RESEARCH ARTICLE https://doi.org/10.7235/HORT.20210057

Development of Plastid InDel Markers to Discriminate Lemons from Other Citrus Groups

Sang Suk Kim[1†], Ho Bang Kim[2†], Kyung Jin Park[3], Jae Wook Hyun[1], Cheol Woo Choi[1], Jae-Ho Joa[1], Seong Beom Jin[1], Eun-Sil Kim[2], and Seung Gab Han[1*]

[1]Citrus Research Institute, National Institute of Horticultural and Herbal Science, Seogwipo 63607, Korea
[2]Life Sciences Research Institute, Biomedic Co., Ltd., Bucheon 14548, Korea
[3]Planning and Coordination Division, National Institute of Horticultural and Herbal Science, Wanju 55365, Korea

*Corresponding author: skhan@korea.kr

†These two authors equally contributed to this research.

Abstract

Lemon (*Citrus limon*), an interspecific hybrid between sour orange and citron, has been widely used as a rootstock along with trifoliate orange. Though lemons are superior to trifoliate orange in terms of their high seed germination rate throughout the year, one of the obstacles to using lemons as rootstocks is the lack of reliable, lemon-specific molecular markers to discriminate buds of the micro-grafted scion from those of the lemon rootstock. In order to obtain lemon-specific molecular markers, we compared the whole-plastid genomes available from four citrus species (lemon, pummelo, sweet orange, and mandarin) and developed seven plastid insertion/deletion (InDel) markers. The plastid InDel markers were applied to 46 citrus accessions that included lemons (17 accessions), grapefruit, mandarin, pummelo, sour orange, orange, papeda, tangor, and tangelo groups. The resulting dendrogram revealed that the citrus accessions used in this analysis could be distinctly classified into seven clusters. Lemons formed a separate cluster and had identical allele sizes for each InDel locus among all accessions investigated. This set of InDel markers could be a useful molecular tool for the rapid and clear discrimination of micro-grafted scions and lemon rootstocks during the production of virus-free citrus trees. The plastid InDel markers with maternal inheritance features can also be used to analyze the phylogenetic origin of various citrus cultivars including lemons.

Additional key words: *Citrus limon*, genotyping, maternal inheritance, organelle genome, rootstock, scion

Received: February 8, 2021
Revised: April 22, 2021
Accepted: May 24, 2021

HORTICULTURAL SCIENCE and TECHNOLOGY
39(5):637-644, 2021
URL: http://www.hst-j.org

pISSN : 1226-8763
eISSN : 2465-8588

This work was supported by the Golden Seed Project (Project No. 213007-05-5-SBQ40), Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS), and Cooperative Research Program for Agriculture Science & Technology Development (Project No. 014473012), Rural Development Administration, Republic of Korea. The authors thank Dr. Chang Jae Oh at Life Sciences Research Institute, Biomedic Co., Ltd. for his critical reading of the manuscript.

Introduction

Lemon [*Citrus limon* (L.) Burm. F.] is considered to be native to the Southern part of the Himalayas in India or Southern China and probably Upper Burma (Nicolosi, 2007). Recent molecular studies suggest that lemons (yellow lemon types) resulted from the interspecific hybridization between sour orange (*C. aurantium* L.) and citron (*C. medica* L.) (Nicolosi et al., 2000; Curk et al., 2015; Curk et al., 2016). Sour orange itself is believed to be a hybrid between pummelo (*C. maxima* Burm.) and mandarin (*C. reticulata* Blanco) (Wu et al., 2018). However, volkamer and rough lemons are the

products of hybridization between mandarin and sour orange (Curk et al., 2016). Lemons are generally divided into three types; 'Eureka', 'Lisbon', and others (Park et al., 2020).

Lemons are cultivated in Mediterranean, subtropical, and intertropical climates worldwide. However, only a few lemon varieties are cultivated worldwide for the production of fresh fruits and essential oils or for use as rootstocks, although substantial genetic diversity exists in this citrus group (Curk et al., 2016). In Korea, lemons have been commercially cultivated since the year 2000 and around 100 farms currently produce lemons from a cultivated area of 30 hectares. This cultivation area is expected to increase due to continuous market demands and farm profitability in Korea. Eureka-type lemons, such as 'Allen', 'Frost', and 'Cook', are mainly cultivated in Korea. It is difficult to clearly identify lemon varieties, even though various types of lemons, including 'Lisbon' and 'Meyer', are also assumed to be introduced into Korea. Microsatellite markers would be a robust molecular tool for identifying citrus cultivars including lemons (Woo et al., 2020).

