Phylogenetic Analysis of Mandarin Landraces, Wild Mandarins, and Related Species in China Using Nuclear LEAFY Second Intron and Plastid trnL-trnF Sequence

Li Yingzhi¹, Cheng Yunjiang, Tao Nengguo, and Deng Xiuxin²
National Key Laboratory of Crop Genetic Improvement, Huazhong Agriculture University, Wuhan, Hubei Province, 430070, P. R. China

ADDITIONAL INDEX WORDS. Citrus reticulata, maximum parsimony, maximum likelihood, incongruence

ABSTRACT. Nucleotide sequences of the second intron of the nuclear LEAFY gene (FLint2) and trnL-trnF region of the chloroplast genome were used to analyze the phylogenetic relationships among eight wild mandarins (Citrus reticulata Blanco), 19 mandarin landraces, and 19 related species of Citrus L. Forty-six FLint2 sequences and 111 trnL-trnF sequences were obtained from 46 ingroup accessions, with an average length of 1059.7 and 776.7 bp respectively. Phylogeny reconstructions were conducted separately for these two data sets using maximum parsimony and maximum likelihood criteria. Monophyly of mandarins was supported by both of these data sets, and in this clade, most mandarin landraces formed an unresolved polytomy, whereas ‘Jiangyong 1’, ‘Chongyi A2’, ‘Chongyi A1’ (or ‘Jiangyong 4’ in FLint2 data), and ‘Daoxian 1’ wild mandarins formed a subclade. ‘Mangshan A1’ and ‘Daoxian 5’ wild mandarins were sisters to this mandarin clade. A hybrid origin of five mandarin landraces and several mandarin-related species was suggested as a plausible hypothesis to explain the incongruence between the FLint2 and trnL-trnF data sets.

Relationships among species of Citrus are complicated by several factors such as a high frequency of bud mutation, a long history of cultivation, nucellar embryony, and wide cross-compatibility among species (Moore, 2001). The genus Citrus itself has been described as consisting of 1 to 162 species. The most widely accepted classification systems are those of Swingle (1943) and Tanaka (1977), who recognized 16 and 162 species respectively. A major difference in these two systems is in how mandarins (C. reticulata) are treated; 36 mandarin species are recognized by Tanaka (1969, 1977) whereas only three mandarin species are noted by Swingle’s system (Swingle and Reece, 1967), including C. indica. Whereas only three mandarin species are recognized by Swingle’s (1943) and Tanaka (1977), who recognized 16 and 162 species respectively. A major difference in these two systems is in how mandarins (C. reticulata) are treated; 36 mandarin species are recognized by Tanaka (1969, 1977) whereas only three mandarin species are noted by Swingle’s system (Swingle and Reece, 1967), including C. tachibana (Makino) Tanaka and C. indica Tanaka, wild species from Japan and India respectively. In the mid 1970s, two studies (Barrett and Rhodes, 1976; Scora, 1975) suggested that cultivated Citrus comprises only three basic species: citrus (C. medica L.), mandarin (C. reticulata), and pummelo (C. maxima (Burm.) Merr.). Later, C. halimii B.C. Stone was added as the fourth basic species (Scora, 1988). The status of these species was also supported by studies using biochemical and molecular markers (Fang and Rose, 1997; Fang et al., 1993; Herrero et al., 1996; Luro et al., 1992; Nicolosi et al., 2000).

Mandarins are the most phenotypically heterogeneous group in Citrus. Both monoembryonic and polyembryonic clones exist, as do self-fertile and self-incompatible types (Moore, 2001). High genetic similarity among members of some mandarin cultivar groups has been detected in several studies (Coletta Filho et al., 1998; Esen and Scora, 1977; Machado et al., 1996), and great heterogeneity has been revealed within this group as well (Fang et al., 1998; Luro et al., 1995; Nicolosi et al., 2000; Torres et al., 1978). The mandarin group has been proposed to be composed of several genetically different individuals and a great number of hybrids (Coletta Filho et al., 1998).

China is believed to be the center of origin for citrus, and many mandarin landraces and wild mandarins are distributed there. In our previous study, high polymorphism was detected among these mandarin landraces and wild mandarins at simple sequence repeat (SSR) loci of both the chloroplast and nuclear genomes (Li et al., 2006). These genetic resources provide a good opportunity to study the origin and phylogenetic relationship of mandarins.

