Theor Appl Genet (2003) 107:757–767
DOI 10.1007/s00122-003-1311-3

Sang-Min Chung · Jack E. Staub

The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa

Received: 16 September 2002 / Accepted: 31 March 2003 / Published online: 25 June 2003

Abstract Although universal or consensus chloroplast primers are available, they are limited by their number and genomic distribution. Therefore, a set of consensus chloroplast primer pairs for simple sequence repeats (ccSSRs) analysis was constructed from tobacco (*Nicotiana tabacum* L.) chloroplast sequences. These were then tested for their general utility in the genetic analysis of a diverse array of plant taxa. In order to increase the number of ccSSRs beyond that previously reported, the target sequences for SSR motifs was set at A or T ($n \geq 7$) mononucleotide repeats. Each SSR sequence motif, along with ±200-bp flanking sequences from the first of each mononucleotide base repeat, was screened for homologies with chloroplast DNA sequences of other plant species in GenBank databases using BLAST search procedures. Twenty three putative marker loci that possessed conserved flanking sequence spans were selected for consensus primer pair construction using commercially available computer algorithms. All primer pairs produced amplicons after PCR employing genomic DNA from members of the Cucurbitaceae (six species) and Solanaceae (four species). Sixteen, 22 and 19 of the initial 23 primer pairs were successively amplified by PCR using template DNA from species of the Apiaceae (two species), Brassicaceae (one species) and Fabaceae (two species), respectively. Twenty of 23 primer pairs were also functional in three monocot species of the Liliaceae [onion (*Allium cepa* L.) and garlic (*Allium sativum* L.)], and the Poaceae [oat (*Avena sativa* L.)]. Sequence analysis of selected ccSSR fragments suggests that ccSSR length and sequence variation could be useful as a tool for investigating the genetic relationships within a genus or closely related taxa (i.e., tribal level). In order to provide for a marker system having significant coverage of the cucumber chloroplast genome, ccSSR primers were strategically "recombined" and named recombined consensus chloroplast primers (RCCP) for PCR analysis. Successful amplification after extended-length PCR of 16 RCCP primer pairs from cucumber (*Cucumis sativus* L.) DNA suggested that the amplicons detected are representative of the cucumber chloroplast genome. These RCCP pairs, therefore, could be useful as an initial molecular tool for investigation of traits related to a chloroplast gene(s) in cucumber, and other closely related species.

Keywords Consensus primer · Chloroplast · Simple sequence repeats · Variable region · Genetic relationships

Communicated by C. Möllers

S.-M. Chung · J. E. Staub (✉)
USDA/ARS, Vegetable Crops Unit, Department of Horticulture, 1575 Linden Drive, University of Wisconsin, Madison, WI 53706, USA
e-mail: jestaub@facstaff.wisc.edu
Tel.: 608-262-0028
Fax: 608-262-4743

Introduction

Molecular markers have been used for analysis of genetic diversity and germplasm organization as well as for genetic similarity estimation (Staub et al. 1996). Unlike nuclear-derived genetic markers, markers constructed from chloroplast genomes are highly conserved (i.e., low nucleotide substitution rates when compared to the nuclear genome). Such markers developed from chloroplast DNA (cpDNA) are desirable for taxonomic and phylogenetic relationship studies at various systematic levels (Wolfe et al. 1987).

DNA sequence information of organelle genomes is necessary for the development of organellar genome markers. Comprehensive organelle DNA sequence information, however, is not broadly available in most species. To overcome this limitation, "consensus" (synomous universal) markers have been developed which possess DNA sequences that are homologous to conserved regions of the plant chloroplast and mitochondrial genomes (Taberlet et al. 1991; Demesure et al. 1995; Dumolin et al. 1997). For instance, consensus polymerase chain reaction (PCR) primer pairs have been constructed from conserved coding sequences of

organellar DNA in liverwort (*Marchantia polymorpha* L.) (Ohyama et al. 1986), rice (*Oryza sativa* L.) (Hiratsuka et al. 1989; Taberlet et al. 1991) and tobacco (*Nicotiana tabacum* L.) (Taberlet et al. 1991; Demesure et al. 1995; Dumolin et al. 1997; Grivet et al. 2001). The PCR products obtained from these primers have subsequently been used in either direct sequencing experiments or the derivation of restriction fragment length polymorphisms (RFLPs) for use in taxonomic, phylogenetic and phylogeographic studies at the intraspecific level (Demesure et al. 1996; King and Ferris 1998; Mohanty et al. 2000).

Chloroplast markers, however, could not be fully utilized for the assessment of closely related chloroplast genomes until highly polymorphic genome regions were identified (McCauley 1995). Chloroplast simple sequence repeats (SSRs or microsatellites; cpSSRs) were developed for such genetic analyses in the 1990s (Powell et al. 1995; Vendramin et al. 1996; Vendramin and Ziegenhagen 1997). This technology is based on highly polymorphic regions and has been used for the genetic analyses of the chloroplast genome of *Glycine* (Powell et al. 1996), *Hordeum* (Provan et al. 1999), *Oryza* (Provan et al. 1996; Ishii and McCouch 2000), *Pinus* (Powell et al. 1995), and *Solanum* (Bryan et al. 1999).

The development of such cpSSR markers has been achieved through the use of chloroplast DNA sequence information that is resident in GenBank (http://www.ncbi.nlm.nih.gov). Weising and Gardner (1999) developed ten consensus chloroplast microsatellite primers (ccmp), and concluded that such primers had utility for discrimination of closely related genotypes in a broad array of plant species. They used A or T mononucleotide repeats ($n \geq 10$) as SSR motifs identified in the tobacco chloroplast genome (Shinozaki et al. 1986) as a development strategy, and demonstrated that most of the derived primer pairs were functional and descriptive as genetic markers in the Actinidiaceae, Brassicaceae, and Solanaceae.

A substantial number of consensus chloroplast markers distributed uniformly across a genome is, however, a requirement for comprehensive genetic analysis. The consensus chloroplast primers developed from tobacco by Weising and Gardner (1999) and Grivet et al. (2001) are limited in their coverage of the tobacco chloroplast genome (i.e., from bp position 1 to 86,694 and from 154,185 to 155,939). Consequently, the development of additional consensus chloroplast SSR (ccSSR) primers which define uncharacterized positions (i.e., from 86,694 to 154,185) of this genome would be desirable; especially in plant species whose sequence information is not now available. Therefore, experiments were designed to develop an expanded set of ccSSR primers from *N. tabacum* that would have utility for the genetic analysis of chloroplast genomes (objective 1). Studies were then conceived that evaluated the general efficacy of the primers developed in economically important plant species for genetic analysis (objective 2). Lastly, these ccSSR primers were strategically recombined [designated as recombined consensus chloroplast primers (RCCPs)] for comprehensive cucumber chloroplast genetic analysis (objective 3). These RCCPs were designed for complete coverage of the chloroplast genome, and thus provide for increased utility in genetic analyses of the cucumber chloroplast genome.