Trifoliate orange [*Poncirus trifoliata* (L.) Raf.] rootstocks have been widely used in the citrus industry, including in Korea, due to several merits including a compact canopy, fruit quality, cold hardiness, and resistance to pathogens and pests. The rootstock trifoliate is quite suitable for satsuma mandarin (*C. unshiu* Marc.) that is the most prevalent citrus in Korea, accounting for approximately 87% of the citrus production in the year 2019. Recent trends for diversification of commercial citrus cultivars resulted in increased demand for more rootstock varieties (Kawase et al., 1987; Donadio et al., 2019). Various rootstock varieties such as lemons and their hybrids, lime, citrange, citrumelo, and Cleopatra mandarin were developed (URL https://citrusvariety.ucr.edu/citrus/rootstocks.html) and have been used in the citrus industry due to the drawbacks of trifoliate orange rootstocks, including incompatibility with several cultivars, virus problems, and nonvigorous growth (Kawase et al., 1987). In addition to these drawbacks, the seed germination rate of trifoliate orange rapidly decreases after May, whereas lemon seeds maintain a high germination rate throughout the year, guaranteeing high seedling yield and quality for rootstock production. However, a critical problem is that it is very hard to identify the genetic origin of new shoots when the micro-grafted scions are germinated on lemon rootstocks.

Several types of DNA-based molecular markers, such as simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs), and Insertions/Deletions (InDels) have been developed and used for the identification of nucellar/zygotic individuals (Woo et al., 2019), the efficiency test of hybrid embryo rescue (Kim et al., 2020), species/cultivar identification (Woo et al., 2020; Jin et al., 2020), and phylogenetic origin analysis in citrus (Curk et al., 2016). SSRs and InDel markers from the mitochondria and/or chloroplast genomes are widely used for the identification and authentication of crop cultivars (Curk et al., 2016; Park et al., 2017; Park et al., 2020; Roy et al., 2020).

In this study, we developed reliable, lemon-specific, multi-locus InDel markers from the comparative analysis of citrus plastid genomes including lemon, pummelo, sweet orange, and mandarin. This set of markers would be a useful molecular genetic tool for the rapid and clear discrimination of micro-grafted scion cultivars and lemon rootstocks during the production of virus-free citrus trees and for the guarantee of nursery stock quality in the citrus industry. The InDel markers from the maternally inherited plastid genome can also be used to analyze the phylogenetic origin of various citrus cultivars including lemons.

Materials and Methods

Plant Materials and Genomic DNA Extraction

The citrus samples used in this research were obtained from two public institutes in the Republic of Korea: Citrus Research Institute, National Institute of Horticultural and Herbal Science, Rural Development Administration, and Agricultural Research and Extension Services, Jeju Special Self-Governing Province. Leaf tissues were rinsed with running tap water and then stored at −70°C until use. Genomic DNA (gDNA) was purified from the leaf tissue using the Biomedic® Plant gDNA Extraction Kit (Biomedic Co., Ltd., Bucheon, Korea). The quantity and quality of the purified gDNA were determined by the DeNovix DS-11+ spectrophotometer (DeNovix, Wilmington, DE, USA) and agarose gel electrophoresis.

Mining of Polymorphic Plastid InDel loci and Primer Design

In order to develop polymorphic InDel markers in the organelle genomes of citrus, we compared the complete plastid genomes of four citrus species, *C. limon* (GenBank accession No. NC_034690), *C. maxima* (NC_034290), *C. reticulata* (NC_034671), and *C. sinensis* (NC_008334) by multiple sequence alignment using the CLC Genomics Workbench (ver. 6.8.4; Qiagen, Aarhus, Denmark). Putative polymorphic InDel markers were preliminarily screened by routine PCR using gDNAs from representative varieties of the four citrus species. The PCR reaction and cycling conditions were described in our previous report (Woo et al., 2019). PCR products were separated on a 2.5% (w/v) agarose gel to confirm PCR amplification and polymorphism among the varieties tested.

