Tree Genetics & Genomes (2018) 14:6
https://doi.org/10.1007/s11295-017-1223-z

ORIGINAL ARTICLE

Phylogenetic relationships of Aurantioideae (Rutaceae) based on RAD-Seq

Yukio Nagano[1] · Takashi Mimura[2] · Nobuhiro Kotoda[2] · Ryoji Matsumoto[2] · Atsushi J. Nagano[3,4,5] · Mie N. Honjo[3] · Hiroshi Kudoh[3] · Masashi Yamamoto[6]

Received: 31 July 2017 / Revised: 27 November 2017 / Accepted: 25 December 2017

Abstract

The economically and nutritionally important genus *Citrus* belongs to the subfamily Aurantioideae in the family Rutaceae. Here, we analyzed the phylogenetic relationships of the subfamily Aurantioideae based on RAD-Seq. The RAD-Seq data produced phylogenetic trees with high support values, clear discriminations based on branch length, and elucidations of early branching events. Our genetic classification corresponded well with the classical morphological classification system and supported the subdivision of Citreae, one of two tribes of the Aurantioideae, into three subtribes—Triphasiinae, Citrinae, and Balsamocitrinae. Additionally, it was largely consistent with the subdivision of Clauseneae, the other tribe of the Aurantioideae, into three subtribes—Micromelinae, Clauseninae, and Merrillinae; the exception was *Murraya paniculata*. With the exception of members of primitive citrus fruit trees, namely, *Severinia buxifolia* and *Hesperethusa crenulata*, lower-level morphological groupings under subtribes based on genetic and morphological classifications corresponded well. The phylogenetic relationship between Asian "true citrus fruit trees" (genera *Citrus*, *Poncirus*, and *Fortunella*) and Australian/New Guinean citrus fruit trees (genera *Microcitrus*, *Eremocitrus*, and *Clymenia*) was inconsistent between present classification based mainly on the nuclear genome and the previous classification based on the chloroplast genome. This inconsistency may be explained by chloroplast capture. Our findings provide a valuable insight into the genetic relationships of the subfamily Aurantioideae in the family Rutaceae.

Keywords Citrus · Aurantioideae · Rutaceae · RAD-Seq · Phylogeny · Genetic classification

Communicated by W.-W. Guo

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11295-017-1223-z) contains supplementary material, which is available to authorized users.

✉ Yukio Nagano
nagano@cc.saga-u.ac.jp

[1] Analytical Research Center for Experimental Sciences, Saga University, 1 Honjo-machi, Saga 840-8502, Japan

[2] Faculty of Agriculture, Saga University, 1 Honjo-machi, Saga 840-8502, Japan

[3] Center for Ecological Research, Kyoto University, 509-3 2-chome, Hirano, Otsu, Shiga 520-2113, Japan

[4] JST PRESTO, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

[5] Faculty of Agriculture, Ryukoku University, 1-5 Yokotani, Seta Oe-cho, Otsu, Shiga 520-2194, Japan

[6] Faculty of Agriculture, Kagoshima University, 1-21-24 Korimoto, Kagoshima 890-0065, Japan

Introduction

Citrus is an economically and nutritionally important genus and a member of the subfamily Aurantioideae in the family Rutaceae. In addition to *Citrus* species, some other members of Aurantioideae are used for various purposes such as foods, spices, medicinal chemicals, cosmetics, and garden trees. Therefore, it is important to understand the phylogenetic relationships of the subfamily Aurantioideae.

Classical morphology divided the subfamily Aurantioideae into two tribes—Clauseneae and Citreae (Swingle and Reece 1967). Recent molecular classifications, using small numbers of DNA sequences, also clearly discriminated Citreae from Clauseneae (Bayer et al. 2009; Penjor et al. 2013; Schwartz et al. 2015). Morphological classification subdivided each tribe into three subtribes. The tribe Clauseneae consists of Micromelinae, Clauseninae, and Merrillinae, and the tribe

Published online: 02 January 2018

Citreae consists of Triphasiinae, Citrinae, and Balsamocitrinae (Swingle and Reece 1967). However, recent molecular studies did not clearly reveal this level of classification (Bayer et al. 2009; Penjor et al. 2013; Schwartz et al. 2015), probably because they did not have the capability to determine earlier branching events.