Materials and methods

Development of consensus chloroplast SSRs (ccSSRs)

To expand the number of candidate consensus chloroplast SSR loci beyond that currently available (Weising and Gardner 1999) (objective 1), mononucleotide repeats of A, C, G or T ($n \geq 7$) in the *N. tabacum* chloroplast sequences (accession number: CHNTXX, 155,939 bp) (Shinozaki et al. 1986) were identified using Genetool software (BioTools Inc, Edmonton, Alberta, Canada). This information was downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov) for use as a reference in the construction of primer pairs. Each SSR motif sequence, including ±200-bp flanking sequences from the first of each mononucleotide base repeat, was screened for homologies with chloroplast DNA sequences of other plant species in GenBank databases using BLAST search procedures (Altschul et al. 1990). Twenty three sequence regions (putative marker loci) that showed conserved flanking sequences were selected for consensus primer-pair (ccSSR; Table 1) construction. Primer sequences were chosen from tobacco chloroplast DNA sequences employing Genetool software, and were modified according to the degenerations of mismatch nucleotides between tobacco sequences and the aligned chloroplast DNA sequences from BLAST search results.

Plant materials and DNA isolation

To evaluate the efficacy of the 23 ccSSRs (objective 2), economically important plant families that had not previously been subjected to complete chloroplast genome-sequencing projects were examined (Ohyama et al. 1986; Shinozaki et al. 1986; Hiratsuka et al. 1989; Sato et al. 1999; Hupfer et al. 2000; Schmitz-Linneweber et al. 2001). This included representative accessions from the Apiaceae [carrot (*Daucus carota* L.; USDA inbred 'B493') from P.W. Simon, University of Wisconsin-Madison, Wis., and parsley (*Petroselinum crispum* Mill.; 'Giant Italian') from Garden City Seeds, Hamilton, Mont.]; Brassicaceae [cabbage (*Brassica oleracea* L.; 'Copenhagen market') from Meredith Company, Chattanooga, Tenn.]; Cucurbitaceae [bottle gourd (*Lagenaria siceraria* (Molina) Stand.) from The Cucurbit Network, Miami, Fla.; cucumber (*Cucumis sativus* L.; line GY-14) from Clemson University, Clemson, S.C.; melon (*Cucumis melo* L.; 'Top Mark') from Hollar Seed Company, Rocky Ford, Colo.; luffa (*Luffa quinquefida* (Hook. & Arn.) Seemann.), squash (*Cucurbita pepo* L.) represented by five cultivars of *Cucurbita pepo* L. ('Connecticut Field', 'Orange Ball', 'Ozarkana', 'Texana', and 'Vegetable Marrow Bush'), and watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nak) from The Cucurbit Network]; Fabaceae [bean (*Phaseolus vulgaris* L.; 'Improved tendergreen') and pea (*Pisum sativum* L.; 'Mr. Big') from W. Atlee Burpee and Company, Warminster, Pa.]; Liliaceae [garlic (*Allium sativum* L.; 'PI515971') from P.W. Simon and onion (*Allium cepa* L.; 'Bringham yellow globe 15–23') from I. Goldman, University of Wisconsin-Madison, Wis.]; Poaceae [oat (*Avena sativa* L.; 'Ogle') from H. Kaeppler, University of Wisconsin-Madison, Wis.]; and Solanaceae [egg plant (*Solanum melongena* L.; 'Long purple') from Plantation Company, Norton, Mass., pepper (*Capsicum annuum* L.; 'California wonder') from Excel Company Chattanooga, Tenn., and tomato (*Lycopersicon esculentum* and *Lycopersicon hirsutum*) from R. Chetelat, University of California-Davis, Calif.]. Tobacco (*N.*

Table 1 Consensus chloroplast simple sequence repeat (ccSSR) primers developed from the *N. tabacum* L. chloroplast genome

Name	Forward (F), Reverse (R) primers (5′ to 3′)[a]	T_M (°C)[b]	SSR motif[c]	Position[d]	Gene names[e]	Expected size (bp)[f]
ccSSR-1	F TCAAATGATACATAGTGCGATACA R AATAAAGGATTTCTAACCATCTT	51 46	(T)10	3,702-3,725 3,875-3,853	TrnK	174
ccSSR-2	F AATCCTGGACGTGAAGAATAA R AATCCCTCTCTTTCCGTTGA	49 50	(A)11	8,488-8,508 8,647-8,628	TrnQ-TrnS	160
ccSSR-3	F CCAAAAGCTGACATAGATGTTA R TTTCATTCGGCTCCTTTATG	49 48	(T)11	10,041-10,062 10,147-10,128	TrnG	107
ccSSR-4	F AGGTTCAAATCCTATTGGACGCA R TTTTGAAAGAAGCTATTCARGAAC	53 49–51	(T)8	10,479-10,501 10,683-10,660	TrnR-AtpA	205
ccSSR-5	F TCTGATAAAAAACGAGCAGTTCT R GAGAAGGTTCCATCGGAACAA	50 52	(T)10	16,771-16,793 17,040-17,020	Rps2-RpoC2	270
ccSSR-6	F CGACCAATCCTTCCTAATTCAC R AGAAAAGMAAGGATATGGGCTC	53 51–53	(T)8	26,724-26,745 27,022-27,001	RpoB	299
ccSSR-7	F CGGGAAGGGCTCGKGCAG R GTTCGAATCCCTCTCTCTCCTTTT	57–59 56	(T)11	36,812-36,829 37,160-37,137	PsbC-TrnS	349
ccSSR-8	F TTGATCTTTACGGTGCTTCCTCTA R TCATTACGTGCGACTATCTCC	54 52	(T)5C(T)17	44,933-44,956 45,181-45,161	Ycf3	249
ccSSR-9	F GAGGATACACGACAGARGGARTTG R CCTATTACAGAGATGGTGYGATTT	56–59 52–54	(A)13	45,985-46,008 46,157-46,134	Ycf3	173
ccSSR-10	F TCTAGGATTTACATATACAACAT R CATCATTATTGTATACTCTTTCA	46 46	(A)13	57,267-57,289 57,415-57,393	AtpB-RbcL	149
ccSSR-11	F TTGGCTACTCTAACCTTCCC R ACCATAGAAACGAWGGAACCCACT	52 56	(T)6C(T)14	71,536-71,555 71,700-71,677	Rpl20-ClpP	165
ccSSR-12	F CCAAAAACTTGGAGATCCAACTAC R TTCCATAGATTCGATCGTGGTTTA	54 52	(A)8	76,437-76,460 76,685-76,662	PsbB-PsbT	249
ccSSR-13	F AGTCTGAAACCRAGKGGATTTATT R TAGTCATTAGTAAAGCCGARGTSA	51–54 52–54	(T)9	85,864-85,887 86,127-86,104	Rps3-Rpl22	264
ccSSR-14	F GGGTATAATGGTAGATGCCC R GCCGTAGTAAATAGGAGAGAAA	52 51	(T)14	86,557-86,576 86,756-86,735	Rps19-Rpl2	200
ccSSR-15	F GCTTATGACCTCCCCCTCTATGC R TGCATTACAGACGTATGATCATTA	59 51	(T)9	88,062-88,084 88,325-88,302	Rpl2-Rpl23	264
ccSSR-16	F TACGAGATCACCCCTTTCATTC R CCTGGCCCAACCCTAGACA	53 55	(T)7C(T)2	104,961-104,982 105,083-105,065	TrnL	123
ccSSR-17	F CACACCAATCCATCCCGAACT R GGTGCGTTCCGRGGTGTGA	54 55–58	(A)13	109,627-109,647 109,862-109,844	Rrn5-TrnR	236
ccSSR-18	F TCGTTGGATTTCTTCDGGACATTT R CCCAATATCATCATACTTACRTGC	52–54 52–54	(A)8	116,517-116,540 116,780-116,757	Ycf5	264
ccSSR-19	F CTATGCAGCTCTTTTATGYGGATC R TCCARGTAATAAATGCCCAAGTT	54–56 50–52	(T)8	116,781-116,804 117,115-117,093	Ycf5	335
ccSSR-20	F CCGCARATATTGGAAAAACWACAA R GCTAARCAAATWGCTTCTGCTCC	51–52 53–55	(A)8	118,974-118,997 119,284-119,262	NdhD-PsaC	311
ccSSR-21	F CCACCCCGTCTCSACTGGATCT R AAAAATAGCTCGACGCCAGGAT	60 53	(T)13	132,649-132,670 132,928-132,907	TrnR-Rrn5	280
ccSSR-22	F CCGACCTAGGATAATAAGCYCATG R GGAAGGTGCGGCTGGATC	56–57 55	(T)8	138,213-138,236 138,402-138,385	TrnL-16SrRNA	190
ccSSR-23	F AYGGRGGTGGTGAAGGGAG R TCAATTCCCGTCGTTCGCC	53–58 53	(A)14	155,747-155,765 24-6	Rp12-TrnH	217