Genotyping by M13-tailed PCR and Data Analysis

The M13-tailed PCR method was used for genotype analysis using the selected polymorphic InDel markers (Schuelke, 2000). The PCR reaction condition was the same as in our previous report (Woo et al., 2019). The cycling conditions for PCR amplification followed the protocol described previously by Schuelke (2000). Fragment analysis of the PCR products was described previously (Kim et al., 2012). Calling of allele sizes was performed using the GeneMapper software (ver. 4.0; Applied Biosystems, Foster City, CA, USA). The unweighted pair group method with arithmetic mean (UPGMA) dendrogram was constructed using MEGA software (v. 7.0) (Kumar et al., 2016), which is embedded in PowerMarker (v.3.25; Liu and Muse, 2005). Genetic parameters such as major allele frequency, number of alleles, genetic diversity, and polymorphic information content were measured by calculating the shared allele frequencies using PowerMarker software (v. 3.25; Liu and Muse, 2005).

Results and Discussion

Development of Polymorphic InDel Markers from Citrus Plastid Genomes

Since the complete chloroplast genome of *C. sinensis* (L.) Osbeck 'Ridge Pineapple' was first reported in 2006 (Bausher et al., 2006), the complete plastid genomes of nine citrus species are currently available from the organelle genome resources of the National Center of Biotechnology Information (NCBI) (URL https://www.ncbi.nlm.nih.gov/

Table 1. List of primers for the amplification of seven InDel loci

Locus	Primer name	Primer sequence (5' to 3')	Expected amplicon size (bp)	Target region[z]
Limon#01	Limon#01-F2	GCGCATACCAACAATATCAT	261	310 to 570 (trnH-GUG/psbA)
	Limon#01-R2	TGCATGAACGTAATGCTCAT		
Limon#13	Limon#13-F1	AGGGTCGGTCTTGAAACA	325	34,198 to 34,522 (trnE-UUC/trnT-GGU)
	Limon#13-R1	AAGGCCAAAAAGCCCCTT		
Limon#15	Limon#15-F1	CTAACAATTACGAGAATCTAG	351	49,783 to 50,133 (trnT-UGU/ trnL-UAA)
	Limon#15-R1	CGAATTAGAATAGAGCAAATTT		
Limon#16	Limon#16-F2	AGTGATATGGCTCGCCATA	400	51,821 to 52,220 (trnF-GAA/ ndhJ)
	Limon#16-R1	ATGCCTGAAAGTTGGATAGG		
Limon#21	Limon#21-F1	TATCGAGGGGCTTTTCTTC	252	75,351 to 75,602 (2nd intron of clpP)
	Limon#21-R1	ATCAAAATCGGGCGAATCC		
Limon#22	Limon#22B-F1	TAGTGTCCTTGCCCATGA	255	84,574 to 84,828 (rps8/rpl14)
	Limon#22B-R1	AACTCGAGTTTTTGGTGC		
Limon#23	Limon#23-F2	TGTAGACCCTCGCAATAGTT	325	87,397 to 87,721 (rps3/rpl22)
	Limon#23-R2	ATGGGTCCTACTGCGAAA		

[z]Location on C. limon plastid, complete genome (160,101 bp, GenBank accession No. NC_034690)

genome/browse#!/organelles/citrus). Recently, the complete chloroplast genome sequence of a medicinal landrace citrus 'Jinkyool' (*C. sunki*) in Jeju, Korea was also reported (Yoo et al., 2020). From the comparative *in silico* analysis of the whole-plastid genomes of *C. limon*, *C. maxima*, *C. sinensis*, and *C. reticulata*, we mined 20 putative major InDel loci (> 2 bp) specific to lemon or the other citrus species. From the primary screening by routine PCR using several citrus accessions of the four citrus species, we finally selected seven plastid InDel loci out of eight candidates (Table 1). Amplicon sizes of the selected InDel loci ranged from 252 (Limon#21) to 400 base pairs (bps) (Limon#16). All of the selected InDel loci were located in inter-genic spacers or introns (Table 1 and Suppl. Fig. 1s). This result corresponds to the fact that high-resolution of DNA barcodes for species identification and phylogenetic analysis in plants have been primarily found in the non-coding regions of the plastid genome (Hollingsworth et al., 2011; Jiao et al., 2019). The forward primers were tagged with 18-mer of the M13 sequence for efficient fragment analysis of the seven InDel loci by the cost-effective M13-tailed PCR method (Schuelke, 2000) (Suppl. Table 1s).