The data size used in previous molecular analyses was small, i.e., small numbers of markers or short DNA sequences. Therefore, these previous studies did not discriminate the earlier branching events that separated major taxa. In addition, it was difficult to reveal genetic relationships among closely related species. Next-generation sequencing produces a large amount of data; hence, it is worthwhile using it to reveal the genetic relationships within the subfamily Aurantioideae. In comparison with whole genome sequencing, restriction site-associated DNA sequencing (RAD-Seq) can analyze many samples inexpensively and simultaneously, and it is suitable for studying the relationships between genetically similar individuals (Baird et al. 2008). The original report of RAD-Seq (Baird et al. 2008) has more than 1000 Google Scholar citations, and the technique has been applied to many types of model and non-model organisms (Andrews et al. 2016). In our case, we used this method to analyze *Citrus* species (Penjor et al. 2014, 2016), which are genetically similar because of their cross-compatibilities. The relationships between genetically distant individuals at the genus and/or family level have been analyzed using RAD-Seq data with the new software PyRAD (Eaton 2014). The clustering method of PyRAD does not use the reference genome and allows for lower similarity thresholds and inclusion of indels. Therefore, it is possible to analyze relationships within the subfamily Aurantioideae using RAD-Seq. The RAD-Seq method and the programs required to analyze the resulting data were reviewed by Andrews et al. (2016).

In the present study, we analyzed the genetic relationships within the subfamily Aurantioideae based on RAD-Seq. The obtained phylogenetic trees elucidated branching order well. Herein, we describe new findings and confirm previous findings. Our molecular classification is consistent with the classical morphological classification (Swingle and Reece 1967).

Materials and methods

Plant materials

Thirty-five species from 22 genera of the subfamily Aurantioideae were used in this study (Table 1). The materials have been preserved at the Faculty of Agriculture, Saga University, and the Faculty of Agriculture, Kagoshima University.

RAD-Seq analysis

The DNA purification procedure was identical to that described by Penjor et al. (2014). The method used to create the library for double-digest RAD-Seq was identical to that described by Sakaguchi et al. (2015), which is a modification of the original double-digest RAD-Seq (Peterson et al. 2012) as follows: *Bgl*II was used as the first restriction site adjacent to the binding site of the primer to read a single-end sequence, and *Eco*RI was used as the second restriction site adjacent to the binding site to read an index sequence. The library was sequenced with 49 bp single-end reads in one lane of an Illumina HiSeq2000 (Illumina, San Diego, CA, USA) by BGI Hong Kong. At BGI, the raw data were modified using the following two steps: (1) reads that were polluted by adapter sequences were deleted, and (2) reads that contained > 50% low-quality bases (quality value ≤ 5) or > 10% N bases were removed.

Phylogenetic analysis based on the pyRAD program

Reads were further cleaned using the process_shortreads program of the Stacks package (version 1.46) with -c (clean data, remove any read with an uncalled base) and -q (discard reads with low-quality scores) options (Catchen et al. 2011, 2013). For the construction of phylogenetic trees, multiple alignments were created using the pyRAD program (version 3.0.66) (Eaton 2014). In multiple alignments c80m12, c85m12, and c90m12, Wclust (clustering threshold as a decimal) was set to 0.80, 0.85, and 0.90, respectively, with Mindepth (min coverage for a cluster) = 6, NQual (max # sites with quality < 20) = 4, MinCov (min samples in a final locus) = 12, and MaxSH (max inds with shared hetero site) = 3. In multiple alignments c80m4, c85m4, and c90m4, Wclust was set to 0.80, 0.85, and 0.90, respectively, with Mindepth = 6, NQual = 4, MinCov = 4, and MaxSH = 3. In these calculations, the default parameter of maximum depth filtering was used. For all six multiple alignments, phylogenetic trees based on maximum likelihood were constructed using the RAxML program (version 8.2.10) (Stamatakis 2014) (-f = a, -x = 12,345, -*p* = 12,345, -N (bootstrap value) = 1000, and -m = GTRGAMMA). For multiple alignments c80m12, c85m12, and c90m12, phylogenetic trees based on Bayesian inference were constructed using the MrBayes program (version 3.2.2) (Ronquist and Huelsenbeck 2003) (lset nst = 6, rates = invgamma, mcmc ngen = 100,000, samplefreq = 1000, nchains = 4, and savebrlens = yes). In each analysis, the midpoint was used as a root. The number of phylogenetically informative sites was calculated using MEGA version 7 (Kumar et al. 2016).