Molecular markers offer a way to solve many of the longstanding problems in the classification and breeding of Citrus (Moore, 2001). DNA sequences of the chloroplast genome are widely used in the field of molecular plant systematics. The trnL intron and trnL-trnF intergenic spacer exhibit a higher level of sequence variation among closely related species than the coding region, and vary in length and substitution rates (Clegg et al., 1994; Taberlet et al., 1991); therefore, they are more useful at lower taxonomic levels (Gielly and Taberlet, 1994). Recently, it was demonstrated that the trnL-trnF intergenic spacer is informative and effective in determining phylogenetic relationships of Korean Citrus species (Jung et al., 2005). Several other regions of the chloroplast genome have been also used to elucidate the phylogenetic relationships of Rutaceae (Chase et al., 1999; Morton et al., 2003; Samuel et al., 2001; Scott et al., 2000). On the other hand, few nuclear DNA regions have been explored, with the exception of the nuclear ribosomal DNA region, which has been sequenced on a...
Table 1. Taxa and their sources used in the study of phylogenetic relationships of mandarin landraces, wild mandarin, and related species using trnL-F and LEAFY second intron (FLint2) sequences and number of FLint2 clones sequenced.

<table>
<thead>
<tr>
<th>Accession name</th>
<th>Taxa*</th>
<th>Collecting places</th>
<th>FLint2 clones sequenced (no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangyong 1 w-mandarin</td>
<td>Citrus reticulata Blanco</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>Jiangyong 4 w-mandarin</td>
<td>C. reticulata</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>Daoxian 1 w-mandarin</td>
<td>C. reticulata</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>Daoxian 5 w-mandarin</td>
<td>C. reticulata</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>Chongyi A1 w-mandarin</td>
<td>C. reticulata</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>Chongyi A2 w-mandarin</td>
<td>C. reticulata</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>Mangshan B0 w-mandarin</td>
<td>C. reticulata</td>
<td>D</td>
<td>3</td>
</tr>
<tr>
<td>Mangshan A1 w-mandarin</td>
<td>C. reticulata</td>
<td>D</td>
<td>2</td>
</tr>
<tr>
<td>Hongju l-mandarin</td>
<td>C. reticulata</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Bendiguangji l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>4</td>
</tr>
<tr>
<td>Guoqingwuhaosatsumac-mandarin</td>
<td>C. reticulata</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Bendizao l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>2</td>
</tr>
<tr>
<td>Ponkan c-mandarin</td>
<td>C. reticulata</td>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>Shatanju l-mandarin</td>
<td>C. reticulata</td>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>Ruju l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>2</td>
</tr>
<tr>
<td>Manju l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>2</td>
</tr>
<tr>
<td>Manguohong l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>2</td>
</tr>
<tr>
<td>Chazhigan l-mandarin</td>
<td>C. reticulata</td>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>Nanfengjiumi l-mandarin</td>
<td>C. reticulata</td>
<td>F</td>
<td>2</td>
</tr>
<tr>
<td>Tuju l-mandarin</td>
<td>C. reticulata</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Jiangjinsuanju l-mandarin</td>
<td>C. reticulata</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Kuigan l-mandarin</td>
<td>C. reticulata</td>
<td>E</td>
<td>14</td>
</tr>
<tr>
<td>Ougan l-mandarin</td>
<td>C. reticulata</td>
<td>E</td>
<td>4</td>
</tr>
<tr>
<td>Choupigan l-mandarin</td>
<td>C. reticulata</td>
<td>A</td>
<td>14</td>
</tr>
<tr>
<td>Huapiju l-mandarin</td>
<td>C. reticulata</td>
<td>A</td>
<td>3</td>
</tr>
<tr>
<td>Huapigan l-mandarin</td>
<td>C. reticulata</td>
<td>A</td>
<td>4</td>
</tr>
<tr>
<td>Qingyi (yuzu?)</td>
<td>C. junos Sieb.ex Tan. ?</td>
<td>H</td>
<td>4</td>
</tr>
<tr>
<td>Valencia sweet orange</td>
<td>C. sinensis (L.) Osbeck</td>
<td>E</td>
<td>3</td>
</tr>
<tr>
<td>Shatianju pummelo</td>
<td>C. maxima (Burm.) Merr.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Eureka lemon</td>
<td>C. limon (L.) Burm. f.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Rangpur lime</td>
<td>C. limonia Osbeck</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>Red lemon</td>
<td>C. limonia</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>2 citron</td>
<td>C. medica L.</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Tibet citron</td>
<td>C. medica</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>861 citron</td>
<td>C. medica</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Yuzu</td>
<td>C. junos Sieb.ex Tan.</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>Ichang papeda</td>
<td>C. ichangensis Swing.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Honghe papeda</td>
<td>C. hongheensis Y.M. Ye et al.</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Dazhongcheng papeda</td>
<td>C. macrospora T.C.</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>Trifoliate orange</td>
<td>Poncirus trifoliata (L.) Raf.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Zhique (trifoliate orange?)</td>
<td>P. trifoliata ?</td>
<td>H</td>
<td>3</td>
</tr>
<tr>
<td>Wild hongkong kumquat</td>
<td>Fortunella hindsii (Champ. ex Benth.) Swing.</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>Meiwa kumquat</td>
<td>F. crassifolia Swing.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Calamondin</td>
<td>C. madurensis Loure.</td>
<td>E</td>
<td>3</td>
</tr>
<tr>
<td>Australian round lime</td>
<td>Microcitrus australis (A. Cunn. ex Mudie) Swing.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Australian finger lime</td>
<td>M. australasica (F. Muell.) Swing.</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Chinese box orange</td>
<td>Severinia buxifolia (Poir.) Tenore</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>Wamppee</td>
<td>Clausena lancea (Lour.) Skeels</td>
<td>E</td>
<td>2</td>
</tr>
</tbody>
</table>

