresh root with a Plant Genomic DNA Mini Kit (Viogene, Sunnyvale, CA, USA), according to the manufacturer's protocol. The PCR mix contained approximately 100 to 200 ng of total DNA, 1 μ M of each primer, 200 mM of each deoxynucleotide, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM of MgCl₂, and 1.25 units of *Taq* polymerase. Double-stranded DNA was amplified, after incubation at 94°C for 2 min, by 45

Currier	T and line	Nos. analysed hill ¹⁾					
Species	Locality	2006	2007	2008	2009	_	
Curcuma aromatica (Kochi)	Kochi Prefecture, Japan	6	6	4	4	_	
C. aromatica (Tanegashima)	Tanegashima Island, Kagoshima Prefecture, Japan	6	6	6	4		
C. longa (Kochi)	Kochi Prefecture, Japan	6	6	6	4		
C. longa (Tanegashima)	Tanegashima Island, Kagoshima Prefecture, Japan	6	6	6	4		
C. longa (Wakayama A)	Wakayama Prefecture, Japan	6	6	4	4		
C. longa (Wakayama B)	Wakayama Prefecture, Japan	6	4	4	4		
C. longa (Wakayama C)	Wakayama Prefecture, Japan	2)	6	4	4		
C. longa (Okinawa A)	Okinawa Prefecture, Japan			4	4		
C. longa (Okinawa B)	Okinawa Prefecture, Japan			4	4		
C. longa (Indonesia A)	Bogol, West Java, Indonesia	4	4	4	4		
C. longa (Indonesia B)	Bogol, West Java, Indonesia		4	4	4		
C. longa (Indonesia C)	Bogol, West Java, Indonesia				4		
C. longa (Vietnam A)	Vietnam				4		
C. longa (Vietnam B)	Vietnam				4		
C. zedoaria	Unknown	4	4	4	4		

1) Four or six hills containing average-sized shoots per line were harvested; 2) Not cultivated.

cycles of incubation at 94°C for 1.5 min, 48°C for 2 min, and 72°C for 3 min, with final extension at 72°C for 15 min. We amplified four regions from cpDNA, namely *matK*, *rpl16* intron 2, *petB* intron 1, and *petB* intron 2, with primers designed by Johnson and Soltis [13] and Nishizawa and Watano [9]. After amplification, reaction mixtures were subjected to electrophoresis in 1% low melting-temperature agarose gels for separation of specific amplified products. We sequenced the purified PCR products using a Big Dye Terminator Cycle Sequencing Kit (ABI PRISM, Perkin Elmer Applied Biosystems, Foster City, CA, USA) and an ABI PRISM 3100-Avant Genetic Analyzer according to the manufacturer's instructions. The primers used for sequencing were the same as those used for amplification.

To construct a phylogenetic tree based on *matK* sequences of *Curcuma* and its allied species, the amplified regions were aligned using ClustalW [14] and were improved manually using MEGA 4 [15]. Phylogenetic relationships were analysed using the neighbour-joining (NJ) method with PAUP* 4.08b [16]. The NJ analyses were performed using MEGA 4 with Kimura's two-parameter model. For the NJ analyses, bootstrapping with 1000 pseudo-replicates was chosen to examine the robustness

of the clades and their phylogenetic relationships. In addition, a network tree of microsatellite regions of cpDNA was constructed with Network 1.4.1.2 (Fluxus Technology Ltd., Suffolk, UK) using the median-joining method [17].

3. Results

3.1. Curcumin Content

We mainly describe results of the cultivation experiments for 2009, since data for the previous years of the study (2006, 2007, and 2008) showed similar trends.

Table 4 shows the curcumin content in the maturity period for the years 2006 through 2009. With respect to the curcumin content of primary branch rhizomes in the maturity period, total curcuminoid content, in order from highest to lowest, was as follows: South Asian *C. longa* (3198.4-2315.1 mg/100 g), with the exception of Indonesia B > domestic *C. longa* (305.0-392.2 mg/100 g), with the exception of Wakayama B > Indonesia B (309.5 mg/100 g) > *C. aromatica* (121.9-126.9 mg/100 g). Indonesia A and B and Vietnam A and B showed a high content of accumulated curcuminoid in the rhizome. However, *C. longa* (Wakayama B) and *C. zedoaria*