Phylogenetic analysis based on the Stacks program

For the de novo analysis, the cleaned reads were analyzed using the denovo_map.pl script of the Stacks package

Table 1 Species used to analyze the genetic relationships within Aurantioideae

Tribe	Subtribe	Group	Latin name (common name/accession name)	Source	Accession no.
Clauseneae	Micromelinae		*Micromelum minutum* (Forst.) Wt. & Arn.	Saga Univ.	8650
	Clauseninae		*Clausena anisata* (Willd.) Hook. f.	Saga Univ.	8612
			Clausena harmandiana (Pierre) Guill.	Saga Univ.	8613
			Clausena lansium (Lour.) Skeels (wanpee)	Saga Univ.	8611
			Glycosmis citrifolia (Willd.) Lindl.	Saga Univ.	8601
			Glycosmis pentaphylla (Retz.) Correa (orangeberry)	Saga Univ.	8600
			Murraya koenigii (L.). Spreng. (curry tree)	Saga Univ.	8622
			Murraya paniculata (L.) Jack.	Saga Univ.	8621
	Merrillinae		*Merrillia caloxylon* (Ridl.) Swing.	Saga Univ.	8640
Citreae	Triphasiinae	Triphasia	*Paramignya lobata* Burkill	Saga Univ.	8350
			Triphasia trifolia (Burm. f.) P. Wils. f.	Saga Univ.	8500
	Balsamocitrinae	Tabog	*Swinglea glutinosa* (Blanco) Merr	Saga Univ.	8420
		Bael fruit	*Aegle marmelos* (L.) Corr. (bael)	Saga Univ.	8400
			Afraegle paniculata (Schum.) Engl.	Saga Univ.	8411
		Wood apple	*Feronia limonia* (L.) Swing.	Saga Univ.	8450
			Feroniella oblata Swing.	Saga Univ.	8460
	Citrinae	Primitive citrus fruit trees	*Hesperethusa crenulata* (Roxb.) Roem.	Saga Univ.	8320
			Severinia buxifolia (Poir.) Tenore (Chinese box-orange)	Saga Univ.	8340
		Near citrus fruit trees	*Atalantia bilocularis* (Roxb.) Wall. ex Skeels	Saga Univ.	8312
			Atalantia ceylanica (Arn.) Oliv.	Saga Univ.	8310
			Atalantia monophylla DC.	Saga Univ.	8314
			Atalantia roxburghiana Hook. f.	Saga Univ.	8316
			Atalantia spinosa (Willd.) Tanaka	Saga Univ.	8315
			Citropsis gabunensis (Engl.) Swing. & M. Kell.	Saga Univ.	8300
			Citropsis gilletiana Swing. & M. Kell.	Saga Univ.	8302
			Citropsis schweinfurthii (Engl.) Swing. & M. Kell.	Saga Univ.	8301
		True citrus fruit trees	*Clymenia polyandra* (Tan.) Swing.	Saga Univ.	8280
			Eremocitrus glauca (Lindl.) Swing.	Saga Univ.	8251
			Microcitrus australasica (F. Muell.) Swing.	Saga Univ.	8203
			Fortunella japonica (Thunb.) Swing. (round kumquat)	Saga Univ.	8002
			Poncirus trifoliata (L.) Raf. "Standard" (trifoliate orange)	Saga Univ.	8100
			Citrus micrantha Wester (papeda/biasong)	Saga Univ.	7004
			Citrus medica L. (citron/Maru busshukan)	Saga Univ.	5001
			Citrus maxima (Burm.) Merr. (pummelo/Mato Buntan)	Saga Univ.	3202
			Citrus reticulata Blanco (mandarin/Yoshida Ponkan)	Kagoshima Univ.	–

(version 1.46) with the m (minimum number of identical raw reads required to create a stack) option of 3. For the reference-based analysis, the cleaned reads were aligned with the reference genomes using bowtie2 (Langmead and Salzberg 2012) with -q –no-unal -L 15 options, and aligned data were analyzed using the ref_map.pl script of the Stacks package with default parameters. As reference genomes, *Citrus sinensis* (sweet orange) (Xu et al. 2012) was used for the analysis of nine members of "true citrus fruit" trees, and *Severinia buxifolia* (*Atalantia buxifolia*) (Wang et al. 2017) was used for the analysis of six members of *Severinia* and *Atalantia*. The populations program of the Stacks package was used to create multiple alignments within the cluster using the options –phylip and –phylip_var. In this program, the -p (minimum number of populations a locus must be present in to process a locus) option was set to 5 in the analysis of nine members of "true citrus fruit trees" and to 3 in the analysis of six members of *Severinia* and *Atalantia*. The phylogenetic trees were constructed as described above.