[a] Y (= C or T), R (= A or G), M (= A or C), K (= T or G), W (= A or T) and S (= C or G)
[b] T_M = melting temperature as calculated by Oligo Calculator Version 3.01 (http://www.basic.nwu.edu/biotools/oligocalc.html)
[c] Bold mononuleotide repeats indicates SSR motifs the same as that reported by Weising and Gardner (1999)
[d] Position according to the tobacco chloroplast genome (accession number = CHNTXX) (http://www.ncbi.nlm.nih.gov)
[e] Names of chloroplast genes where consensus chloroplast SSR (ccSSR) primer sequences are located according to the tobacco chloroplast genome
[f] Expected size according to the tobacco chloroplast genome

tabacum L.; 'Havana 38'), obtained from P.W. Simon, was used as a positive control to compare amplicons after PCR with their expected sizes (Table 2).

Seven to ten seeds of each accession were planted in a greenhouse at the University of Wisconsin, Madison. Seedling leaf tissue of each accession was collected and held at –80 °C until extraction. DNA was extracted according to Staub et al. (1996), and then quantified using a mini-fluorometer (model TD-360; Turner Designs, Sunnyvale, Calif.).

Table 2 Base pair-length sizes produced after PCR using 23 consensus chloroplast SSRs (ccSSRs) in a diverse array of plant species

Reference number	Species	Size of amplification products (bp)																						
		ccSSR number																						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	*Nicotiana tabacum* L.	174	160	107	205	270	299	349	249	173	149	165	249	264	200	264	123	236	264	335	311	280	190	217
2	*Solanum lycopersicum* L. 'lycopersicum Hirsutum'	176	160	111	205	270	299	349	239	161	150	155	251	264	207	264	127	232	264	335	310	285	190	317
3	*Solanum lycopersicum* L. 'lycopersicum esculentum'	176	160	109	205	268	299	352	239	163	150	153	250	264	207	264	127	234	264	335	310	285	190	316
4	*Solanum melongena* L.	174	165	111	201	275	299	349	239	173	150	159	255	264	205	264	127	249	264	341	310	300	190	302
5	*Capsicum annuum* L.	171	159	109	214	264	299	346	239	174	148	165	249	264	214	264	127	233	264	335	311	275	190	316
6	*Avena sativa* L.	N/A[a]	N/A	58	214	305	299	N/A	237	177	N/A	124	230	231	N/A	264	209	236	264	352	313	252	187	208
7	*Allium cepa* L.	N/A	N/A	80	202	N/A	299	249	N/A	171	159	114	231	239	N/A	264	345	198	264	365	313	240	185	196
8	*Allium sativum* L.	N/A	N/A	80	202	N/A	299	249	N/A	171	159	115	237	239	N/A	264	345	204	264	365	313	253	190	196
9	*Daucus carota* L.	N/A	N/A	100	206	247	299	346	222	154	139	160	256	285	202	264	356	234	264	367	309	278	185	302
10	*Petroselinum crispum* Mill.	N/A	N/A	102	192	N/A	299	340	245	173	N/A	151	256	N/A	193	266	345	235	264	367	N/A	280	185	N/A
11	*Pisum sativum* L.	174	206	87	247	258	299	317	249	154	164	152	243	N/A	N/A	271	360	225	264	377	322	280	362	N/A
12	*Phaseolus vulgaris* L.	N/A	175	82	249	290	299	308	265	138	193	169	208	N/A	197	264	353	236	266	380	318	269	180	N/A
13	*Brassica oleracea* L.	163	154	99	N/A	232	299	288	234	153	156	165	253	265	194	264	213	228	267	377	348	270	112	326
14	*Cucumis sativus* L.	176	144	94	628	252	299	339	233	177	142	159	236	280	197	264	220	234	270	358	343	280	185	258
15	*Cucumis melo* L.	176	144	97	612	253	299	333	231	181	142	159	236	280	202	264	219	231	270	358	342	277	185	290
16	*Citrullus lanatus* (Thunb.) Matsum. & Nakai	180	144	94	615	252	299	338	231	179	147	158	234	289	202	264	220	243	270	358	343	290	185	262
17	*Lagenaria siceraria* (Molina) Stand. 'Birdhouse'	175	144	94	645	251	299	338	232	179	147	162	237	289	202	264	220	231	270	358	344	279	185	262
18	*L. quinquefida* (Hook. & Arn.) Seemann	169	141	94	645	250	299	349	234	176	146	162	228	289	202	264	360	233	270	358	339	280	185	260

Table 2 (continued)