Species	matK			
Species	Accession No.	Reference		
Curcuma aeruginosa	AF478840	Kress et al. (2002)		
C. amarissima	AB047751	Cao et al. (Unpubl.)		
C. aromatica	AB047731	Cao et al. (2001)		
C. aromatica (Kochi)	AB551929	this study		
C. aromatica (Tanegashima)	AB551929	this study		
C. aromatica (Okinawa)	AB551929	this study		
C. attenuata	GQ248110	Hollingsworth et al. (Unpubl.)		
C. bicolor	AF478837	Kress et al. (2002)		
C. chuanezhu	AB047736	Cao et al. (2001)		
C. chuanhuangjiang	AB047732	Cao et al. (2001)		
C. chuanyujin	AB047733	Cao et al. (2001)		
C. comosa	AF478838	Kress et al. (2002)		
C elata	AB047747	Cao et al (Unpubl)		
C. axiana	AB047750	Cao <i>et al.</i> (Unpubl.)		
C. Exigui	AD047744	Coo at al. (2001)		
C. kwangsiensis A	AD047745	Cao et al. (Lemul)		
C. kwangstensis B	AB047745	this study		
<i>C. longa</i> (Tanagashima)	AB551930	this study		
C_{1} long (Wakayama A)	AB551930	this study		
C. longa (Wakayama A)	AB551930	this study		
C_{1} longa (Wakayama D)	AB551930	this study		
$C \log a$ (Wakayama C)	AB551930	this study		
C_{longa} (Okinawa R)	AB551930	this study		
C_{longa} (Indonesia A)	AB551930	this study		
$C \ longa (Indonesia R)$	AB551930	this study		
<i>C. longa</i> (Indonesia <i>D</i>)	AB551930	this study		
C_{longa} (Vietnam A)	AB551930	this study		
<i>C. longa</i> (Vietnam B)	AB551930	this study		
C longa	AB047738	Cao $et al.$ (2001)		
C phaeocaulis	AB047735	$C_{20} \text{ at } al (2001)$		
C. proceeding A	AD047741	$C_{ab} et al. (Uppubl.)$		
C. roscorana B	AB047741 AE478830	$K_{ress} at al. (2002)$		
C. siehusnensie A	AD047720	$C_{00} \neq d \left(\text{Uppubl} \right)$		
C. sichuanensis A	AB047739			
C. sichuanensis B	AB047740	Cao <i>et al.</i> (Unpubl.)		
C. thorelii	AF4/8841	Kress <i>et al.</i> (2002)		
C. wenyujin	AB047746	Cao <i>et al.</i> (2001)		
C. xanthorrhiza	AB047752	Cao et al. (Unpubl.)		
C. yunnanensis	AB047749	Cao et al. (Unpubl.)		
C. zedoaria	AB551932	this study		
C. zedoaria A	AB047734	Cao et al. (2001)		
C. zedoaria B	AB047743	Cao <i>et al.</i> (2001)		
Outgroup				
Boesenbergia rotunda	AF478827	Kress et al. (2002)		
Cautleya spicata	AF478834	Kress et al. (2002)		
Cornukaempferia aurantiflora	AF478835	Kress et al. (2002)		
Curcumorpha longiflora	AF478842	Kress et al. (2002)		
Kaempferia marginata	AB232054	Sitthithaworn and Komatsu (Unpubl.)		
Scaphochlamys hiloba	AF478889	Kress <i>et al.</i> (2002)		
Zingiher winga	AB0/7755	$C_{a0} et el (Unpubl.)$		
Lingiver miogu	AD047733	Cao ei ui. (Onpubl.)		

Table 2. Accession numbers of matK using phylogenetic analysis of Curcuma and outgroup taxa.

с ·	Т				
Species	rps16 intron 2	petB intron 1	petB intron 2	Haplotype	
Curcuma zedoaria	a ¹⁾	e	g	А	
C. aromatica (Kochi)	b	f	i	В	
C. aromatica (Tanegashima)	b	f	i	В	
C. longa (Kochi)	с	e	j	D	
C. longa (Tanegashima)	с	e	j	D	
C. longa (Wakayama A)	с	e	j	D	
C. longa (Wakayama B)	с	e	j	D	
C. longa (Wakayama C)	с	e	j	D	
C. longa (Okinawa A)	с	e	j	D	
C. longa (Okinawa B)	с	e	j	D	
C. longa (Indonesia A)	d	f	h	С	
C. longa (Indonesia B)	с	e	j	D	
C. longa (Indonesia C)	d	f	h	С	
C. longa (Vietnam A)	d	f	h	С	
C. longa (Vietnam B)	d	f	h	С	

Table 3. Accession numbers based on network analysis of Curcuma longa and allied species.

1) Type of chloroplast DNA. Accession numbers; a: AB557651, b: AB557648, c: AB557650, d: AB557649, e: AB557653. f: AB557652, g: AB557657, h: AB557655, i: AB557654, j: AB557656. Haplotype is corresponded to **Figure 3**.

showed little curcuminoid content. The inside color of the rhizome in *C. longa* (Wakayama B) and *C. zedoaria* was white and purple, respectively, whereas *C. longa*, which showed an accumulation of curcuminoid, had a yellow-colored rhizome (**Figure 1**).

In order from highest to lowest of each content of curcumin 1, 2, and 3 showed a similar trend with the order of total curcuminoid content. With respect to the ratio of its curcumin 1, 2, and 3 in domestic C. longa (with the exception of Wakayama B) was 69%-71%, 21%-23%, and 8%-9%, respectively. However, the ratio of curcuminoid content in South Asian C. longa (with the exception of Indonesia B) was 46%-52% (curcumin 1), 26%-28% (curcumin 2), and 22%-26% (curcumin 3), indicating a low percentage of curcuminoid in curcumin 1 and a high percentage of curcuminoid in curcumin 3, compared with domestic C. longa. The ratio of curcumin 1, 2, and 3 in Indonesia B differed from that in Indonesia A and C and more closely resembled domestic C. longa, despite the fact that Indonesia A, B, and C originated in the same country. In C. aromatica, the ratio of curcumin 1, 2, and 3 was 47%-53%, 47%-52%, and 0-1%, respecttively. C. aromatica also had a high percentage of curcumin 2.