Results

Phylogenetic trees of the subfamily Aurantioideae

We used double-digest RAD-Seq of 35 species to reveal the relationships within the subfamily Aurantioideae (Table 1). The library for double-digest RAD-Seq was created using the restriction enzymes *Bgl*II and *Eco*RI. The region adjacent to *Bgl*II was sequenced with 49 bp single-end reads. Using the quality-filtered reads shown in Supplementary Table 1, the PyRAD program created six multiple alignments by changing the clustering thresholds (80, 85, or 90%) and number of samples (species) in a final locus (four or 12).

Based on these variables, the multiple alignments were named c80m4, c85m4, c90m4, c80m12, c85m12, and c90m12, and they contained 652,582; 641,156; 599,272; 189,271; 184,680; and 159,021 aligned sequences, respectively. The numbers of phylogenetically informative sites in multiple alignments c80m4, c85m4, c90m4, c80m12, c85m12, and c90m12 were 32,999; 30,041; 19,717; 16,484; 15,258; and 9979; respectively. Multiple alignments c80m4, c85m4, and c90m4 extracted the locus conserved among only four species; hence, they produced larger numbers of aligned sequences than did c80m12, c85m12, and c90m12, which extracted the locus conserved among 12 species. Analysis of the nucleotide compositions (Supplementary Tables 2–7) revealed that each alignment contained fewer heterozygous sites and deletions than homologous sites.

Multiple alignments c80m12, c85m12, and c90m12 were used to construct phylogenetic trees based on maximum likelihood (Supplementary Fig. 1, Supplementary Fig. 2, and Fig. 1, respectively) and Bayesian inference (Supplementary Fig. 3, Supplementary Fig. 4, and Fig. 2, respectively). Owing to poor computer performance, c80m4, c85m4, and c90m4 were used to construct phylogenetic trees based on maximum likelihood (Supplementary Fig. 5, Supplementary Fig. 6, and Supplementary Fig. 7, respectively) but not on Bayesian inference. It was difficult to include the RAD-Seq data of the outgroup (distantly related species) in these phylogenetic analyses; hence, the midpoint was used as a root.

The obtained phylogenetic trees had similar topologies and generally shared several characteristics. (1) The subfamily Aurantioideae was subdivided into two tribes—Clauseneae and Citreae. (2) The tribe Citreae was subdivided into three subtribes—Triphasiinae, Citrinae, and Balsamocitrinae. (3) The tribe Clauseneae was subdivided into three subtribes—Micromelinae, Clauseninae, and Merrillinae—although *Murraya paniculata* was an exception. (4) The subtribe Balsamocitrinae was subdivided into three groups—wood apple, tabog, and bael fruit. (5) The subtribe Citrinae was subdivided into three groups—a group containing the genera *Hesperethusa* and *Citropsis*, a group containing the genera *Severinia* and *Atalantia*, and the "true citrus fruit trees"—and the proposal to discriminate the two groups "near citrus fruit trees" and "primitive citrus fruit trees" was not supported. (7) "True citrus fruit trees" formed a monophyletic group.