* c-mandarin, mandarin cultivar; l-mandarin, mandarin landrace; w-mandarin, wild mandarin.
* Taxonomic treatment of mandarin followed Swingle’s system (Swingle and Reece, 1967).

* A, Jiangyong, Hunan Province, China; B, Daoxian, Hunan Province, China; C, Chongyi, Jiangxi Province, China; D, Mangshan, Hunan Province, China; E, Citrus Research Institute of Huazhong Agricultural University, Wuhan, Hubei Province, China; F, Huangyan, Zhejiang Province, China; G, Sihui, Guangdong Province, China; H, Nanzhen, Shannxi Province, China; I, Citrus Research Institute of Chinese Academy of Agricultural Sciences, Chongqing, China.
large scale. From this region, especially the internal transcribed spacers (ITS1 and ITS2) are extensively applied for phylogeny reconstruction at low taxonomic levels. However, ITS sequences are sometimes unsuitable for phylogenetic studies as a result of high sequence divergence (Wilson, 2003), extensive length variations between copies (Liston et al., 1996), paralogy resulting from high sequence divergence (Wilson, 2003), extensive sequence divergence; therefore, Chinese box orange was used as the outgroup. Outgroup comparisons with wampee were not amenable to use for phylogenetic comparison of FLnt2 sequences because of significant sequence divergence; therefore, Chinese box orange was used as the outgroup.

**Material and Methods**

**Plant materials and DNA extraction.** Samples from 39 biotypes from seven species of *Citrus*, (among them, 18 were mandarin landraces and eight were wild mandarins) were collected. An additional seven accessions were also included to represent related genera: *Fortunella* Swingle, *Poncirus* Raf., and *Microcitrus* Swingle. Among mandarin accessions, *Hua* (Cheng et al., 2005), *Li* (Cheng et al., 2005), and relative accessions. The two data sets were analyzed to reconstruct independent phylogenetic trees with maximum likelihood and parsimony criteria to ascertain the robustness of their nodes.

**Polymerase chain reaction amplification and purification of its products.** Two primers, LFY1 5’-ATGC CGTGATTTCTTGATCC 3’ and LFY2 5’-TGGCATCAATA TCCCAACTT 3’, were used in the amplification of *FLnt2* for all accessions (N.G. Tao and X.X. Deng, unpublished). The chloroplast trnL-trnF region was amplified using c and f primers (Taberlet et al., 1991).

The target regions were amplified via polymerase chain reaction (PCR) in a Peltier Thermal Cycler-200 thermocycler (MJ Research, Watertown, MA). For each PCR reaction, ~50 ng total DNA was included in a 20-μL reaction, consisting of 0.2 mm deoxynucleosides (dNTPs) (Shanghai Sangon Biological Engineering Technology and Services Corp., Shanghai, China). 0.1 μm forward and reverse primers (Shanghai Sangon), 1× PCR buffer (Fermentas Canada, Burlington, ON, Canada), 1.5 mm magnesium chloride (Fermentas), and 1 U Taq DNA polymerase (Fermentas). Thirty-five cycles of a three-step PCR followed by a final extension at 72 °C for 10 min were performed: denaturation at 94 °C for 1 min, primer annealing at 55 °C for 40 s, and primer extension at 72 °C for 1 min.