The relative order of curcuminoid content in the primary branch rhizomes was similar throughout the four years of the study. Morishita *et al.* [18] reported that year-to-year variation in curcuminoid content is more vigorous than that of the varietal differences in turmeric. In the present study, the difference of curcuminoid content in *Curcuma* species had not only varietal difference but also year-to-year variation.

3.2. DNA Analysis

To construct the molecular phylogenetic tree of *Curcuma* and its allied species, we determined sequences of the *matK* gene of *Curcuma* cpDNA and seven outgroups of *Boesenbergia rotunda*, *Cautleya spicata*, *Cornukaempferia aurantiflora*, *Curcumorpha longiflora*, *Hedychium greenei*, *Scaphochlamys biloba*, and *Zingiber mioga*. The lengths of the *matK* gene of *Curcuma* species varied from 1831 bp (*C. longa* (Wakayama B)) to 1846 bp (*C. thorelii*).

The results of the phylogenetic analysis indicated that *Curcuma* species consists of a monophyletic group (**Figure 2**), with the genus *Curcuma* primarily divided into two monophyletic groups: clade 1 and clade 2. Clade 1, located at the most basal monophyletic position of *Cur*-

Development of a Molecular Marker to Identify a Candidate Line of Turmeric (*Curcuma longa* L.) with a High Curcumin Content