Discrimination of lemons from Other Citrus Groups using Polymorphic InDel Markers

In this study, we used 46 citrus accessions belonging to lemon, grapefruit, mandarin, pummelo, sour orange, orange, papeda, tangor, and tangelo groups, which were obtained from two public institutes in Korea (Table 2). Seven polymorphic InDel markers were applied to the citrus accessions to determine their genotypes (Suppl. Table 2s). Table 3 shows the genetic characteristics of the plastid InDel loci based on the genotype analysis of 46 citrus accessions. A total of 23 alleles, ranging from 2 (Limon#13, Limon#21, and Limon#22) to 5 (Limon#15 and Limon#23) per locus, were observed among the 46 accessions, with an average of 3.3 alleles per locus. Major allele frequency (M_{AF}) varied from 0.37 to 0.76. The average genetic diversity (GD) value was 0.53, ranging from 0.36 to 0.74, and the average polymorphism information content (PIC) value was 0.47, ranging from 0.30 (Limon#13, Limon#21, and Limon#22) to 0.69 (Limon#23).

Table 2. Citrus samples used in this study

Citrus group	Species/Cultivar name	Abbreviation
Lemon	*C. limon* 'Allen'	LM-AN
	C. limon 'Frost Eureka'	LM-FN
	C. limon 'Eureka'	LM-UR
	C. limon 'Lisbon'	LM-RM
	C. limon 'Jeramon'	LM-JRM
	C. limon 'Dr. Strong Lisbon'	LM-DS
	C. limon 'Frost Lisbon'	LM-FL
	C. limon 'Genoa'	LM-JN
	C. limon 'Interdonato'	LM-ID
	C. limon 'Lapithiotiki'	LM-LT
	C. limon 'Limonero Fino 49'	LM-LP
	C. limon 'Limonero Messina'	LM-LM
	C. limon 'Utt Allen Eureka'	LM-AAE
	C. limon 'Villafranca'	LM-VF
	C. limon 'Yenben Lisbon'	LM-YB
	C. limon 'Prior Lisbon'	LM-PL
	C. limon 'Meyer 806'	LM-MY
Grapefruit	*C. paradisi*	GF
	C. paradisi 'Redblush'	RB
	C. paradisi 'Ray Ruby'	RR
Mandarin	*C. reticulata*	BG
	C. erythrosa	DJK
	C. platymamma	BK
	C. clementina 'Oroval'	OV
	C. sunki	JG
Pummelo	*C. maxima* 'Mato Buntan'	MM
	C. maxima 'Banbeiyu'	MB
Sour orange	*C. obovoidea*	KG
	C. natsudaidai 'Tachibana Orange'	LHO
	C. natsudaidai 'Whanggumhagyul'	HH
Orange	*C. sinensis* 'Sanguinelli'	SG
	C. sinensis 'Delfino'	DP
	C. sinensis 'Moro'	MR
	C. sinensis 'Bream Tarocco'	BTR
Papeda	*C. sphaerocarpa* 'Oita Kobosu'	DB
	C. wilsonii 'Inchangkyul'	IC
	C. junos	SYJ
	C. sudachi	SDC
Tangor	*Citrus hybrid* 'Hinoyutaka'	BP
	Citrus hybrid 'Kiyomi'	CK
	Citrus hybrid 'Okitsu no. 46'	HJ46
	Citrus hybrid 'Harehime'	HM
	Citrus hybrid 'Winter Prince'	PR
Tangelo	*Citrus hybrid* 'Satonokaori'	STNK
	Citrus hybrid 'Seminole'	SN
	Citrus hybrid 'Fallglo'	FG

Table 3. Characteristics of the seven plastid InDel loci based on the genotype analysis of 46 citrus accessions including 17 lemon resources

Locus	SS	NOBS	Availability	N_G	M_{AF}	N_A	GD	PIC
Limon#01	46	46	1	4	0.59	4	0.58	0.52
Limon#13	46	46	1	2	0.76	2	0.36	0.30
Limon#15	46	46	1	5	0.37	5	0.73	0.68
Limon#16	46	46	1	3	0.59	3	0.56	0.50
Limon#21	46	46	1	2	0.76	2	0.36	0.30
Limon#22	46	46	1	2	0.76	2	0.36	0.30
Limon#23	46	46	1	5	0.37	5	0.74	0.69
Mean	46	46	1	3.3	0.60	3.3	0.53	0.47

SS, sample size; N_{OBS}, number of observations; Availability is defined as 1-OBS/n, where OBS is the number of observations and n is the number of individuals sampled.; N_G, Genotype number; M_{AF}, major allele frequency; N_A, number of alleles; GD, genetic diversity. GD, often referred to as expected heterozygosity, is defined as the probability that two randomly chosen alleles from the population are different.; PIC, polymorphism information content.