Polymerase chain reaction products were separated on a 1.0% Tris-Acetate-EDTA (TAE)-agarose gel, cut out of the gel, and cleaned using E.Z.N.A. Gel Extraction Kit (Omega Bio-Tek, Doraville, GA) following the manufacturer’s protocols. The purified PCR products were cloned into the Takara pMD18-T vector (Takara Bio, Shiga, Japan), and inserts from one to four transformed colonies were sequenced with M13– and M13+ primers.

All sequences were determined at Beijing SunBiotech Corp. (Beijing, China), which used an ABI 3770 automate sequencer (Applied BioSystems, Foster City, CA). Sequences were edited in Sequencher 4.1 (GeneCodes Corp., Ann Arbor, MI).

**Sequence alignments.** For FLnt2, determined sequences were compared with LEAFY complete sequences of sweet orange *C. sinensis* (L.) Osbeck, National Center for Biotechnology Information accession no. AF338976 and the exon–intron boundaries were identified. Sequences were aligned using Clustal X (Thompson et al., 1997) and adjusted manually as needed. All sequences were deposited in GenBank databases, and aligned data matrices with the NEXUS file format (Maddison et al., 1997) were submitted to TreeBASE (Morell, 1996). These are available from the first author upon request.

Pairwise sequence divergence was calculated using Kimura’s (1980) two-parameter (K2P) method in MEGA 3.0 (Kumar et al., 2004), with gaps/missing data completely deleted.

**Outgroup selection.** The data sets were polarized by the outgroup method. For the trnL-trnF data set, wampee and Chinese box orange were used as the outgroup. Outgroup comparisons with wampee were not amenable to use for phylogenetic comparison of FLnt2 sequences because of significant sequence divergence; therefore, Chinese box orange was used as the outgroup.

**Maximum parsimony analysis.** The maximum parsimony analyses were conducted using a heuristic search with tree bisection-reconnection (TBR) branch swapping, collapse of zero-length branches, and accelerated transformation. Characters were equally weighted; character states were specified as unordered. Random taxon addition (100 replicates) was used to search for multiple islands of trees. The strict consensus tree was generated based on the equally most-parsimonious trees produced by heuristic search. Bootstrap analysis (Felsenstein, 1985) using heuristic search with random addition on 1000 bootstrap replicates was performed to determine relative support for various clades found in the parsimony analysis. All analyses were performed using PAUP* 4.0b10 for Windows (Swofford, 2000).

In maximum parsimony analysis, we first treated gaps as missing data, and then treated them as fifth state to determine whether there were any phylogenetic topology differences between the two methods.

**Maximum likelihood analysis.** In all maximum likelihood analyses, heuristic searches with 500 replicates of random taxon addition, TBR branch swapping, and the MultiTrees option were used. The best-fitting evolutionary model for each data in

the maximum likelihood analysis was determined by the akaike information criterion using Modeltest 3.7 (Posada and Cran-dall, 1998). Gaps were treated as missing data and were excluded from analysis in maximum likelihood analysis.

**Kishino–Hasegawa test and Shimodaira–Hasegawa test.** To test which tree was best supported by sequence data sets, we used the Kishino–Hasegawa test (KH test) and the Shimodaira–Hasegawa test (SH test) as implemented in PAUP* 4.0b10 (Swofford, 2000), with a model for each of the two data sets estimated using Modeltest 3.7 (Posada and Crandall, 1998), the reestimated log likelihood method, and 1000 bootstrap replicates to compare different trees.

### Results

**Variations among trnL-trnF sequences.** Forty-eight sequences were obtained in this study. The trnL intron and trnL-trnF intergenic spacer of *Citrus* and its relatives ranged in size from 1009 bp [dazhonghongcheng papeda (*C. macrosperma* T.C. Guo et Y.M. Ye)] to 1083 bp (Tibet citron), with an average of 1059.7 bp. Nucleotide compositions were T, 28.8%; C, 19.3%; A, 33.6%; and G, 18.3%; and there were no differences in nucleotide compositions between mandarin landraces and wild populations.

Pairwise divergence between sequences was 0.00 to 0.018 in ingroup, and was 0.11 to 0.025 between outgroup and ingroup, with an average of 0.006. Other study showed that the greatest sequence divergence among Korean citrus was between sour orange (*C. aurantiom L.*) and *C. natsudaidai* Hay. (0.032) (Jung et al., 2005), different germplasm and different sequence length used in this study resulted in difference in sequence divergences. The transition-to-transversion ratio was 0.8.