Comparison of the phylogenetic trees based on maximum likelihood and Bayesian inference revealed no difference in tree topology for each of the multiple alignments c80m12, c85m12, and c90m12. However, comparison of the trees constructed using the six types of alignments (c80m4, c85m4, and c90m4; c80m12, c85m12, and c90m12) revealed some differences. (1) In all trees except the maximum-likelihood tree based on c80m4 (Supplementary Fig. 5), Micromelinae was an outgroup of a monophyletic clade containing the members of Clauseninae; however, in the maximum-likelihood tree based on c80m4 (Supplementary Fig. 5), *Murraya koenigii* (*Bergera koenigii*, commonly known as the curry tree) was clustered together with Micromelinae (*Micromelum minutum*). (2) In all trees except the maximum-likelihood trees based on c90m4 (Supplementary Fig. 7), *Murraya koenigii* was an outgroup of five species (two *Glycosmis* species and three *Clausena* species); however, in the maximum-likelihood tree based on c90m4 (Supplementary Fig. 7), two *Glycosmis* species formed an outgroup of three *Clausena* species and *Murraya koenigii*, and *Murraya koenigii* was an outgroup of three *Clausena* species. (3) In the maximum-likelihood trees based on c90m12 and c90m4 (Fig. 1 and Supplementary Fig. 7) and the Bayesian inference tree based on c90m12 (Fig. 2), Triphasiinae was an outgroup of Balsamocitrinae and Citrinae; however, in the maximum-likelihood trees based on c80m12, c85m12, c80m4, and c85m4 (Supplementary Figs. 1, 2, 5, and 6) and the Bayesian inference trees based on c80m12 and c85m12 (Supplementary Figs. 3 and 4), Balsamocitrinae was an outgroup of Triphasiinae and Citrinae. (4) In all trees except the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), the branch containing wood apple, tabog, and Citrinae was subdivided into two sub-branches—one was Citrinae, and the other consisted of wood apple and tabog; in comparison, in the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), wood apple was an outgroup of the branch containing tabog and Citrinae, and tabog was an outgroup of Citrinae. (5) There was a conflict in the locations of *Fortunella japonica* (round kumquat) and *Citrus reticulata* (mandarin); in all trees except the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), these two species formed a single branch, and this branch was clustered together with another branch containing Australian/New Guinean "true citrus fruit trees" (*Clymenia*, *Eremocitrus*, and *Microcitrus*); in comparison, in the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), *Citrus reticulata* was an outgroup of the branch containing *Fortunella japonica* and Australian/New Guinean "true citrus fruit trees," and