A total of 23 alleles derived from the seven polymorphic InDel loci were used to analyze the genetic relationships among the 46 citrus accessions. An UPGMA dendrogram was constructed based on the genetic similarity matrices among the accessions. Fig. 1 illustrates the results of the cluster analysis based on the genotype data of the seven plastid InDel loci. The resulting dendrogram revealed that the accessions used in the analysis could be distinctly classified into seven groups, forming a separate lemon cluster. This result suggests that these seven plastid InDel loci comprise a robust marker set to reliably discriminate scion cultivars from lemon rootstocks. Phylogenetic analysis using mitochondrial InDel and chloroplastic SSR markers showed that all *C. limon* accessions investigated clustered to the sour orange group with *C. aurantium* and *C. limetta*, etc., and were separated from the other five groups, which were the citrons (including *C. medica* only), *C. micrantha* (including *C. aurantifolia* and *C. aurata*, etc.), edible mandarins (including *C. kinokuni*, *C. deliciosa*, and *C. clementina*, etc.), wild mandarins (including *C. sunki*, *C. reticulata*, and *C. limonia*, etc.), and pummelos (including *C. maxima* and *C. paradisi*, etc.) (Curk et al., 2016). This report supports our dendrogram forming a lemon group that is separated from other citrus species (Fig. 1). Similar tree topology was also observed from the dendrogram of citrus genetic resources that were constructed using 17 nuclear SSR markers (Woo et al., 2019; Woo et al., 2020).

Molecular genetic evidences using intergeneric sexual hybrids revealed the maternal inheritance pattern of chloroplast DNA in citrus (Moreira et al., 2002; Abkenar et al., 2004). Our data also support a maternal pattern of organelle inheritance in citrus. As an example, cluster B contained three mandarins (BG, JG, and OV), five tangors (BP, CK, HM, PR, and HJ46), and two tangelos (STNK and FG). Tangors and tangelos originated from the genetic crosses of mandarin x *C. sinensis* and mandarin x *C. maxima* (or mandarin x *C. paradisi*), respectively (Park et al., 2020). Especially, three tangors (BP, HM, and PR) and one tangelo (STNK) shared *Citrus* hybrid 'Kiyomi' (CK) as a grandmother, which originated from the genetic combination of *C. unshiu* and *C. sinensis* (Park et al., 2020). According to the Swingle system, *C. unshiu* is considered to be a group of mandarin varieties (Froelicher et al., 2011). Our result suggests that all citrus cultivars of cluster B share mandarin as a female parent, further supporting the maternal inheritance of the chloroplast genome in *Citrus*.

DNA barcoding is an effective tool that enables rapid and accurate identification of plant species. However,

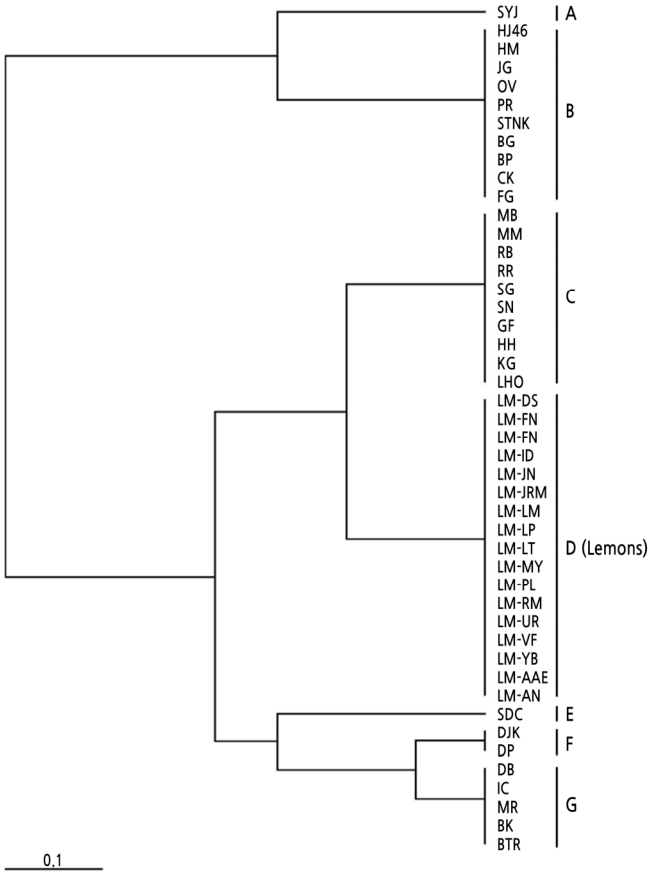


Fig. 1. An UPGMA dendrogram based on the genetic distances among 46 citrus accessions using seven plastid InDel markers.

single-locus DNA barcodes lack adequate variation in closely related taxa. Recent accumulation of whole-plastid genome sequences within closely related taxa has allowed researchers to develop cost-effective and reliable barcodes for accurate plant identification (Li et al., 2015). As shown in this study, reliable, multi-locus barcodes from the comparative analysis of whole-plastid genomes can increase our ability to distinguish closely related horticultural taxa such as citrus.