Year	Species	Curcumin 1 content ¹⁾ Curcumin 2 con		2 content ¹⁾	Curcumin 3	3 content ¹⁾	Curcuminoids content ¹⁾		
	Curcuma aromatica (Kochi)	59.5 ^{efg}	$(47)^{2}$	65.7°	$(52)^{2)}$	1.0 ^d	$(1)^{2)}$	126.2 ^c	(100)
	C. aromatica (Tanegashima)	64.7^{defg}	(53)	56.7°	(47)	0.5^{d}	(0)	121.9 ^c	(100)
	C. longa (Kochi)	274.6 ^c	(70)	87.5°	(22)	30.1 ^d	(8)	392.2°	(100)
	C. longa (Tanegashima)	254.6 ^{cde}	(70)	77.7°	(22)	29.1 ^d	(8)	361.3°	(100)
	C. longa (Wakayama A)	260.3 ^{cde}	(69)	83.9 ^c	(22)	30.3 ^d	(8)	374.5 ^c	(100)
	C. longa (Wakayama C)	268.4 ^{cd}	(69)	89.0 ^c	(23)	32.1 ^d	(8)	389.5°	(100)
2009	C. longa (Okinawa A)	246.6 ^{cde}	(71)	72.7 ^c	(21)	27.8 ^d	(8)	347.1°	(100)
	C. longa (Okinawa B)	213.8 ^{cde}	(70)	63.5°	(21)	27.7 ^d	(9)	305.0°	(100)
	C. longa (Indonesia A)	1588.0^{a}	(52)	813.6 ^a	(26)	657.5 ^b	(22)	3059.1ª	(100)
	C. longa (Indonesia B)	209.1 ^{cdef}	(68)	72.7 ^c	(23)	27.8 ^d	(9)	309.5°	(100)
	C. longa (Indonesia C)	1169.9 ^b	(50)	618.2 ^b	(27)	526.9 ^c	(23)	2315.1 ^b	(100)
	C. longa (Vietnam A)	1427.1ª	(48)	840.1 ^a	(28)	709.5 ^b	(24)	2976.7ª	(100)
	C. longa (Vietnam B)	1474.7^{a}	(46)	884.8^{a}	(28)	838.9 ^a	(26)	3198.4ª	(100)
	C. longa (Wakayama B)	0.5 ^g	(49)	0.7°	(51)	0.0^{d}	(0)	1. ^c	(100)
	C. zedoaria	0.8^{fg}	(70)	0.4 ^c	(30)	0.0^{d}	(0)	1.2 ^c	(100)
	C. aromatica (Kochi)	19.2 ^b	(41)	27.1 ^b	(58)	0.2 ^b	(1)	46.5 ^b	(100)
	C. aromatica (Tanegashima)	15.4 ^b	(42)	20.9 ^b	(58)	0.1^{b}	(1)	36.3 ^b	(100)
	C. longa (Kochi)	239.1 ^{ab}	(67)	86.4 ^{ab}	(24)	32.2 ^{ab}	(9)	357.6 ^{ab}	(100)
	C. longa (Tanegashima)	273.3 ^{ab}	(70)	85.6 ^{ab}	(22)	32.8 ^{ab}	(9)	391.7 ^{ab}	(100)
	C. longa (Wakayama A)	268.4 ^{ab}	(69)	84.2 ^{ab}	(22)	35.8 ^{ab}	(10)	388.3 ^{ab}	(100)
2008	C. longa (Wakayama C)	271.6 ^{ab}	(69)	88.0^{ab}	(22)	35.9 ^{ab}	(9)	395.5 ^{ab}	(100)
	C. longa (Okinawa A)	256.0 ^{ab}	(70)	78.9 ^{ab}	(22)	29.0 ^{ab}	(8)	364.0 ^{ab}	(100)
	C. longa (Okinawa B)	265.0 ^{ab}	(71)	74.8^{ab}	(22)	33.3 ^{ab}	(9)	373.1 ^{ab}	(100)
	C. longa (Indonesia A)	1294.7ª	(49)	652.5ª	(24)	730.6 ^a	(27)	2677.8ª	(100)
	C. longa (Indonesia B)	229.7 ^{ab}	(68)	77.0 ^{ab}	(23)	29.9 ^{ab}	(9)	336.6 ^{ab}	(100)
	C. longa (Wakayama B)	0.4^{b}	(71)	0.3 ^b	(29)	0.0^{b}	(0)	0.7 ^b	(100)
	C. zedoaria	0.5 ^b	(49)	0.6 ^b	(51)	0.0^{b}	(0)	1.0 ^b	(100)
-	C. aromatica (Kochi)	30.60 ^{abc}	(43)	39.35 ^{ab}	(56)	0.87^{ab}	(1)	70.81 ^b	(100)
	C. aromatica (Tanegashima)	37.53 ^{ab}	(48)	38.74 ^{abc}	(51)	1.54 ^{ab}	(2)	77.81 ^{bc}	(100)
	C. longa (Kochi)	260.38 ^{bcd}	(65)	96.32 ^{bcd}	(24)	41.01 ^{bc}	(10)	397.71 ^{ab}	(100)
	C. longa (Tanegashima)	259.06 ^{bcd}	(68)	86.29 ^{bcd}	(23)	36.88 ^{bc}	(10)	382.24 ^{ab}	(100)
2007	C. longa (Wakayama A)	270.04 ^{ab}	(67)	93.36 ^{ab}	(23)	39.65 ^{ab}	(10)	403.05 ^{ab}	(100)
2007	C. longa (Wakayama C)	266.83 ^{abcd}	(67)	97.22 ^{abc}	(24)	40.19 ^{abc}	(10)	404.24 ^{abc}	(100)
	C. longa (Indonesia A)	1276.73 ^a	(53)	687.15 ^a	(25)	813.05 ^a	(23)	2776.93 ^a	(100)
	C. longa (Indonesia B)	149.72 ^{abcd}	(66)	58.11 ^{abcd}	(25)	21.17 ^{abc}	(9)	229.00 ^{abc}	(100)
	C. longa (Wakayama B)	1.18 ^d	(51)	0.99 ^{cd}	(49)	0.10 ^c	(4)	2.27 ^c	(100)
	C. zedoaria	1.28 ^{cd}	(67)	0.38 ^d	(29)	0.08 ^c	(5)	1.74 ^c	(100)
2006	C. aromatica (Kochi)	16.05 ^{bc}	(43)	20.96 ^{ab}	(56)	0.60^{abc}	(2)	37.61 ^{bcd}	(100)
	C. aromatica (Tanegashima)	18.36 ^{abc}	(45)	22.02 ^{ab}	(54)	0.51 ^{bc}	(1)	40.89 ^{abcd}	(100)
	C. longa (Kochi)	226.72 ^{ab}	(69)	68.90 ^a	(21)	31.60 ^{ab}	(10)	327.21 ^{ab}	(100)
	C. longa (Tanegashima)	219.47 ^{ab}	(73)	56.67 ^a	(19)	25.23 ^{ab}	(8)	301.36 ^{abc}	(100)
	C. longa (Wakayama A)	127.72 ^{abc}	(71)	37.66 ^{ab}	(21)	13.87 ^{abc}	(8)	179.25 ^{abcd}	(100)
	C. longa (Indonesia A)	1358.64ª	(48)	679.14 ^a	(24)	811.43 ^a	(28)	2849.21ª	(100)
	C. longa (Wakayama B)	0.26 ^c	(34)	0.43 ^b	(56)	0.07 ^c	(9)	0.76^{d}	(100)
	C. zedoaria	0.43 ^c	(57)	0.33 ^b	(43)	0.00°	(0)	0.76 ^{cd}	(100)

Table 4. Curcumin content at maturity 2006-2009.

1) Curcumin content in primary branch rhizomes (mg/100g). 2) Percentage of curcumin 1, 2, and 3 content to curcuminoids content. Values followed by the same letter in a column are not significantly different at 5% level by one-way ANOVA.