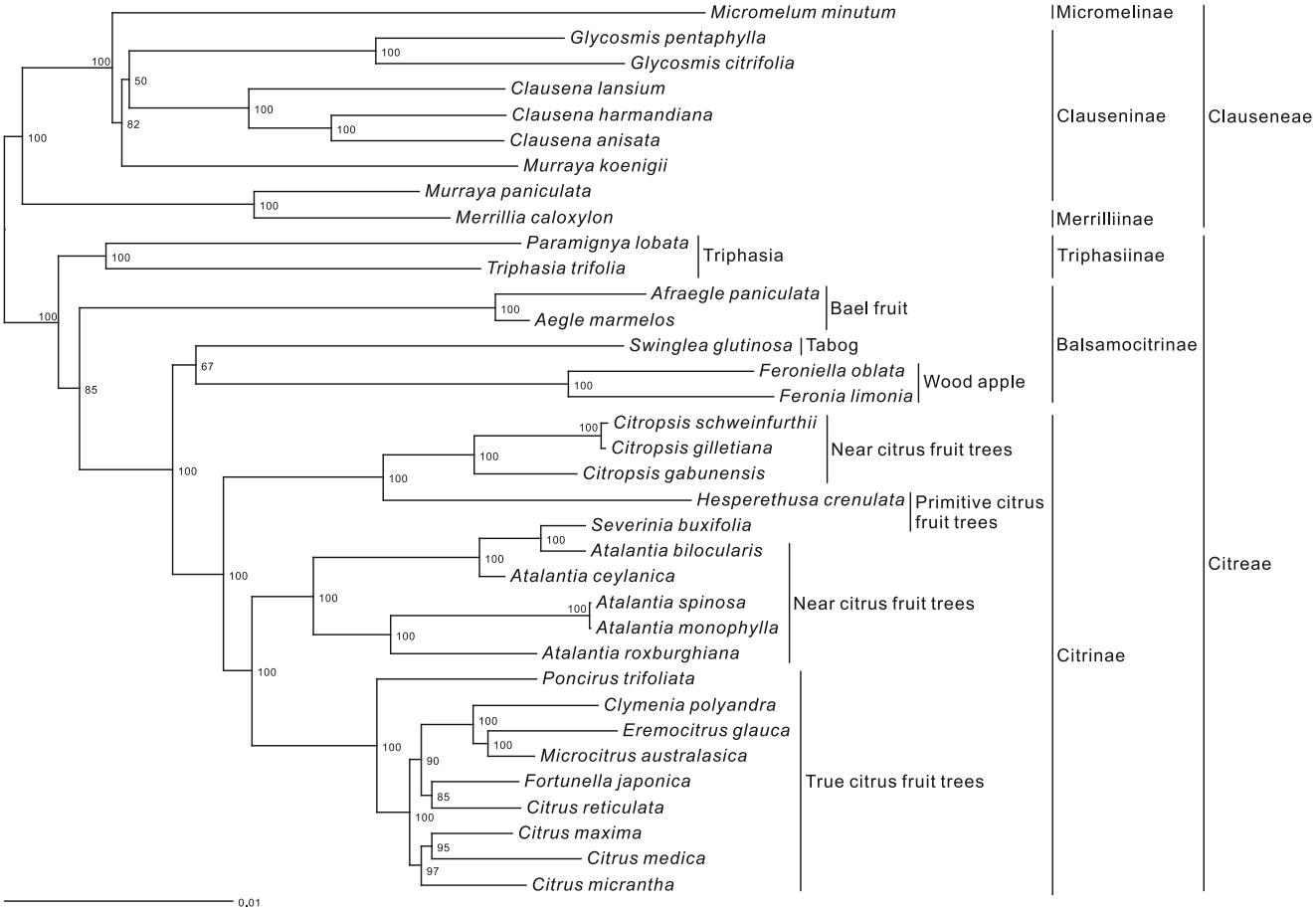

Fig. 1 Phylogenetic tree based on maximum likelihood analysis of 35 species of the subfamily Aurantioideae. Numbers at the nodes indicate bootstrap values (% over 1000 replicates). The scale bar shows the number of substitutions per site. The phylogenetic tree was calculated based on multiple alignment c90m12. The multiple alignment was calculated using the PyRAD program, with Wclust (clustering threshold as a decimal) set to 0.90 and MinCov (min samples in a final locus) set to 12

Fortunella japonica was an outgroup of the branch containing Australian/New Guinean "true citrus fruit trees."

Phylogenetic trees of closely related species

The pyRAD program is useful for creating multiple alignments of distantly related samples (Eaton 2014), and the Hardy-Weinberg principle is not expected to apply. However, in the analysis of closely related samples such as cross-compatible species, the Hardy-Weinberg principle should be considered. The Stacks program is useful for analyzing the RAD-Seq data of closely related samples (Catchen et al. 2011, 2013), and the Hardy-Weinberg principle is expected to apply. The members of "true citrus fruit trees" (*Citrus*, *Fortunella*, *Poncirus*, *Eremocitrus*, *Microcitrus*, and *Clymenia*) are closely related, and cross-compatibility was observed (Iwamasa et al. 1988). Therefore, we analyzed the RAD-Seq data of "true citrus fruit trees" using the Stacks program. *Citrus medica* (citron), *Citrus micrantha* (papeda), *Citrus maxima* (pummelo), and *Citrus reticulata* (mandarin) are considered to be ancestral species, and most other *Citrus* species are considered to be derivatives or hybrids of these four species (Nicolosi et al. 2000; Froelicher et al. 2011; Garcia-Lor et al. 2013; Curk et al. 2014, 2015). We used these four species in our subsequent analysis. One of the multiple alignments was created using a de novo method, and the other multiple alignment was created by aligning the data to the reference genome of *Citrus sinensis* (sweet orange) (Xu et al. 2012). The former and latter calculations extracted 10,783 and 11,143 sites, respectively, and these included 1003 and 1375 phylogenetically informative sites, respectively. The nucleotide compositions of these alignments are shown in Supplementary Tables 8 and 9. We constructed phylogenetic trees based on maximum likelihood (Fig. 3) and Bayesian inference (Supplementary Fig. 8). Comparison of the phylogenetic trees based on maximum likelihood and Bayesian inference revealed no difference in the tree topology for each multiple alignment. However, comparison of the trees constructed using the two types of multiple alignment revealed some differences in the locations of *Fortunella japonica* (round kumquat) and *Citrus reticulata*. In the trees based on the de novo method, *Fortunella japonica* was an outgroup of Australian/New Guinean "true citrus fruit trees," and *Citrus*