Reference number	Species	Size of amplification products (bp)																						
		ccSSR number																						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
19	*Cucurbita pepo* ssp. pepo L. 'Connecticut Field'	170	142	100	637	259	299	339	233	182	155	157	206	299	208	264	219	233	270	358	341	279	185	253
20	*Cucurbita pepo* ssp. pepo L. 'Orange Ball'	170	142	100	637	259	299	339	233	185	155	157	206	299	208	264	219	233	270	358	341	279	185	253
21	*Cucurbita pepo* ssp. pepo L. 'Vegetable Marrow Bush'	170	142	100	637	259	299	339	233	182	155	157	206	299	208	264	219	233	270	358	341	279	185	253
22	*Cucurbita pepo* ssp. ovifera L. 'texana'	170	142	100	637	259	299	339	233	185	154	157	206	299	208	264	219	233	270	358	341	279	185	253
23	*Cucurbita pepo* ssp. ovifera L. 'ozarkana'	170	142	100	637	259	299	339	233	185	154	157	206	299	208	264	219	233	270	358	341	279	185	253

[a] N/A = no amplification

Table 3 Recombined consensus chloroplast primers (RCCPs) derived from *N. tabacum* L. consensus chloroplast SSR primers (ccSSRs)

Name	Primers (5′–3′)[a]	ccSSR name	Position[b]	Expected size (bp)[c]	Detected size (bp)[d]
RCCP-1	F AYGGRGGTGGTGAAGGGAG	23-F	155,747-155,765	10,243	11,789
	R TTTCATTCGGCTCCTTTATG	3-R	10,147-10,128		
RCCP-2	F CCAAAAGCTGACATAGATGTTA	3-F	10,041-10,062	643	1,075
	R TTTTGAAAGAAGCTATTCARGAAC	4-R	10,683-10,660		
RCCP-3	F AGGTTCAAATCCTATTGGACGCA	4-F	10,479-10,501	6562	7,895
	R GAGAAGGTTCCATCGGAACAA	5-R	17,040-17,020		
RCCP-4	F TCTGATAAAAAACGAGCAGTTCT	5-F	16,771-16,793	10,252	12,109
	R AGAAAAGMAAGGATATGGGCTC	6-R	27,022-27,001		
RCCP-5	F CGACCAATCCTTCCTAATTCAC	6-F	26,724-26,745	10,437	12,437
	R GTTCGAATCCCTCTCTCTCCTTTT	7-R	37,160-37,137		
RCCP-6	F CGGGAAGGGCTCGKGCAG	7-F	36,812-36,829	8,370	10,042
	R TCATTACGTGCGACTATCTCC	8-R	45,181-45,161		
RCCP-7	F TTGATCTTTACGGTGCTTCCTCTA	8-F	44,933-44,956	12,483	13,120
	R CATCATTATTGTATACTCTTTCA	10-R	57,415-57,393		
RCCP-8	F TCTAGGATTTACATATACAACAT	10-F	57,267-57,289	14,434	14,600
	R ACCATAGAAACGAWGGAACCCACT	11-R	71,700-71,677		
RCCP-9	F TTGGCTACTCTAACCTTCCC	11-F	71,536-71,555	5,150	3,287
	R TTCCATAGATTCGATCGTGGTTTA	12-R	76,685-76,662		
RCCP-10	F CCAAAAACTTGGAGATCCAACTAC	12-F	76,437-76,460	9,691	11,478
	R TAGTCATTAGTAAAGCCGARGTSA	13-R	86,127-86,104		
RCCP-11	F AGTCTGAAACCRAGKGGATTTATT	13-F	85,864-85,887	2,462	2,413
	R TGCATTACAGACGTATGATCATTA	15-R	88,325-88,302		
RCCP-12	F GCTTATGACCTCCCCCTCTATGC	15-F	88,062-88,084	17,022	16,688
	R CCTGGCCCAACCCTAGACA	16-R	105,083-105,065		
RCCP-13	F TACGAGATCACCCCTTTCATTC	16-F	104,961-104,982	14,324	14,996
	R GCTAARCAAATWGCTTCTGCTCC	20-R	119,284-119,262		
RCCP-14	F CCGCARATATTGGAAAAACWACAA	20-F	118,974-118,997	13,955	14,600
	R AAAAATAGCTCGACGCCAGGAT	21-R	132,928-132,907		
RCCP-15	F CCACCCCGTCTCSACTGGATCT	21-F	132,649-132,670	5,754	6,548
	R GGAAGGTGCGGCTGGATC	22-R	138,402-138,385		
RCCP-16	F CCGACCTAGGATAATAAGCYCATG	22-F	138,213-138,236	17,751	16,248
	R TCAATTCCCGTCGTTCGCC	23-R	24-6		

[a] Y (= C or T), R (= A or G), M (= A or C), K (= T or G), W (= A or T), and S (= C or G)
[b] Position according to tobacco chloroplast genome (accession number = CHNTXX) (http://www.ncbi.nlm.nih.gov)
[c] Expected size is according to that defined in the tobacco chloroplast genome
[d] Size of amplified RCCP fragments of cucumber DNA detected using Labimage software (http://www.labimage.de) and given in Fig. 2

Polymerase chain reaction amplification employing standard and fluorescent procedures

Standard PCR amplifications were carried out using the 23 consensus ccSSR primer pairs developed by BLAST search and primer design procedures (Table 1). All PCR reagents were purchased from Promega Corporation (Madison, Wis.). Each 15-µl reaction volume contained 4.0 mM of $MgCl_2$, 0.3 mM of dNTPs, 15 ng of DNA, 0.45 µM of primers, polymerase buffer and 0.2 units of *Taq* DNA polymerase. Samples were over-laid with mineral oil, and all amplifications were conducted using the following cycling profile: 94 °C for 5 min; 35 cycles of 94 °C for 60 s, 50 °C for 60 s, 72 °C for 60 s; 72 °C for 6 min, and then followed by an indefinite soak at 4 °C.

To confirm PCR amplification, 5 µl of loading dye (0.1% bromophenol blue, 0.1% xylene cyanol FF, 10% Ficol.) was added to each reaction tube. The samples were then electrophoresed in 1.6% agarose gels (20 × 40 cm) containing 0.5 µg/ml of ethidium bromide in 0.5× TBE buffer (0.045 M Tris-borate and 1.0 mM EDTA pH 8.0), for 3.5 h at approximately 170 V. The stained gels were imaged with a digital camera, and recorded using "Gel Expert" (Nucleotech Corporation, San Mateo, Calif.).