Literature Cited

Abkenar A, Isshiki S, Tashiro Y (2004) Maternal inheritance of chloroplast DNA in intergeneric sexual hybrids of "true citrus fruit trees" revealed by PCR-RFLP analysis 79:360-363. doi:10.1080/14620316.2004.11511773

Bausher MG, Singh ND, Lee SB, Jansen RK, Daniell H (2006) The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. BMC Plant Biol 6:21. doi:10.1186/1471-2229-6-21

Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet J-P, Garcia-Lor A, Navarro L, Ollitrault P (2015) Nuclear species-diagnostic SNP markers mined from 454 amplicon sequencing reveal admixture genomic structure of modern citrus varieties. PLoS ONE 10:e0125628. doi:10.1371/journal.pone.0125628

Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P (2016) Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. Ann Bot 117:565-583. doi:10.1093/aob/mcw005

Donadio LC, Lederman IE, Roberto SR, Stucchi ES (2019) Dwarfing-canopy and rootstock cultivars for fruit trees. Rev Bras Frutic 41:e-997. doi:10.1590/0100-29452019997

Froelicher Y, Mouhaya W, Bassene JB, Costantino G, Kamiri M, Luro F, Morillon R, Ollitrault P (2011) New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. Tree Genet Genomes 7:49-61. doi:10.1007/s11295-010-0314-x

Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS ONE 6:e19254. doi:10.1371/journal.pone.0019254

Jiao L, Lu Y, He T, Li J, Yin Y (2019) A strategy for developing high-resolution DNA barcodes for species discrimination of wood specimens using the complete chloroplast genome of three *Pterocarpus* species. Planta 250:95-104. doi:10.1007/s00425-019-03150-1

Jin S-B, Kim HB, Park SM, Kim MJ, Choi CW, Yun S-H (2020) Identification of the 'Haryejosaeng' mandarin cultivar by multiplex PCR-based SNP genotyping. Mol Biol Rep 47:8385-8395. doi:10.1007/s11033-020-05850-4

Kawase K, Iwagaki I, Takahara T, Ono S, Hirose K (1987) Rootstock studies for citrus varieties in Japan. Jpn Agric Res Q 20:253-259

Kim HB, Jeon JH, Han AR, Lee Y, Jun SS, Kim TH, Cho GH, Park PB (2012) Genetic evaluation of domestic walnut cultivars trading on Korean tree markets using microsatellite markers. Afr J Biotechnol 11:7366-7374. doi:10.5897/AJB11.3647

Kim M, Kim SH, Kim HB, Park YC, Song KJ (2020) Some factors affecting the efficiency of hybrid embryo rescue in the 'Shiranuhi' mandarin. Hortic Sci Technol 38:271-281. doi:10.7235/HORT.20200026

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874. doi:10.1093/molbev/msw054

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90:157-166. doi:10.1111/brv.12104

Liu KJ, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128-2129. doi:10.1093/bioinformatics/bti282

Moreira CD, Gmitter Jr FG, Grosser JW, Huang S, Ortega VM, Chase CD (2002) Inheritance of organelle DNA sequences in a *Citrus-Poncirus* intergeneric cross. J Hered 93:174-178. doi:10.1093/jhered/93.3.174

Nicolosi E (2007) Origin and taxonomy. In IA Khan, ed, Citrus: Genetics, Breeding and Biotechnology. CAB International, Oxfordshire, UK, pp 19-43. doi:10.1079/9780851990194.0019

Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E (2000) Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor Appl Genet 100:1155-1166. doi:10.1007/s001220051419

Park HS, Jayakodi M, Lee SH, Jeon J-H, Lee H-O, Park JY, Moon BC, Kim C-K, Wing RA, et al. (2020) Mitochondrial plastid DNA can cause DNA barcoding paradox in plants. Sci Rep 10:6112. doi:10.1038/s41598-020-63233-y