(c)

(d)

Figure 1. Rhizomes of the genus *Curcuma*. (a) *C. aromatica* (typical type); (b) *C. longa* (WakayamaB); (c) *C. longa* (Kochi); (d) *C. zedoaria*. Scale bar = 5 cm.

cuma, includes C. aeruginosa, C. bicolor, and C. thorelii. Clade 2 includes the following 17 species: C. amarissima, C. aromatica, C. chuanezhu, C. chuanhuangjiang, C. chuanyujin, C. comosa, C. elata, C. exigua, C. kwangsiensis, C. longa, C. phaecaulis, C. roscoeana, C. sichuanensis, C. wenyujin, C. xanthorriza, C. yunnanensis, and C. zedoaria. With the exception of C. roscoeana B, clade 2 was further divided into two groups: subclade 1 and subclade 2. Subclade 1 included the following 12 samples: C. chuanezhu, C. comosa, C. elata, C. exigua, C. kwangsiensis B, C. phaeocaulis, C. sichuanensis B, C. xanthorriza, C. yunnanensis B, C. zedoaria, C. zedoaria A, and C. zedoaria B. The remaining Curcuma species were grouped under subclade 2.

Six species, C. aromatica, C. longa, C. kwangsiensis, C. roscoeana, C. sichuanensis, and C. zedoaria, were analysed using more than one sample. Of these, C. kwangsiensis (subclades 1 and 2), C. roscoeana (clade 2 and subclade 1), and C. sichuanensis (subclades 1 and 2) contained paraphyletic assemblages. Although three samples of C. zedoaria were included in the same clade (subclade 1), each sample belonged to a different lineage within the subclade; therefore, this species also was considered to be paraphyletic. Four samples of C. aromatica were included in the monophyletic group with C. chuanyujin and C. chuanhuangjiang in subclade 2. HowDevelopment of a Molecular Marker to Identify a Candidate Line of Turmeric (*Curcuma longa* L.) with a High Curcumin Content



Curcuma species



Figure 2. Phylogenetic tree of *Curcuma* and its allied species using the neighbour-joining (NJ) method. The numbers below the branches indicate the bootstrap value. Black squares indicate high curcumin content. Black circles indicate medium curcumin content. White circles indicate low curcumin content.

ever, because there was no synapomorphic character in this group, we were unable to determine whether C.

aromatica was monophyletic or paraphyletic. In addition, 13 samples of *C. longa* did not share a synapomorphic

Copyright © 2011 SciRes.

AJPS

character that would place them in the monophyletic group; therefore, this species is also considered to show polytomy. We were unable to determine whether *C. longa* was monophyletic or paraphyletic.

To establish the genetic polymorphisms within *C. longa*, we determined the chloroplast microsatellite regions of *rpl*16 intron 2, *petB* intron 1, and *petB* intron 2 of *C. aromatica*, *C. longa*, and *C. zedoaria*. The sequence lengths of *rpl*16 intron 2, *petB* intron 1, and *petB* intron 2 were 42-46 bp, 168-169 bp, and 47-62 bp, respectively. Network relationships of each individual by statistical parsimony showed that *C. longa* included two haplotypes: C and D (**Figure 3**). The genetic relationship between haplotype C and D was distant. It is notable that haplotype C was closely related to *C. aromatica* and *C. zedoaria* despite haplotype C being the most distinct of the *C. longa* haplotypes.

4. Discussion

4.1. Phylogenetic Relationship of *Curcuma* and Its Allied Species

To clarify the phylogenetic relationships of *Curcuma* and its allied species, we determined the nucleotide sequences of the cpDNA *matK* region for all species of the genus examined in this study. Only a few characters were considered in the phylogenetic analysis. We suggest that the degree of genetic differentiation among species of the genus *Curcuma* is small, despite the fact that these species are morphologically diverse, as evidenced by shoot/ root architecture and floral coloration. Therefore, it is likely that the genus lineage has undergone recent, rapid



Figure 3. Parsimony network of *C. longa*, *C. aromatica*, and *C. zedoaria* based on cpDNA *rpl16* intron 2, *petB* intron 1 and *petB* intron 2. Haplotype letters correspond to those in Table 3. White circles indicate the taxa in which haplotypes were found. Black circles between haplotypes represent a mutational step.

radiation. To confirm the scenario of rapid radiation in the genus *Curcuma*, we also determined the high-substitution regions of the regions of the *rps*16 intron (accession number AB557658) and *trnG* (UCC) intron (accession numbers AB557659 and AB557660), and the two intergenic regions of *trnV* (UAC)-*trnM* (CAU) (accession number AB55766) and *petD-rpoA* (accession number AB557662). No variable characters were detected even though the utility of these regions for intraspecific phylogeny was well determined [9]. This result also supports the hypothesis that the evolutionary history of the *Curcuma* genus underwent recent diversification.

Several reports of the rapid radiation of continental plant species have been published [19,20], although most of these examples are reported to have occurred on oceanic islands [21,22]. In most oceanic examples, rapid radiation is hypothesized to have been driven by low levels of competition in new habitats [23]. However, none of the samples used in the present study occur on oceanic islands. Alternatively, the concept of key innovations often is employed to explain the rapid radiation of a lineage [20,23,24]; therefore, the evolution of any key innovations in *Curcuma* may lead to rapid radiation. In the future, more studies should compare the morphologic and/or physiologic characters of *Curcuma* and its allied species to determine the nature of any key innovations.