The aligned sequences, with introduced gaps, had a length of 1157 bp. The matrix contained 827 constant sites, 77 automorphic sites, and 17 parsimony-informative sites for the 48 accessions. Postulated insertions/deletions ranged from 1 to 27 bp in length. Mangshan A1 wild mandarin had a 5-bp insertion, australain round lime [*Microcitrus australis* (A. Cunn. ex Mudie) Swingle] had a 27-bp insertion, and *Fortunella* sp. had a 27-bp insertion. All mandarin landraces, except possible mandarin-related hybrids (Huapigan, Choupigan, Bendiguangju, Kuigan, and Ougan), wild mandarins, except two wild mandarins (Mangshan A1 and Daokian 5), rangpur lime (*C. limonia* Osbeck) and red lemon (*C. limonia*), shared a 6-bp insertion in the same site. Kumquat (*Fortunella* sp.) biotypes, citron (*C. medica*) biotypes, australain round lime (*M. australis*), and australian finger lime [*M. australasica* (F. Muell.) Swing.] shared a 6-bp deletion. Honghe papeda (*C. hongheensis* Y.M. Ye et al.) had a 5-bp insertion. None of these indels were parsimony informative.

After exclusion of indels, the matrix contained 921 bp, including 94 variable sites and 17 parsimony-informative sites. The trnL-trnF data set used to study the phylogeny of Korean
citrus was 372 bp in length, with 67-bp variable sites (Jung et al., 2005). Again, a different germplasm and different sequence length used in these two studies resulted in a difference in sequence divergences.

**SEQUENCE VARIATIONS AMONG LEAFY SECOND INTRON (FLInt2).** Polymerase chain reaction amplification with LFY1 and LFY2 primers produced a single band in all accessions. For each accession, we sequenced more than two clones following the study of phylogenetic utility of FLInt2 in Neillia D. Don and Stephanandra Sieb. et Zucc. (Oh and Potter, 2003). As a result, we obtained 135 sequences from 48 biotypes. The length of sequences of FLInt2 examined in this study varied from 756 to 783 bp, with an average of 776.7 bp. This variation was incited by one poly T tract and several indels in this region. The average nucleotide compositions were A, 26.7%; T, 30.4%; G, 18.3%; and C, 24.6%; and there were also no differences between mandarin landrace populations and wild mandarins.

Pairwise sequence divergence of K2P distance ranged from 0.00 to 0.026 among the ingroup species, and from 0.023 to 0.035 between the ingroup taxa and the outgroup taxa, with an average of 0.012. The transition-to-transversion ratio was 1.0.

Nucleotide sequences of FLInt2 from wampee were highly divergent from other samples; therefore, they were excluded from phylogenetic analysis. Because no significant differences were detected between different clones, only one clone per sample was included in the reduced data set. The final alignment of this reduced data set included 795 characters, of which 36 were phylogenetically informative and 91 were variable. Two insertions and one deletion had to be assumed to align sequences from chinese box orange and australian finger lime respectively.

After exclusion of indels, the data matrix was 753 bp in length, of which 91 sites were variable and 36 sites were parsimony informative.

**PHYLOGENETIC ANALYSIS USING trnL–trnF DATA SET.** We used two data sets for maximum parsimony analysis based on cpDNA sequences. First, the data set of 921 characters, excluding all indels, was analyzed. We obtained 16 equally most-parsimonious trees of 104 steps, a consistency index (CI) of 0.952 and retention index (RI) of 0.926. The strict consensus tree is shown in Fig. 1, with a tree length of 113 steps, a CI of 0.876, and an RI of 0.794.

Second, we added all indels and scored them as a fifth character. Phylogenetic analysis using 1157 characters yielded eight most-parsimonious trees of 515 steps with a CI of 0.680 and an RI of 0.682. A strict consensus tree (CI, 0.655; RI, 0.645; tree length, 534 steps) is shown in Fig. 2, and had more structures and increased branch resolution than the indels-excluded parsimony tree.

The most appropriate model for the trnL–trnF data determined in Modeltest was the TVM+G model with a shape parameter of gamma distribution, $\alpha = 0.4240$, and a substitution matrix of six models. The best tree ($-\ln L = 2595.20961$) in the maximum
likelihood analysis (Fig. 3) using the model showed a few topology differences from the two strict consensus maximum parsimony trees.

The maximum likelihood tree was the best tree in the KH tests and the SH test, and the 16 indel information-excluded most-parsimonious trees were not different significantly from the maximum likelihood tree, whereas the eight indel information-included most-parsimonious trees were rejected for significantly different ($P < 0.05$).