Park SI, Hwangbo K, Gil J, Chung H, Kim HB, Kim OT, Kim SC, Koo SC, Um Y, et al. (2017) Determination of the origin of angelica roots using *Angelica gigas* chloroplast based SSR markers. Korean J Medicinal Crop Sci 25:361-366. doi:10.7783/KJMCS.2017.25.6.361

Park YC, Hong SY, Kang SH, Lee CH, Oh MH, Yang CJ (2020) Citrus genetic resources. Jeju Special Self-Governing Province, Jeju (in Korean)

Roy NS, Jeong U, Na M, Choi I-Y, Cheong EJ (2020) Genomic analysis and a consensus chloroplast genome sequence of *Prunus yedoensis* for DNA marker development. Hortic Environ Biotechnol 61:859-867. doi:10.1007/s13580-020-00265-3

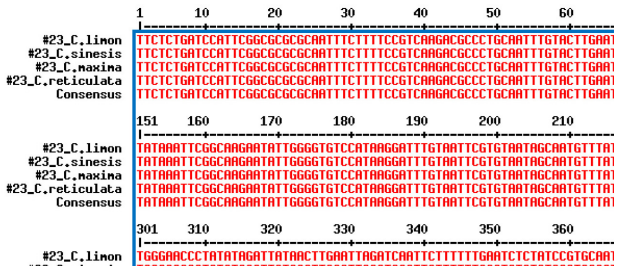
Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233-234. doi:10.1038/72708

Woo JK, Park YC, Lee JW, Yun SH, Kim M, Park S, Lee Y, Song KJ, Kim HB (2019) Evaluation of polyembryony for genetic resources and efficacy of simple sequence repeat markers for the identification of nucellar and zygotic embryo-derived individuals in citrus. Appl Biol Chem 62:30. doi:10.1186/s13765-019-0437-1

Woo JK, Yun S-H, Yi KU, Park YC, Lee H-Y, Kim M, Lee Y, Song KJ, Kim HB (2020) Identification of citrus varieties bred in Korea using microsatellite markers. Hortic Sci Technol 38:374-384. doi:10.7235/HORT.20200036

Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, Domingo C, Tadeo FR, Carbonell-Caballero J, et al. (2018) Genomics of the origin and evolution of Citrus. Nature 554:311-316. doi:10.1038/nature25447

Yoo Y-H, Oh CJ, Shin SC, Seo S, Kim M, Yun S-H, Song KJ, Lee H, Kim HB (2020) Complete chloroplast genome sequence of a medicinal landrace citrus Jinkyool (*Citrus sunki* Hort. ex Tanaka) in Jeju Island, Korea. Mitochondrial DNA Part B 5:3719-3720. doi:10.1080/23802359.2020.1833771



Supplementary Fig. 1s. Continued.

Supplementary Table 1s. List of primers for the genotype analysis of seven InDel loci by the M13-tailed PCR. Bold sequences indicate M13 tag

Target locus	Primer name	Primer sequence (5' to 3')
Limon#01	M13+Lim#01-F2	**TGTAAAACGACGGCCAGT**GCGCATACCAACAATATCAT
	Limon#01-R2	TGCATGAACGTAATGCTCAT
Limon#13	M13+Lim#13-F1	**TGTAAAACGACGGCCAGT**AGGGTCGGTCTTGAAACA
	Limon#13-R1	AAGGCCAAAAAGCCCCTT
Limon#15	M13+Lim#15-F1	**TGTAAAACGACGGCCAGT**CTAACAATTACGAGAATCTAG
	Limon#15-R1	CGAATTAGAATAGAGCAAATTT
Limon#16	M13+Lim#16-F2	**TGTAAAACGACGGCCAGT**AGTGATATGGCTCGCCATA
	Limon#16-R1	ATGCCTGAAAGTTGGATAGG
Limon#21	M13+Lim#21-F1	**TGTAAAACGACGGCCAGT**TATCGAGGGGCTTTTCTTC
	Limon#21-R1	ATCAAAATCGGGCGAATCC
Limon#22	M13+Lim#22B-F1	**TGTAAAACGACGGCCAGT**TAGTGTCCTTGCCCATGA
	Limon#22B-R1	AACTCGAGTTTTTGGTGC
Limon#23	M13+Lim#23-F2	**TGTAAAACGACGGCCAGT**TGTAGACCCTCGCAATAGTT
	Limon#23-R2	ATGGGTCCTACTGCGAAA

Supplementary Table 2s. Allele size data of 46 citrus accessions determined by seven polymorphic InDel markers