4.2. Genetic Markers to Identify Lines with High Curcumin Content

Simple molecular marker assays that do not rely on specialized equipment or reagents underpin routine research activities in many laboratories worldwide. Recent research has shown that the non-coding portion of the plastid genome is more variable than previously anticipated [9]. Mutational hot spots have been reported from intronic or intergenic cpDNA regions of a number of organisms. Such hot spots are usually characterized by the presence of mononucleotide repeats [25] and a high incidence of duplications and other types of indels [26,27]. Short inversions [28] and minisatellites [29,30] have also been identified but seem to occur less frequently. In the present study, we were able to determine the sequencing repeat of three regions in Curcuma species, rpl16 intron 2, *petB* intron 1 and *petB* intron 2, providing evidence for additional and previously unrecognised types of polymorphic regions in the cpDNA of this group.

Identifying mutational hot spots in cpDNA is of utmost interest for studies below the species level. These so-called microsatellites have been used to detect intraspecific chloroplast polymorphisms in various species from a number of plant families [7]. In many cases, the

amplified fragments also contained size-variable sites in the target species. Thus, microsatellites may provide reasonable genetic markers to amplify the polymorphic chloroplast regions from any grasses, crops, fruits, and vegetables of interest. As the position of mutational hot spots in general is usually not conserved across species, these primer pairs will have to be tested in any new taxon on the basis of trial and error. However, the present report could extend the application of microsatellites to the Curcuma species. In fact, we observed three instances of haplotype sharing within C. longa, together with a considerable degree of intraspecific variation. Moreover, it is interesting that C. longa exhibits unique chloroplast haplotypes (e.g. haplotype C) that are involved in lines with high curcumin content (Figure 3). This result indicates that the microsatellite regions are useful for identifying C. longa candidate lines with high curcumin content. Further studies will determine whether more comprehensive sampling and additional evidence will support the results of the present study regarding the identification of C. longa lines with high curcumin content.

In this study, we showed that the cpDNA of C. longa can be divided into two groups, expecting to expose a cryptic species with the appropriate biologic characteristics. Cryptic species are closely related species that differ critically in genetic, physiologic, behavioral, and ecological traits [31], and the abundance of morphologically cryptic or unrecognized species, even in well-known taxa, suggests that there are more species than are currently recognized or estimated [32]. Recent studies have reported that cryptic species seem to be especially common in certain taxa [33-36], and similar evidence for cryptic species was found in our results. Therefore, in order to verify this hypothesis with respect to C. longa, further reciprocal crosses among the haplotypes discussed in the present study may lead to a conclusion of cryptic diversity. In the future, additional research will be needed to determine if the various cryptic species differ in their ecological requirements and tolerance to potential environmental stresses. There is also the need to re-examine the alpha taxonomy of the four species showing cryptic diversity. Without this basic taxonomic research, these cryptic species will remain in taxonomic crypsis [37].

Our limited data set does not allow us to decide the most appropriate genetic marker at the present time. To accurately map the distribution of *C. longa*, and to monitor phylogeographic patterns within this species in more detail, we are currently applying the chloroplast markers developed in the present study to a large number of specimens sampled throughout the distributional range of this species. In addition, genetic and phylogenetic studies in *Curcuma* species have lagged behind those of

other crop species. The network analysis presented here, along with the polymorphic information presented, provide valuable genetic resources for future studies of *C. longa* and related species.

5. Acknowledgements

We wish to thank R. Arakawa and J. Yokoyama, T. Yoshida, A. Matsuzawa, A. Hirata, Y. Muramatsu, M. Saito, R. Ueda, K. Ohga, N. Yokoyama, M. Muroi, and K. Matsuyama for providing much help. This study was partly supported by a grant-in-aid for scientific research from the Ministry of Education, Science, and Culture of Japan (to TF).

REFERENCES

 T. W. Corson and C. M. Crews, "Molecular Understanding and Modern Application of Traditional Medicines: Triumphs and Trials," *Cell*, Vol. 130, No. 5, 2007, pp. 769-774.