To obtain the size of PCR amplified fragments, fluorescent PCR was performed in 8-µl volumes. Each reaction contained 15 ng of template DNA, 4 mM of $MgCl_2$, 0.25 M each of the forward and reverse primers, 0.1 mM of each dNTP, 1 × commercial polymerase buffer, 0.2 units of *Taq* polymerase and either 0.04 µM dUTP [R6-G] (green), 0.04 µM dUTP [R110] (blue) or 0.2 µM dUTP [TAMRA] (yellow). Fluorescent dUTPs were purchased from Perkin Elmer (Applied Biosystems, Incorporated, Fremont, Calif.). PCR products containing each of three different fluorochromes were then pooled for analysis. For fragment analysis of the fluorescent PCR, 1 µl of each PCR reaction was mixed with 1.2 µl of sample loading buffer (80% formamide, 10 mg/ml of blue dextran, 5 mM of EDTA pH 8.0), and 0.3 µl of a molecular-weight standard (ILS 600 Promega standard). Samples were heated to 95 °C for 3 min, and then chilled on ice. Approximately 0.7 µl of each sample was subsequently spotted onto a membrane comb (Gel Company, San Francisco, Calif.) and loaded on a gel containing 5% LongRanger (FMC Bioproducts, Rockland, Maine) polyacrylamide/6 M urea according to the manufacturer's instructions. Each gel was electrophoresed on a PE Applied Biosystems 377XL DNA sequencing apparatus (Fremont, Calif.) at 3,000 V with 2,400 scans per h in 36-cm well-to-read plates. Data were collected using the DNA Sequencer Data Collection v. 2.0 computer program, and subsequently analyzed using the GeneScan v. 3.1 algorithms (PE Applied Biosystems, Incorporated, Fremont, Calif.).

Sequencing of selected ccSSR fragments

To inverstigate whether the observed ccSSR length varations were due to SSR variations or other insertion/deletion events, four ccSSR fragments (i.e., 12, -17, -19 and -20) were chosen from primers developed herein that targeted cp SSR tobacoo locations from $n \geq 10$ (ccSSR-17) and $10 > n \geq 7$ (ccSSR-12, -19 and -20) (Table 1). Amplied fragments from luffa, lagenaria, squash ('Connecticut Field') and watermelon template DNAs were sequenced at least two times using a Perkin Elmer ABI 377 (Perkin Elmer, Boston, Mass.) sequencer. Both strands were sequenced for each fragment, and then tobacco sequences from GenBank were added to ccSSR fragment sequences to perform the alignment analyses.

For direct DNA sequencing reactions, standard PCR products were initially sized by agarose-gel electrophoresis. The excess dNTPs and unincorporated primers were then removed from the remaining PCR products using the ExoSAP-IT kit (USB Company, Cleveland, Ohio) according to the manufacture's protocol. Subsequently, 2 µl of each cleansed PCR product was added to a 8-µl sequencing reaction mixture containing 2 µl of 2.5 × reaction buffer (5 × in 400 mM Tris pH 9, 10 mM of $MgCl_2$), 2 µl of BigDye enzyme mix (Applied Biosystems, Foster City, Calif.), 0.5 µl of forward or reverse primer (20 pmol/µl), and 3.5 µl of water. Sequencing reactions were run on a Perkin-Elmer 9700 thermal cycler using the following cycling protocol: 50 cycles at 95 °C for 20 s, 50 °C for 30 s, 60 °C for 4 min, and then 72 °C for 7 min. The dye terminators were then removed through MicroBioSpin columns (BioRad, Hercules, Calif.). GeneTool software (BioTools Incorporated, Edmonton, Alberta, Canada, http://biotools.com) was used to align accession sequences for the detection of sequence polymorphisms among the accessions employed herein.

Amplifications of significant portions of the chloroplast genome by recombined consensus chloroplast primers (RCCP)

To determine ccSSR primer utility for amplification of a significant portion (coverage) of the chloroplast genome after "extended-length" PCR, forward and reverse primers of ccSSRs were recombined according to their base-pair positions in the tobacco chloroplast genome (objective 3). For instance, a forward primer from ccSSR-22 and a reverse primer from ccSSR-23 were combined to form a primer mixture designated as RCCP-16 for extended-length PCR experiments employing the cucumber DNA template (Table 3). This extended-length PCR procedure proceeded in a 15-µl volume as follows: 15 ng of line GY-14 cucumber DNA was reacted with 0.5 units of *Ex Taq* polymerase (PanVera Co. Madison, Wis.), 1.5 µl of 10× *Ex Taq* Buffer (contained 20 mM of $MgCl_2$), 1.2 µl of 25 mM dNTP, and 1.2 µl (2.5 µM) of each RCCP pair (Table 3). Each PCR was accomplished using the thermocycling profile of 94 °C for 5 min; 30 cycles of 94 °C for 60 s, annealing temperature 50 °C for 60 s, 68 °C for 17 min, and then 72 °C for 6 min. The samples were subsequently electrophoresed in 1.6% agarose gels (20 × 40 cm) containing 0.5 µg/ml of ethidium bromide in 0.5× TBE buffer (0.045 M of Tris-borate and 1.0 mM of EDTA pH 8.0) for 4.5 h at approximately 170 V. The size of all amplicons after extended-length PCR employing 16 RCCP pairs from cucumber DNA was determined using Labimage software (http://www.labimage.de), and then compared with expected band sizes from the tobacco chloroplast genome (Shinozaki et al. 1986).

Results

Development of consensus chloroplast SSRs (ccSSRs)

To expand the number of previously reported ccSSR primer pairs (Weising and Gardner 1999), an A, C, G, or T ($n \geq 7$), a mononucleotide repeat primer-construction strategy was used employing the *N. tabacum* chloroplast genome as a reference. No primer pairs were developed from C and G mononucleotide repeats because their flanking sequences did not show sufficient sequence conservation for consensus primer construction (data not presented). Nevertheless, 15 ccSSR primer pairs and eight ccSSR primer pairs were developed from $T \geq 7$ and $A \geq 7$ mononucleotide repeats, respectively (Table 1). Among those 23 ccSSRs, 13 and ten ccSSR primer pairs were constructed from $n \geq 10$ and $10 > n \geq 7$ mononucleotide repeats, respectively (Table 1).

To test the utility of consensus primers for genetic analysis, 23 accessions of eight plant families were analyzed using 23 ccSSR primer pairs (Table 2). Thirteen of the ccSSR primer pairs developed by the $n \geq 10$ strategy detected bp-length polymorphisms in 23 accessions. However, one (ccSSR-6) of ten ccSSR primer pairs developed by the $10 > n \geq 7$ strategy failed to detect bp-length polymorphisms in all 23 accessions (Table 2). All ccSSR primer pairs produced products after PCR using template DNA from selected species in the Cucurbitaceae (6) and the Solanaceae (4) (Table 2). Sixteen, 22 and 19 of the initial 23 primer pairs were successively amplified by PCR using template DNA from species of the Apiaceae (2), Brassicaceae (1), and Fabaceae (2), respectively. Twenty of 23 primer pairs were functional in detecting banding morphotypes using DNA from monocotyledonous species in the Liliaceae (2) and the Poaceae (1).