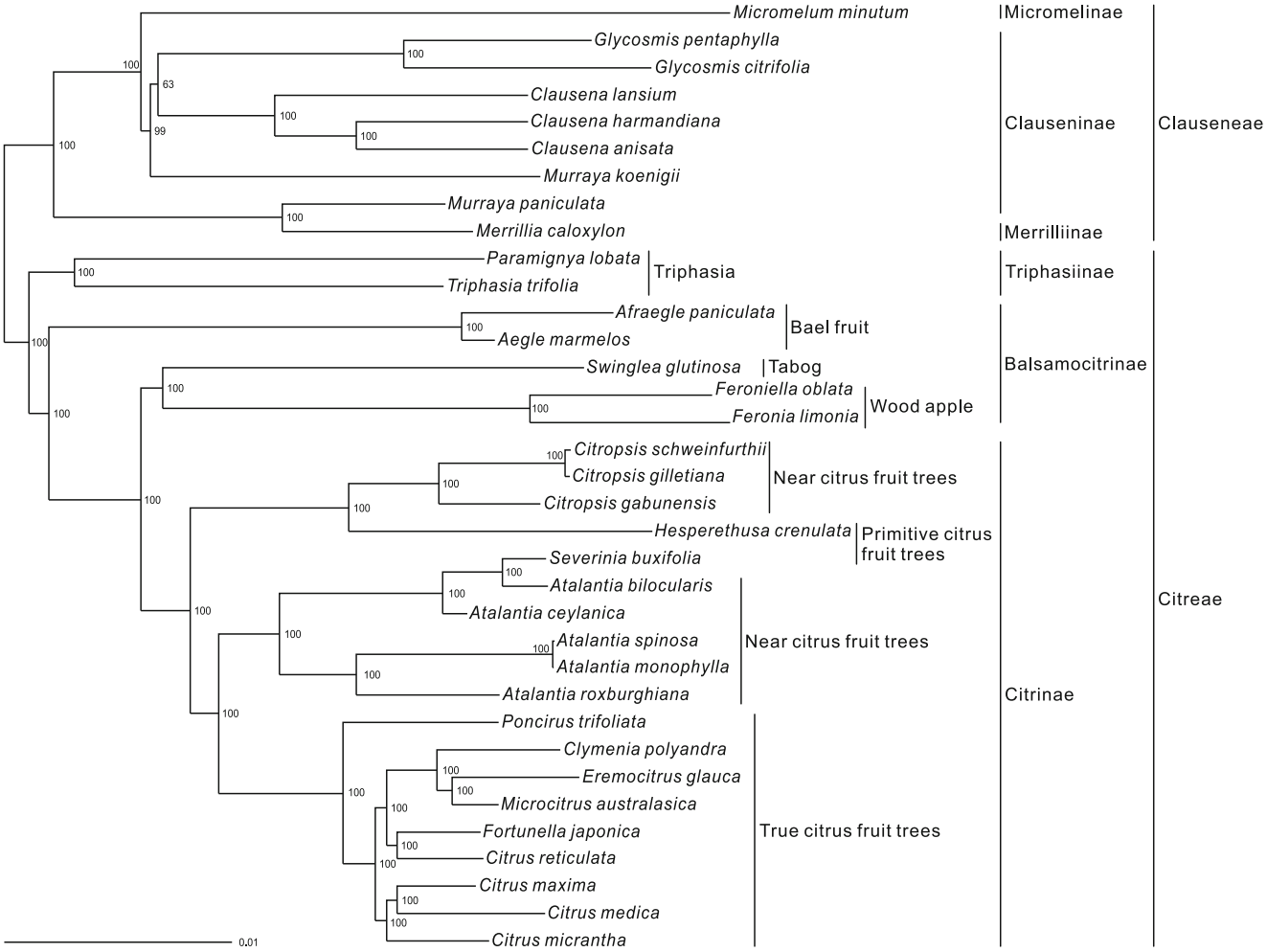

Fig. 2 Phylogenetic tree based on Bayesian inference analysis of 35 species of the subfamily Aurantioideae. Numbers at the nodes indicate posterior probabilities. The scale bar shows the number of substitutions per site. The phylogenetic tree was calculated based on multiple alignment c90m12

reticulata was an outgroup of the remaining three *Citrus* species (Fig. 3a and Supplementary Fig. 8A). However, in the trees based on reference genome-based method, *Fortunella japonica* and *Citrus reticulata* formed a single branch, and this branch was clustered together with another branch containing Australian/New Guinean "true citrus fruit trees"; this finding was consistent with most of the pyRAD-based phylogenetic trees, except the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4).