doi:10.1016/j.cell.2007.08.021

- [2] S. Singh, "From Exotic Spice to Modern Drug?" *Cell*, Vol. 130, No. 5, 2007, pp. 765-768. doi:10.1016/j.cell.2007.08.024
- [3] Y. Sasaki, H. Fushimi, H. Cao, S. Q. Cai and K. Komatsu, "Sequence Analysis of Chinese and Japanese Curcuma Drugs on the 18S rRNA Gene and *trn*K Gene and the Application of Amplification-Refractory Mutation System Analysis for Their Authentication," *Biological & Pharmaceutical Bulletin*, Vol. 25, No. 12, 2002, pp. 1593-1599. doi:10.1248/bpb.25.1593
- [4] M. Minami, K. Nishio, Y. Ajioka, H. Kyushima, K. Shigeki, K. Kinjo, K. Yamada, M. Nagai, K. Satoh and Y. Sakurai, "Identification of *Curcuma* Plants and Curcumin Content Level by DNA Polymorphisms in the *trnS-trnfM* intergenic Spacer in Chloroplast DNA," *Journal of Natural Medicines*, Vol. 63, No. 1, 2009, pp. 75-79. doi:10.1007/s11418-008-0283-7
- [5] K. Aoi, K. Kaburagi, T. Seki, T. Tobata, M. Sarak and M. Kuroyanagi, "Studies on the Cultivation of Turmeric (*Curcuma longa L.*) I: Varietal Differences in Rhizome Yield and Curcuminoid Content," *Bulletin of National Institute of Hygienic Sciences*, Vol. 104, 1986, 124-128.
- [6] M. N. Tamura, J. Yamashita, S. Fuse and M. Hamaguchi, "Molecular Phylogeny of Monocotyledons Inferred from Combined Analysis of Plastid *matK* and *rbcL* Gene Sequences," *Journal of Plant Research*, Vol. 117, No. 2, 2004, pp. 109-120. doi:10.1007/s10265-003-0133-3
- [7] D. E. Soltis and P. S. Soltis, "Choosing an Approach and an Appropriate Gene for Phylogenetic Analysis," In: D. E. Soltis, P. S. Soltis and J. J. Doyle, Eds. *Molecular Systematics of Plants II*. Kluwer Academic Publishers, Boston, 1998, pp. 1-42.
- [8] R. G. Olmstead and P. A. Reeves, "Evidence for the Polyphyly of the Scrophulariaceae Based on Chloroplast *rbcL* and *ndhF* Sequences," *Annals of the Missouri Bo-*

tanical Garden, Vol. 82, No. 2, 1995, pp. 176-193. doi:10.2307/2399876

- [9] T. Nishizawa and Y. Watano, "Primer Pairs Suitable for PCR-SSCP Analysis of Chloroplast DNA in Angiosperms," *Journal of Phytogeography and Taxonomy*, Vol. 48, No. 1, 2000, pp. 67-70.
- [10] S. A. Kelchner, "The Evolution of Non-Coding Chloroplast DNA and Its Application in Plant Systematics," Ann. Annals of the Missouri Botanical Garden, Vol. 87, No. 4, 2000, pp. 482-498. doi:10.2307/2666142
- [11] D. Posada and K. A. Crandall, "Intraspecific Gene Gene-Alogies: Trees Grafting into Networks," *Trends in Ecol*ogy & Evolution, Vol. 16, No. 1, 2001, pp. 37-45. <u>doi:10.1016/S0169-5347(00)02026-7</u>
- [12] M. Sato, K. Shimura and K. Hashizume, "Quality Valua-Tion of Turmeric Powders on Market," *Mie Prefectural Health and Environment Research Institution*, Vol. 6, 2004, pp. 52-54.
- [13] L. A. Johnson and D. E. Soltis, "matK DNA Sequences and Phylogenetic Reconstruction in Saxifragaceae s.str.," Systematic Botany, Vol. 19, No. 1, 1994, pp. 143-156. doi:10.2307/2419718
- [14] J. D. Thompson, D. G. Higgins and T. J. Gibson, "CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Positions-Specific Gap Penalties and Weight Matrix Choice," *Nucleic Acids Research*, Vol. 22, No. 22, 1994, pp. 4673-4680. doi:10.1093/nar/22.22.4673
- [15] K. Tamura, J. Dudley, M. Nei and S. Kumar, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software. Version 4.0," *Molecular Biology and Evolution*, Vol. 24, No. 8, 2007, pp. 1596-1599. doi:10.1093/molbev/msm092
- [16] D. L. Swofford, "paup*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4," Sinauer Associates, Sunderland, MA. 2001.
- [17] H. J. Bandelt, P. Forster and A. Röhl, "Median-Joining Networks for Inferring Intraspecific Phylogenies," *Molecular Biology and Evolution*, Vol. 16, No. 1, 1999, pp. 37-48.
- [18] T. Morishita, H. Yamaguchi, K. Degi, T. Shimizu and H. Nakagawa, "Varietal Difference and Year-to-Year Variation of the Contents of Curcuminoids in Turmeric," *Japanese Journal of Crop Sciences*, Vol. 75, 2006, pp. 130-131. doi:10.1626/jcs.75.335
- [19] A. B. Utelli, B. A. Roy and M. Baltisberger, "Molecular and morphological analysis of European Aconitum species (Ranunculaceae)," *Plant Systematics and Evolution*, Vol. 224, No. 3-4, 2000, pp. 195-212. doi:10.1007/BF00986343
- [20] S. T. Malcomber, "Phylogeny of *Gaertinera* Lam. (Lubiaceae) Based on Multiple DNA Markers: Evidence of Rapid Radiation in a Widespread, Morphologically Diverse Genus," *Evolution*, Vol. 56, No. 1, 2002, pp. 42-57.