All mandarin landraces, except five possible hybrid biotypes (Choupigan, Bendiguangju, Kuigan, Ouegan, and Huapigan), formed a moderately supported monophyletic clade (hereafter called the mandarin clade I) in both parsimony [Figs. 1 and 2; bootstrap percentage (BP) 71% to 91%] and likelihood (Fig. 3) trees. An exception was found for the Tuju mandarin landrace, which was placed outside this mandarin clade in the indel information-excluded parsimony tree (Fig. 2).

Five possible hybrid-origin mandarin landraces—Choupigan, Bendiguangju, Kuigan, Ouegan, and Huapigan—were in a clade with sweet orange ($C. sinensis$), trifoliate orange [$Poncirus trifoliata$ (L.) Raf.], pummelo, honghe papeda, and Eureka lemon [$C. limon$ (L.) Burm. f.] in the indel information-excluded parsimony consensus tree (no >50% bootstrap value supported) and maximum likelihood tree, indicating that these mandarin biotypes had a closer relationship with pummelo and sweet orange.

Wild mandarins had the same topology in phylogenetic trees constructed with all three methods. All the wild mandarins, with the exception of Daoxian 5 and Mangshan A1, were nested within the mandarin clade. Jiangyong 1, Daoxian 1, Chongyi A1, and Chongyi A2 wild mandarins were loosely associated with Jiangjinsuanju mandarin landrace (BP, 62% to 63%); Mangshan B0 wild mandarin and Jiangyong 4 wild mandarins were sisters to these four wild mandarins (BP, 71% to 91%). Daoxian 5 wild mandarin and Mangshan A1 wild mandarin were sisters to the mandarin clade I.

Rangpur lime and red lemon were contained in the mandarin clade I in all parsimony and likelihood analyses of the trnL-trnF sequence (BP, 71% to 91%), indicating a closer relationship of these biotypes with mandarins in the chloroplast genome.

Qingpi (yuzu?) and yuzu ($C. junos$ Siebold ex Tanaka) formed a weakly supported (BP, 65% to 64%) clade in the phylogenetic trees constructed by all three methods, showing a close relationship between them. Citron formed a monophyletic clade in all parsimony and likelihood trees (BP, 68% to 75%), and 2 citron and Tibet citron formed a subclade (BP, 53% to 59%). Results of parsimony analysis with indel information included showed a sister relationship of australian round lime to the subclade of 2 citron/Tibet citron, but this was weakly supported (Fig. 2, BP, 53%).

In the indel information-excluded parsimony tree (Fig. 1), relationships of meiwa kumquat ($Fortunella crassifolia$ Swing.), wild hongkong kumquat [$F. hindsii$ (Champ. ex Benth.) Swing.], and calamondin ($C. madurensis$ Lour.) were unresolved and could not form a clade. However, in the
parsimony tree with indels treated as a fifth character, meiwa kumquat and calamondin branched together with a bootstrap support of 89% (Fig. 2); and, in the maximum likelihood tree (Fig. 3), wild hongkong kumquat, meiwa kumquat, and calamondin formed an unresolved polytomy clade.

**Phylogenetic analysis using data set Flint2.** We used two data sets for maximum parsimony analysis based on Flint2 sequences. First, all indels were excluded. A heuristic search found 12 equally most-parsimonious trees of 104 steps, a CI of 0.904, and an RI of 0.940. The 50% major-rule consensus tree and bootstrap values are shown in Fig. 4, with a tree length of 111 steps, a CI of 0.847, and an RI of 0.898.

Second, we added all indels and treated them as the fifth character. Phylogenetic analysis yielded 348 most-parsimonious trees of 177 steps with a CI of 0.797 and an RI of 0.870. A 50% major-rule consensus tree (CI, 0.750; RI, 0.830; tree length, 188 steps) is shown in Fig. 5, and a few differences in topology and decreased branch resolution were found between this tree and that obtained by the first analyses.

The best evolutionary model for the Flint2 data determined in Modeltest was HKY+G with the shape parameter of gamma distribution, $\alpha = 0.5279$, and a transition-to-transversion ratio of 1.3502. The best tree ($-\ln L = 1935.55849$) in the maximum likelihood analysis (Fig. 6) using the model showed the same topology as the strict consensus maximum parsimony tree derived from the indel information-excluded data matrix.

The KH test and the SH test found the maximum likelihood tree to be the best tree, and 12 indel information-excluded most-parsimonious trees and several indel information-included most-parsimonious trees were not different significantly from the maximum likelihood tree.