No.	Citrus group	Species/Cultivar name	Abbr.	Markers						
				Limon#1	Limon#13	Limon#15	Limon#16	Limon#21	Limon#22 B	Limon#23
1	Lemon	*C. limon* 'Allen'	LM-AN	279	343	369	418	270	272	343
2		*C. limon* 'Frost Eureka'	LM-FN	279	343	369	418	270	272	343
3		*C. limon* 'Eureka'	LM-UR	279	343	369	418	270	272	343
4		*C. limon* 'Lisbon'	LM-RM	279	343	369	418	270	272	343
5		*C. limon* 'Jeramon'	LM-JRM	279	343	369	418	270	272	343
6		*C. limon* 'Dr. Strong Lisbon'	LM-DS	279	343	369	418	270	272	343
7		*C. limon* 'Frost Lisbon'	LM-FL	279	343	369	418	270	272	343
8		*C. limon* 'Genoa'	LM-JN	279	343	369	418	270	272	343
9		*C. limon* 'Interdonato'	LM-ID	279	343	369	418	270	272	343
10		*C. limon* 'Lapithiotiki'	LM-LT	279	343	369	418	270	272	343
11		*C. limon* 'Limonero Fino 49'	LM-LP	279	343	369	418	270	272	343
12		*C. limon* 'Limonero Messina'	LM-LM	279	343	369	418	270	272	343
13		*C. limon* 'Utt Allen Eureka'	LM-AAE	279	343	369	418	270	272	343
14		*C. limon* 'Villafranca'	LM-VF	279	343	369	418	270	272	343
15		*C. limon* 'Yenben Lisbon'	LM-YB	279	343	369	418	270	272	343
16		*C. limon* 'Prior Lisbon'	LM-PL	279	343	369	418	270	272	343
17		*C. limon* 'Meyer 806'	LM-MY	279	343	369	418	270	272	343
18	Grape fruit	*C. paradisi*	GF	279	343	371	418	270	272	347
19		*C. paradisi* 'Redblush'	RB	279	343	371	418	270	272	347
20		*C. paradisi* 'RayRuby'	RR	279	343	371	418	270	272	347
21	Mandarin	*C. reticulata*	BG	273	328	380	419	262	277	364
22		*C. erythrosa*	DJK	278	343	371	417	270	272	345
23		*C. platymamma*	BK	278	343	370	417	270	272	345
24		*C. clementina* 'Oroval'	OV	273	328	380	419	262	277	364
25		*C. sunki*	JG	273	328	380	419	262	277	364
26	Pomelo	*C. maxima* 'Mato Buntan'	MM	279	343	371	418	270	272	347
27		*C. maxima* 'Banbeiyu'	MB	279	343	371	418	270	272	347
28	Sour orange	*C. obovoidea*	KG	279	343	371	418	270	272	347
29		*C. natsudaidai* 'Tachibana Orange'	LHO	279	343	371	418	270	272	347
30		*C. natsudaidai* 'Whanggumhagyul'	HH	279	343	371	418	270	272	347
31	Orange	*C. sinensis* 'Sanguinelli'	SG	279	343	371	418	270	272	347
32		*C. sinensis* 'Delfino'	DP	278	343	371	417	270	272	345
33		*C. sinensis* 'Moro'	MR	278	343	370	417	270	272	345
34		*C. sinensis* 'Bream Tarocco'	BTR	278	343	370	417	270	272	345
35	Papeda	*C. sphaerocarpa* 'Oita Kobosu'	DB	278	343	370	417	270	272	345
36		*C. wilsonii* 'Inchangkyul'	IC	278	343	370	417	270	272	345
37		*C. junos*	SYJ	278	328	370	419	262	277	360
38		*C. sudachi*	SDC	280	343	373	419	270	272	345
39	Tangor	*Citrus* hybrid 'Hinoyutaka'	BP	273	328	380	419	262	277	364
40		*Citrus* hybrid 'Kiyomi'	CK	273	328	380	419	262	277	364
41		*Citrus* hybrid 'Okitsuno.46'	HJ46	273	328	380	419	262	277	364
42		*Citrus* hybrid 'Harehime'	HM	273	328	380	419	262	277	364
43		*Citrus* hybrid 'Winter Prince'	PR	273	328	380	419	262	277	364
44	Tangelo	*Citrus* hybrid 'Satonokaori'	STNK	273	328	380	419	262	277	364
45		*Citrus* hybrid 'Seminole'	SN	279	343	371	418	270	272	347
46		*Citrus* hybrid 'Fallglo'	FG	273	328	380	419	262	277	364