- [21] B. G. Baldwin, "Adaptive Radiation of Hawaiian Silvers-Word Alliance: Congruence and Conflict of Phylogenetic Evidence from Molecular and Non-Molecular Investiga-Tions," In: T. J. Givnish and K. J. Sytsma, Eds., *Molecular Evolution and Adaptive Radiation*, Cambridge University Press, Cambridge, 1997, pp. 103-128.
- [22] J. Francisco-Ortega, D. J. Crawford, A. S. Guerra and R. K. Jansen, "Origin and Evolution of Argyranthmum (Asteraceae: Anthemideae) in Macaronesia," In: T. J. Givnish and K. J. Sytsma, Eds., Molecular Evolution and Adaptive Radiation, Cambridge University Press, Cambridge, 1997, pp. 407-431.
- [23] K. F. Liem, "Key Evolutionary Innovations, Differential Diversity, and Symecomorphosis," In: M. Nitecki, Ed., *Evolutionary Innovations*, University of Chicago Press, Chicago, 1990, pp. 147-170.
- [24] S. A. Hodges and M. L. Arnold, "Columbines: A Geographically Widespread Species Flock," *Proceedings of the National Academy of Sciences of the USA*, Vol. 91, No. 11, 1994, pp. 5129-5132. doi:10.1073/pnas.91.11.5129
- [25] J. Provan, N. Soranzo, N. J. Wilson, J. W. McNicol, M. Morgante and W. Powell, "The Use of Uniparentally Inherited Simple Sequence Repeat Markers in Plant Population Studies and Systematics," In: P. M. Hollingsworth, R. M. Bateman and R. J. Gornall, Eds., *Molecular Systematics and Plant Evolution*, Taylor & Francis, London, 1999, pp. 1-19.
- [26] B. R. Morton and M. T.Clegg, "A Chloroplast DNA Mutational Hotspot and Gene Conversion in a Noncoding Region near *rbcL* in the Grass Family (Poaceae)," *Current Genetics*, Vol. 24, No. 4, 1993, pp. 357-365. doi:10.1007/BF00336789
- [27] D. A. Johnson and J. Hattori, "Analysis of a Hotspot for Deletion Formation within the Intron of the Chloroplast *trnI* Gene," *Genome*, Vol. 39, No. 5, 1996, pp. 999-1005. doi:10.1139/g96-124
- [28] S. A. Kelchner and J. F. Wendel, "Hairpins Create Minute Inversions in Non-Coding Regions of Chloroplast DNA," *Current Genetics*, Vol. 30, No. 3, 1996, pp. 259-262. doi:10.1007/s002940050130
- [29] D. Cafasso, G. Pellegrino, A. Musacchio, A. Widmer and S. Cozzolino, "Characterization of a Minisatellite Repeat Locus in the Chloroplast Genome of *Orchis palustris* (Orchidaceae)," *Current Genetics*, Vol. 39, No. 5-6, 2001, pp. 394-398. doi:10.1007/s002940100226
- [30] R. A. King and C. Ferris, "A Variable Minisatellite Sequence in the Chloroplast Genome of Sorbus L. (Rosaceae: Maloideae)," *Genome*, Vol. 45, No. 3, 2002, pp. 570-576. doi:10.1139/g02-018
- [31] D. H. Darling and J. H. Werren, "Biosystematics of Nasonia (Hymenoptera: Pteromalidae): Two New Species Reared from Birds' Nests in North America," Annals of the Entomological Society of America, Vol. 83, No. 3, 1990, pp. 352-370.
- [32] E. M. Barratt, R. Deaville, T. M. Burl, M. W. Bruford, G.

Jones, P. A. Racey and R. K. Wayne, "DNA Answers the Call for Pipistrelle Bat Species," *Nature*, Vol. 387, 1997, pp. 138-139. doi:10.1038/387138b0

- [33] I. D. Hogg, M. I. Stevens, K. E. Schnabel and M. A. Chapman, "Deeply Divergent Lineages of the Widespread New Zealand Amphipod Paracalliope Fluviatilis Revealed Using Allozyme and Mitochondrial DNA Analyses," *Freshwater Biology*, Vol. 51, No. 2, 2006, pp. 236-248. doi:10.1111/j.1365-2427.2005.01491.x
- [34] T. Tully, C. A. D'Haese, M. Richard and R. Ferriere, "Two Major Evolutionary Lineages Revealed by Molecular Phylogeny in the Parthenogenetic Collembola Species Folsomia Candida," *Pedobiologia*, Vol. 50, No. 2, 2006, pp. 95-104. doi:10.1016/j.pedobi.2005.11.003
- [35] D. Bickford, D. J. Lohman and N. S. Sodhi, "Cryptic

Species as a Window on Diversity and Conservation," *Trends in Ecology & Evolution*, Vol. 22, No. 3, 2007, pp.148-155. doi:10.1016/j.tree.2006.11.004

- [36] T. L. Finston, M. S. Johnson, W. F. Humphreys, S. M. Eberhard and S. A. Halse, "Cryptic Speciation in Two Widespread Subterranean Amphipod Genera Reflects Historical Drainage Patterns in an Ancient Landscape," *Molecular Ecology*, Vol. 16, No. 2, 2007, pp. 355-365. doi:10.1111/j.1365-294X.2006.03123.x
- [37] B. C. Schlick-Steiner, B. Seifert and C. Stauffer, "Without Morphology, Cryptic Species Stay in Taxonomic Crypsis Following Discovery," *Trends in Ecology & Evolution*, Vol. 22, No. 8, 2007, pp. 391-392. doi:10.1016/j.tree.2007.05.004