The mandarin landraces clustered as a large group (BP, 56%), with two possible mandarin related hybrids (Kuigan and Choupigan) falling outside this clade (hereafter called the mandarin clade II). Both parsimony and likelihood analysis revealed that Choupigan mandarin landrace and Shatiyantu pummelo together formed a strongly supported monophyletic group (BP, 83% to 91%), and that the Kuigan mandarin landrace is a sister to the Choupigan/Shatiyantu pummelo clade supported by a low bootstrap value (BP, 57% or <50% BP). Tuju and Manju mandarin landraces formed a clade within this mandarin clade, with 61% to 67% BP support, and likelihood and indel-included parsimony analyses place Chongyi A1 wild mandarin as sister to this clade (Tuju/Manju), but no greater than 50% BP support was found in the parsimony tree.

All the wild mandarin accessions, except Mangshan A1 wild mandarin and Daoxian 5 wild mandarin, nested within the mandarin clade II. And, in this clade, Jiangyong 1 wild mandarin, Daoxian 1 wild mandarin, Jiangyong 4 wild mandarin, and Chongyi A2 wild mandarin branched together (BP, 52% to 66%), with Jiangyong 1 being sister to other three biotypes (BP, 50% to 62%). Mangshan A1 wild mandarin and Daoxian 5 wild mandarin were sisters to the mandarin clade II.

In this study, clustering of australian round lime (M. australis) with Mangshan A1 wild mandarin was the most distinct in the Flint2 data set analysis. In all phylogenetic trees
resulting from the FLint2 analysis, Mangshan A1 wild mandarin and australian round lime formed a highly supported clade (BP, 100%). They are significantly different in morphology and geographical distribution, so this result is difficult to explain.

Results of indel-excluded parsimony and maximum likelihood analysis showed a sister-group relationship of ichang papeda (*C. ichangensis* Swing.) and dazhongcheng papeda (*C. macrosperma*) to the Qingpi/yuzu clade, but this is not well supported by parsimony analysis (BP, 62%).

Rangpur lime and valencia sweet orange (*C. sinensis*) rested among biotypes belonging to mandarin in parsimony and likelihood trees, and therefore were more related to mandarin than to other species.

Eureka lemon and all three biotypes of citron formed a strongly supported polytomy (BP, 96% to 98%), and red lemon is sister to this clade supported by a relatively high bootstrap value (BP, 73% to 76%).

As in the *trnL-trnF* data set, Qingpi and yuzu clustered together in all the maximum parsimony and maximum likelihood analyses of the FLint2 data set (BP, 63% to 81%), again showing a close relationship between these two accessions.

Honghe papeda, meiwa kumquat, and calamondin produced a well-supported polytomy clade in both the parsimony (BP, 84% to 86%) and likelihood trees, and only in the indel information-included parsimony tree, meiwa kumquat and calamondin could form a weakly supported subclade (BP, 52%). The position of wild hongkong kumquat as sister to these three biotypes was only moderately supported by indel information-included parsimony analysis (BP, 74%).

Different from cpDNA analyses results, in which, the relationship between trifoliate orange and ‘Zhi-que’ (trifoliate orange?) was unresolved, in nrDNA analyses results, these two accessions formed a highly supported clade in all maximum parsimony and maximum likelihood trees (BP, 97% to 98%).

**Discussion**

**Relationships among wild mandarin and mandarin landraces.** Both the phylogenetic analysis of nrDNA and cpDNA sequence data showed that wild mandarins were not a monophyletic group, and were divided into three groups.

Sister relationships of Mangshan A1 wild mandarin and Daoxian 5 wild mandarin to all the nonhybrid mandarin landraces were supported by both the cpDNA and nrDNA analysis. Jiangyong 1, Daoxian 1, and Chongyi A2 wild mandarins formed a subclade within the mandarin clade I and II in cpDNA (BP, 62% to 63%) and nrDNA (BP, 52% to 66%); however, cpDNA analysis placed Chongyi A1 wild mandarin in the mandarin clade I whereas nrDNA analysis placed Jiangyong 4 wild mandarin in the mandarin clade II. The other accessions of wild mandarin, Jiangyong 4 and Mangshan B0, Jiangyong 1, and Mangshan B0 were sisters to this wild mandarin subclade in cpDNA and nrDNA respectively.

A relatively loose association of Jiangjinsuanju mandarin landrace and wild mandarins (Jiangyong 1, Daoxian 1, Chongyi A1, Chongyi A2) was revealed by cpDNA analyses (BP, 62% to 76%).
Chongyi A1 wild mandarin appeared to be sister to the Manju/Tuju mandarin landrace clade in nrDNA analyses results (Figs. 5 and 6), but this was not well supported (BP, 67%).