Two ccSSR pairs (ccSSR-9 and -10) allowed for the detection of bp-length polymorphisms among five cultivars within *C. pepo* (Table 2). Likewise, 15 of 23 ccSSR primer pairs allowed visualization of bp-length polymorphisms in three Cucurbitaceae genera of the Benincaseae tribe (*Citrullus, Lagenaria* and *Luffa*) (Table 2).

ccSSR fragment-sequence analysis of Cucurbitaceae species

Fragments obtained from ccSSR-12, -17, -19 and -20 employing Cucurbitaceae species were sequenced to investigate the hypothesis that the length variations detected are due to either SSR variations or other insertion/deletion events (Fig. 1). The targeted tobacco SSR motif (ccSSR-17; A_{13}) was detected at the same position that includes SSR length variations among the Cucurbitaceae species examined (Fig. 1). Length variation of the ccSSR-20 (A_8) motif was detected but not in the targeted position [i.e., that of the tobacco SSR motif (A_8)]. Nevertheless, both sequence analyses of ccSSR-12 (A_8) and -19 (T_8) did not detect the predicted SSR varation found in amplified regions among the Curcurbiteace species examined. Both ccSSR-12 and -17 alignments indicate that other insertion/deletion events exist in regions amplified by these primer pairs. Homoplasy of length varation due to different insertion events in different positions was observed only between luffa (TTCAAAA) and squash (TCAATAT) sequences of ccSSR-17 (Fig. 1).

Fig. 1 Alignment of ccSSR fragment sequences from Cucurbitaceae species amplified by primer pairs ccSSR-12, -17, -19 and -20, and the tobacco sequence from GenBank (http://www.ncbi.nlm.nih.gov). The length and sequence polymorphisms within the Cucurbitaceae are shown in *bold and italic*, and the SSR motifs originally targeted from the tobacco chloroplast genome (Table 1) are *bold and underlined*. Uncertain mononucleotides (i.e., resulting from multiple sequencing peaks or missing data) are designated as *N*

ccSSR-12

```
watermelon CTCTGTTC-CTTT-TTTC--CTTTATTTT-TTGATTTGAC----ATAGA--TAGGGGTAGCGGATAAATCTTG 73
lagenaria  CTCTGTTC-CTTT-TTTC--CTTTATTTT-TTGATTTGAC----ATAGA--TAGGG-TAGCGGATAAATCTTG
luffa      CT---------TT-TTTC--CTTTATTTT-TTGATTTGAC----ATAGA--TAGGG-TACCGGATAAATCTTG
squash     CTCTGTTC-CTTT-TTTC--CTTTATTTT-TTGATTTGAC----ATAGA--T--------------------
tobacco    CTTTGGTATCTTTCTTTCGCCCTTATTTTCTTTCTTTTACTTTTATTGACATAGGG-TACCAGAGAAATCTTT

watermelon ATTCGGATTGCCGACTTTTCGTTTCTTTGACTCTTGTCTTGGTCTTTTTTTT----AT--CCGTAAAAAGA-- 146
lagenaria  ATTCGGATTGCTGACTTTTCATTTCTTTGACTCTTGTCTTGGTCTTTTTTTT----AT--CCGTAAAAAGA--
luffa      ATTCGGATTGCTGACTTTTCATTTCTTTGACTCTTGTCTTGGTCTTTTTTTT----AT--CCGGAAAAAGA--
squash     ---------GCTGACTTTTCATTTCTTTGACTCTTATCTTGGTCTTTTTTTT----AT--CCGGAAAAAGA--
tobacco    ATTTGAAT--CA-ACTT--CGTTT-TT--ACTCTTGT-TCGTTCTTTATCCGGAAGATGACAAAAAAAAGAAA
```

ccSSR-17

```
watermelon TAATTCTTATTACTTTT-CAATATTCAATATCAATATGGAAAAAAAAAA---TGAAAAATGAAAAGGTCGTCT 73
lagenaria  TAATTCTTATTACTTTT-CAATAT------------GGAAAAAAAAAAA-TGAAAAATGAAAAGGTCGTCT
luffa      TAATTCTTATTACTTTT-CAATAT------------GGAAAAAAAAAAAATTCAAAATGAAAAGGTCGTCT
squash     TAATTCTTATTACTTTT-CAATATTCAATAT------GGAAAAAAAAAAAA-------TGAAAAGGTCGTCT
tobacco    TCATTCTTATTACTTTTCAATATG-AAAA-------CGAAAAAAAAAAAATGAAAAATCAAAAGGTCGTTT
```

ccSSR-19

```
watermelon TACATTTC-AAAAAGCTATA-AGAATTTTTGGTAAAAACAATAATTTCTTAAATGCGTTATTTTCCTTTGATG 73
lagenaria  TACATTTC-AAAAAGCTATA-AGAATTTTTGGTAAAAACAATAATTTCTTAAATGCGTTATTTTCCTTTGATG
luffa      TACATTTC-AAAAAGCTATA-AGAATTTTTGGTAAAAACAATAATTTATTAAATGCGTTATTTTCCTTTGATG
squash     TACATTTC-AAAACTCTATA-AGANTTTTTGGTAAAAACAATCATTTATTAAATGCGTTATTTTCCTTTGATG
tobacco    TACATTTCGAAAAAAC-AGACAGCTTTTTTA-TAAGAGCAATGGTTTTTTAAACGAGTCATTTTTCTT-G--G

watermelon AGATCCAATACATGAACGAAGGGAACAATTTTCTAAGAAACACTTCCTTTTTTTCTTCGAAGAATTATTATAA 146
lagenaria  AGATCCAATACATGAATGAAGGGAACAATTTTTTAAGAAACACTTCCTTTTTTTCTTCGAAGAATTATTATAA
luffa      AGATCCAATACATGAACGAAGGGAACAATGGTTTAGGAAACACTTCTTTTTTTTCTTCGAAGAATTATTATAA
squash     AGATCCAATACATGAACGANGAGAACAATGTTTAAGAAACACTTATTTTTTTTCTTATAATTAATTATTATAA
tobacco    -G----------TGAA----------AATGTTTTACAAAATACTTCTTTTTTTTCTGCTAAAAATTATTACAG
```

ccSSR-20

```
watermelon ATGAAATAAAAAAAA----TATATATATGTATATTTATTTTCGAGCACGAGTTTTTGTCGGTAAAAAAGAAATC 73
lagenaria  ATGAAATAAAAAAAAAAATATATATATGTATATTTATTTTCGAGCACGAGTTTTTGTCGGTAAAAAAGAAATC
luffa      ATGAAATAAAAAAAA----ATCTATATGTATATTTATTTTCGAGCACGAGTTTTTGTCGGTAAAAAATAAATC
squash     ATGAAATAAAAAAANA--TATNTATATGTATATTTATTTTCGAGCACGAGTTTTTGTCGGTAAAAAAGAAATC
tobacco    AT---AT-----------------TGTG-------ATTTTCGAGCAC-AGGTTT-GTCGGTAAAAAA--AATT
```