**Relationships of Citrus and Its Relatives.** The most commonly invoked cause for incongruence among nuclear and plastid DNA markers is hybridization. Because of the widely accepted belief that most *Citrus* species have been derived from hybridization (Federici et al., 1998; Handa et al., 1986; Moore, 2001), we attempted to use the incongruence between a cpDNA data set and an nrDNA data set to decide which species are hybrids.

Red lemon was believed to be a hybrid between citron and mandarin (Fang et al., 1993). In this study, red lemon was contained within a clade including lemon (*C. limon*) and citron (*C. medica*) in the analysis of the nrDNA data set, whereas it nested among mandarins in the analysis of the cpDNA data set. From this point of view, we propose that red lemon is a hybrid between mandarin and lemon or citron, with the former as the maternal parent and the latter as the male parent—a view consistent with previous suggestions that citron (*C. medica*) is probably a parent of limes and lemons (Barrett and Rhodes, 1976; Federici et al., 1998; Malik et al., 1974; Scona, 1975).

A citron × mandarin hybrid ancestry was proposed for rangpur lime in a previous study (Nicolosi et al., 2000). However, we could not solve the relationship of rangpur lime to mandarin and citron in this analysis. It has a close association with mandarin in nrDNA data set analysis, but uncertainty in the cpDNA data set analysis. In indel information-excluded parsimony tree and maximum likelihood tree, it clustered with all nonhybrid mandarin landraces, but in indel information-included parsimony tree, it branched together with the Tuju mandarin landrace (BP, 71%) as a part of a large clade including lemon, papeda, and pummelo (BP, 56%).

Swingle (1943) suggested calamondin to be a hybrid between kumquat and mandarin, whereas some studies did not support kumquat as its parent (Torres et al., 1978), suggesting that *Fortunella* species were a single, independent group (Rahman and Nito, 1994), with a single maternal ancestor (Abkenar et al., 2004). Our results showed that, in both nrDNA analysis and cpDNA analysis, meiwa kumquat and calamondin formed a monophyletic clade, and a sister relationship of honghe papeda was moderately supported by nrDNA analysis. Another kumquat species in this study, wild hongkong kumquat, was sister to these two kumquat species, but this was weakly supported by both the nrDNA (Fig. 5; BP, 74%) and cpDNA (Fig. 3). Despite the fact that the relationship of calamondin and kumquat was not resolved in this study, the relationship of calamondin with mandarin was not supported by our results, although sample errors (e.g., undetected sequences in cloning procedures) cannot be ruled out as possibilities.

Both the *trnL-trnF* data set and the *Flint2* data set support Qingpi as being the same species as yuzu, and it was different significantly from mandarin landraces both in this study and in a previous study based on nuclear and chloroplast SSR data (Li et al., 2006). Besides, in *Flint2* analysis, these two biotypes formed a clade with ichang papeda and dazhongcheng papeda.

![Fig. 6. Maximum likelihood reconstruction (-ln L = 1935.5849) of phylogeny under the HKY+G model for *Flint2* sequences in mandarins and related *Citrus* species. The best-fitting model was selected using Modeltest 3.7 (Posada and Crandall, 1998).](image)
Previously, Swingle and Reece (1967) suggested that yuzu is a hybrid between ichang papeda and mandarin. Our results support the nuclear genome of yuzu coming from papeda, but the origin of their chloroplast genomes was not resolved in this study.

Another two biotypes found to be the same species were trifoliate orange and Zhiquie, despite divergence in their cpDNA sequences. These results were consistent with results from nuclear simple sequence and chloroplast simple sequence analysis (Y.Z. Li and X.X. Deng, unpublished).

Huapigan, Choupigan, Kuigan, Benidgesuqian, and Ougan mandarin landraces have been proposed to be of hybrid origin, with pummelo/orange and mandarin being parents (Cheng et al., 2005; Li et al., 2006). This study further supported this hypothesis, because Huapigan, Benidgesuqian, and Ougan clustered with mandarin accessions in the analysis of nrDNA, whereas distributed in the pummelo/orange clade in the analysis of cpDNA indicating that mandarin was their male parent and pummelo/orange was their female parent. However, FLev2 analysis placed Kuigan, Choupigan, and pummelo in a clade. We proposed that only the maternal copy of the gene is present in these biotypes, and sampling of other low-copy genes may provide further information.

In this study, valencia sweet orange had a close association with mandarin in nrDNA, whereas it had a close association with pummelo in cpDNA, supporting the theory that sweet oranges are a hybrid between mandarin and pummelo (Barrett and Rhodes, 1976; Luro et al., 1995; Nicoliost et al., 2000), with mandarin as the male parent and pummelo/orange as the female parent.

### Literature Cited