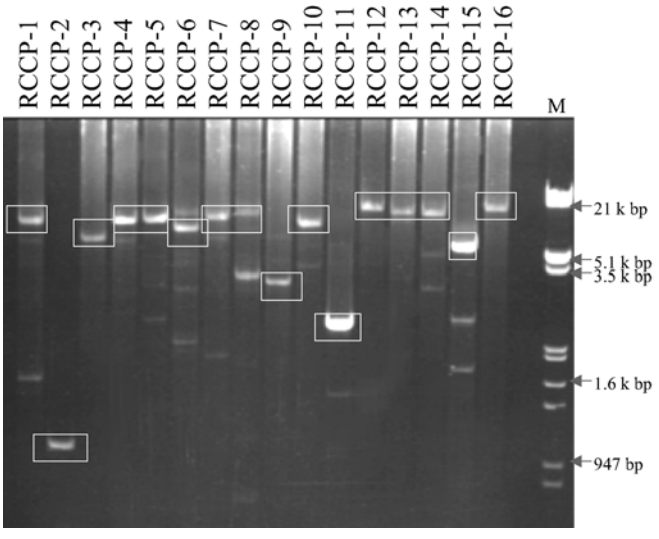

Fig. 2 Amplifications of cucumber DNA (GY-14) by 16 recombined consensus chloroplast primer (RCCP) pairs using an extended-length PCR technique. Lane M designates the Lambda DNA *Eco*RI + *Hin*dIII size ladder. Bands showing similar expected size based on the tobacco chloroplast genome (*N. tabacum* L.) are indicated by *boxes*

Amplifications of extended portions of the chloroplast genome by recombined consensus chloroplast primers (RCCPs)

All 16 RCCP pairs produced products after extended-length PCR using cucumber as a template DNA (Fig. 2). Although some RCCP pairs produced multiple amplicons, the brightest amplicons (except RCCP-8) from each RCCP pair possessed similar and expected size bands based on predictions from tobacco chloroplast-genome analysis (Fig. 2; Table 3).

Discussion

A $n \geq 7$ mononucleotide repeat strategy was used herein to construct ccSSR primer pairs from the tobacco chloroplast genome. Twenty three ccSSR primer pairs were developed and tested for their utility in genetic analysis. Their potential utility was determined by the assessment of polymorphism level in a broad array of plant species including five cultivars of *C. pepo* and three genera (*Citrullus, Lagenaria* and *Luffa*) of the tribe Benincaseae in the Cucurbitaceae family. These ccSSR primers were then recombined and analyzed for comprehensive cucumber chloroplast genetic analysis.

Weising and Gardner (1999) developed ten ccmp pairs using a $n \geq 10$ threshold strategy for the detection of mononucleotide repeats present in the tobacco chloroplast genome. In that study, the two criteria were applied to the selection of the most promising consensus chloroplast primers. Initially, only those loci were considered for which putative primer target regions flanking the mononucleotide repeat were sufficiently conserved in plant species. Secondly, consideration was given to sequences internal to the primer binding sites which harbored a poly (A) or poly (T) tract not only in tobacco, but also in other species used in database searches for their sequence homology. These criteria allowed for an increased chance of detecting size variation in other species. In the present study, a bp threshold of $n \geq 7$ was established for mononucleotide repeats to allow for the identification of additional consensus primer pairs for SSR analysis.

Many of the primer sequences reported herein are different from those reported by Weising and Gardner (1999). Nevertheless, the genomic locations of eight of 23 of the chloroplast SSR motifs characterized are the same as those of Weising and Gardner [i.e., ccSSR-1 (ccmp1), -2 (ccmp2), -3 (ccmp3), -5 (ccmp5), -8 (ccmp6), -10 (ccmp7), -11 (ccmp8), and -14 (ccmp10) (primer name in parentheses); Table 1]. We did not develop a new primer sequence for the ccmp4 SSR motif in this study because its primer sequences satisfied our criteria for consensus chloroplast development (i.e., primer sequences are highly conserved among plant species in the GenBank database). Therefore, given its conserved nature, this primer pair will likely be valuable for genetic-relationship analysis. The ccmp9 SSR primer motif, however, was not included in our study because its primer sequences were not located in the conserved region identified by our BLAST analyses (data not presented).

The primer pair of ccmp8 developed by Weising and Gardner (1999) was not a useful discriminator among some of the plant-species comparisons made herein. Thus, more conserved primer sequences were designed from their ccmp constructs (i.e., ccmp1, 2, 3, 5, 6, 7, 8, and 10). For example, the discriminatory performance of the ccmp8 SSR motif was altered (i.e., our ccSSR-11) to improve its ability to assess genetic differences in the species array examined herein (Table 2). The remaining 15 primer pairs that were constructed are unique and identify loci that have not been defined in Weising and Gardner (1999), the only previous report of consensus chloroplast SSR primer construction.

The hypothesis that $n \geq 7$ mononucleotide repeats in the *N. tabacum* chloroplast genome could also be used as indicators of regions of high chloroplast DNA variation was supported by the data presented herein. For instance, except for ccSSR-6, the ccSSR primer pairs (ccSSR-4, -6, -12, -13, -15, -16, -18, -19, -20, and -22) that were developed by the $10 > n \geq 7$ threshold strategy detected bp-length variations among the species examined. Likewise, the remaining 13 ccSSRs (ccSSR-1, -2, -3, -5, -7, -8, -9, -10, -11, -14, -17, -21, and -23) developed by the $n \geq 10$ threshold were also functional in the detection of polymorphisms among the species examined (Tables 1 and 2).

Weising and Gardner (1999) suggested that, for monocotyledonous species, it might be desirable to design an independent set of consensus primers based on nucleotide variation found in the completely sequenced rice or maize chloroplast genomes. Only three of 23 ccSSR primer pairs (ccSSR-1, -2 and -14) developed from the *N. tabacum* chloroplast genome used herein, failed to produce amplicons in all three monocotyledon species (Table 2). This result suggests that if, during BLAST searches, the conserved flanking sequence information of chloroplast SSR motifs from dicotyledonous species can be adequately aligned with the sequence information of monocot species presently existing in GenBank database, then consensus primers applicable to both dicotyledonous and monocotyledonous species can be constructed.

Sequence analysis of ccSSR fragments

Orti et al. (1997) suggested that there is a risk associated with employing SSRs originating from nuclear DNA for phylogenetic analysis. This risk stems from the inconsistency between bp-length variations among putative alleles and the known phylogenetic relationships among such alleles (Orti et al. 1997). For instance, if homoplasy exists in SSR length variations among a given set of taxa, their true taxonomic relationships might be indefinable by SSR analysis. Moreover, when a relatively small marker array (i.e., <25 markers) is applied to genetic-relationship analysis, banding morphotype misclassification by a few markers (i.e., one to three) may lead to errors in species classification and identity relationship estimation (Staub et al. 1996, 1997).

Sequence analysis of ccSSR fragments indicated that polymorphisms based on length variations might be due to insertion or deletion events of short DNA fragments instead of variations in the originally targeted mononucleotide repeat length (Fig. 1). However, the usefulness of such ccSSR primers is not limited to genetic-relationship studies since insertion or deletion events have a value for detecting and deciphering the evolution of the chloroplast genome. In the present study, homoplasy of length variation in sequence alignments was not detected at the tribe level (i.e., the Benincaseae tribe to include watermelon, luffa, and lagenaria), suggesting their potential utility for genetic analyses at the intra-tribe level. However, length homoplasy was detected at the family level [i.e., the Cucurbitaceae to include luffa (Benincaseae tribe) and squash (Cucurbiteae tribe)].

The morphological (Jeffrey 1962, 1964), geographical, cytological and biochemical (Dunnill and Fowden 1965; Jeffrey 1990; Walters et al. 1991) relationships among genera within the Cucurbitaceae remain largely obscure. Sequencing of PCR products has been informative for identification of polymorphisms for taxonomic and phylogenetic analyses (Jobst et al. 1998; Sanjur et al. 2002).

A comprehensive assessment of the systematics of this family would require more-rigorous analyses (e.g., examination of chloroplast sequence and primer-specific fragment-length variation). The chloroplast primer set documented herein would be an appropriate tool for a more-extended genetic and evolutionary analyses of the Cucurbitaceae given its universal applicability for sequence-based phylogeny studies.

Amplifications of extended portions of the chloroplast genome by recombined consensus chloroplast primers (RCCPs)

Dumolin et al. (1997), Weising and Gardner (1999) and Grivet et al. (2001) reported the construction of ccmp and consensus primers that spanned specific but not complete portions of the *N. tabacum* chloroplast. There was a need to provide additional coverage to span the chloroplast genomic region from bp position 86,695 to 154,185. Thus, consensus chloroplast primers were developed and strategically arrayed as RCCPs (Table 3) to provide for complete coverage of the tobacco chloroplast genome (1 to 155,939 bp) through extended-length amplification. The amplification products after extended-length PCR using 16 RCCP pairs and cucumber template DNA having similar bp-lengths as expected for the tobacco chloroplast genome, indicate complete coverage of the cucumber chloroplast genome (Fig. 2). However, for some RCCPs (i.e., RCCP-8 and -15; Fig. 2) brightly staining unpredicted amplicons were observed. These bands could be representative of chloroplast DNA sequences that have been transferred to the nuclear genome via insertion-events during species evolution (Yuan et al. 2002). This hypothesis could be tested through sequence analysis of the unexpected amplicons. Regardless of such anomalies, the RCCP set constructed and documented herein, might be useful for the genetic investigation of traits under cytoplasmic control such as chilling injury in cucumber which is maternally controlled (Chung et al. 2003).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bryan GJ, McNicoll J, Ramsay G, Meyer RC, De JWS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. Theor Appl Genet 99:859–867

Chung S, Staub JE, Fazio G (2003) Inheritance of chilling injury: a maternally inherited trait in cucumber. J Am Soc Hort Sci (in press)

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–131

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evolution 50:2515–2520

Dumolin LS, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 6:393–397

Dunnill PM, Fowden L (1965) The amino acids of seeds of the Cucurbitaceae. Phytochemistry 4:933–944

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185–194

Hupfer H, Swiatek M, Hornung S, Herrmann RG, Maier RM, Chiu WL, Sears B (2000) Complete nucleotide sequence of the *Oenothera elata* plastid chromosome, representing plastome I of the five distinguishable Euoenothera plastomes. Mol Gen Genet 263:581–585

Grivet D, Heinze B, Vendramin GG, Petit RJ (2001) Genome walking with consensus primers: application to the large single-copy region of chloroplast DNA. Mol Ecol Notes 1:345–349

Ishii T, McCouch SR (2000) Microsatellites and microsynteny in the chloroplast genomes of *Oryza* and eight other Gramineae species. Theor Appl Genet 100:1257–1266

Jeffrey C (1962) Notes on Cucurbitaceae, including a proposed new classification of the family. Kew Bull 15:337–371

Jeffrey C (1964) A note on pollen morphology in the Cucurbitaceae. Kew Bull 17:473–476

Jeffrey C (1980) A review of the Cucurbitaceae. Bot J Linn Soc 81:233–247

Jeffrey C (1990) An outline classification of the Cucurbitaceae. In: Biology and utilization of the Cucurbitaceae. Cornell University Press, Ithaca, NY, pp 449–463

Jobst J, King K, Hemleben V (1998) Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of the family Cucurbitaceae. Mol Phylog Evol 9:204–219

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* L. Gaertn. Mol Ecol 7:1151–1161

McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. Trends Ecol Evol 10:198–202

Mohanty A, Martin JP, Aguinagalde I (2000) Chloroplast DNA diversity within and among populations of the allotetraploid *Prunus spinosa* L. Theor Appl Genet 100:1304–1310

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Complete nucleotide sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Plant Mol Biol Rep 4:148–175

Orti G, Pearse DE, Avise JC (1997) Phylogenetic assessment of length variation at a microsatellite locus. Proc Natl Acad Sci USA 94:10,745–10,749

Powell W, Morgante M, McDevitt R, Vendramin G, Rafalski J (1995) Polymorphic simple-sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. Proc Natl Acad Sci USA 92:7759–7763

Powell W, Morgante M, Doyle JJ, McNicol JW, Tingey SV, Rafalski AJ (1996) Genepool variation in the genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics 144:792–803

Provan J, Corbett G, Waugh R, McNicol JW, Morgante M, Powell W (1996) DNA fingerprints of rice (*Oryza sativa*) obtained from hypervariable chloroplast simple-sequence repeats. Proc Roy Soc Lond B 263:1275–1281

Provan J, Russell JR, Booth A, Powell W (1999) Polymorphic chloroplast simple-sequence repeat primers for systematic and population studies in the genus *Hordeum*. Mol Ecol 8:505–511

Sanjur OI, Piperno DR, Andres TC, Wessel-Beaver L (2002) Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: implications for crop plant evolution and areas of origin. Proc Natl Acad Sci USA 99:535–540

Sato S, Nakmura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res 6:283–290

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chungwongse J, Obokata J, Yam-

aguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049

Schmitz-Linneweber CR, Maier M, Alcaraz J, Cottet A, Herrmann RG, Mache R (2001) The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and gene organization. Plant Mol Biol 45:307–315

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. Hortscience 31:729–740

Staub JE, Serquen FC, McCreight JD (1997) Genetic diversity in cucumber (*Cucumis sativus* L.). III. An evaluation of Indian germplasm. Genetic Resour Crop Evol 44:315–326

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1110

Vendramin GG, Ziegenhagen B (1997) Characterization and inheritance of polymorphic plastid microsatellites in *Abies*. Genome 40:857–864

Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol Ecol 5:595–598

Walters TW, Decker Walters DS, Posluszny U, Kevan PG (1991) Determination and interpretation of comigrating allozymes among genera of the Benincaseae (Cucurbitaceae). Syst Bot 16:30–40

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9–19

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNA. Proc Natl Acad Sci USA 84:9054–9058

Yuan Q, Hill J, Hsiao J, Moffat K, Ouyang S, Cheng Z, Jiang J, Buell CR (2002) Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. Mol Gen Genomics 267:713–720