The pyRAD-based trees showed that *Severinia buxifolia* (*Atalantia buxifolia*) and five species of *Atalantia* (*Atalantia bilocularis*, *Atalantia ceylanica*, *Atalantia monophylla*, *Atalantia roxburghiana*, and *Atalantia spinosa*) are closely related. Therefore, we analyzed the RAD-Seq data of these six species using the Stacks program, and we produced two types of multiple alignments using the de novo and reference genome-based methods. As a reference genome, we used the genome sequence of *Severinia buxifolia* (*Atalantia buxifolia*) (Wang et al. 2017). The former and latter calculations extracted 7958 and 6836 sites, respectively, and these included 1479 and 1810 phylogenetically informative sites, respectively. The nucleotide compositions of these alignments are shown in Supplementary Tables 10 and 11. We constructed phylogenetic trees based on maximum likelihood (Supplementary Fig. 9) and Bayesian inference (Supplementary Fig. 10). Comparison of the phylogenetic trees based on maximum likelihood and Bayesian inference revealed no difference in the topology of the trees for each multiple alignment. Furthermore, there was no conflict between the pyRAD-based trees and the Stacks-based trees.

Discussion

Relationships between the tribe Clauseneae and the tribe Citreae

The trees constructed in our present study clearly discriminate Citreae from Clauseneae, supporting previous molecular studies (Bayer et al. 2009; Penjor et al. 2013; Schwartz et al.

2015). These previous studies showed that members of Clauseneae formed an outgroup of Citreae, and the members of Clauseneae did not belong to the clade Citreae; the root was located between the clade containing *Murraya paniculata* and *Merrillia caloxylon* and the clade containing the other members of Clauseneae. The root of the trees constructed in our present study may be incorrect, because we did not include the outgroup in the RAD-Seq analysis; hence, we placed the root at the midpoint. In some previous studies, the polytomous clade containing *Murraya paniculata* and *Merrillia caloxylon* was included in the tribe Citreae (Samuel et al. 2001; Morton et al. 2003; Morton 2009; Oueslati et al. 2016); however, our present study and other previous studies (Bayer et al. 2009; Penjor et al. 2013; Schwartz et al. 2015) do not support these observations.

Relationships within the tribe Clauseneae

Classical morphology subdivided the tribe Clauseneae into three subtribes—Micromelinae, Clauseninae, and Merrillinae (Swingle and Reece 1967). In the present study, we analyzed one member of Micromelinae, (*Micromelum minutum*), one member of Merrillinae (*Merrillia caloxylon*; monotypic genus), and seven members of Clauseninae (two *Murraya* species, two *Glycosmis* species, and three *Clausena* species). Without taking into consideration the position of *Murraya paniculata*, in all trees except the maximum-likelihood tree based on c80m4 (Supplementary Fig. 5), our molecular classification subdivides the tribe Clauseneae into three subtribes—Merrillinae is an outgroup of Micromelinae and Clauseninae, and Micromelinae is an outgroup of monophyletic clade containing the members of Clauseninae; however, in the maximum-likelihood tree based on c80m4 (Supplementary Fig. 5), *Murraya koenigii* (*Bergera koenigii*) is clustered together with Micromelinae (*Micromelum minutum*), although the support value is low. In previous molecular studies (Bayer et al. 2009; Penjor et al. 2013; Schwartz et al. 2015), *Micromelum minutum* was nested in the subtribe Clauseninae. Thus, our present molecular classification is consistent with the morphological classification (Swingle and Reece 1967), in which the three subtribes were separated. However, in the present study, we analyzed only one species of Micromelinae; hence, further studies with additional species are required.

Next, we compared our present classification within the tribe Clauseneae with that of our previous molecular study (Penjor et al. 2013). Some findings are consistent as follows: (1) three species of the genus *Clausena*—*Clausena anisata*, *Clausena harmandiana*, and *Clausena lansium* (wanpee)—belong to a monophyletic clade; (2) two species of the genus *Glycosmis*—*Glycosmis citrifolia* and *Glycosmis pentaphylla* (orangeberry)—belong to a monophyletic clade; and (3) *Murraya paniculata* is clustered with *Merrillia caloxylon*

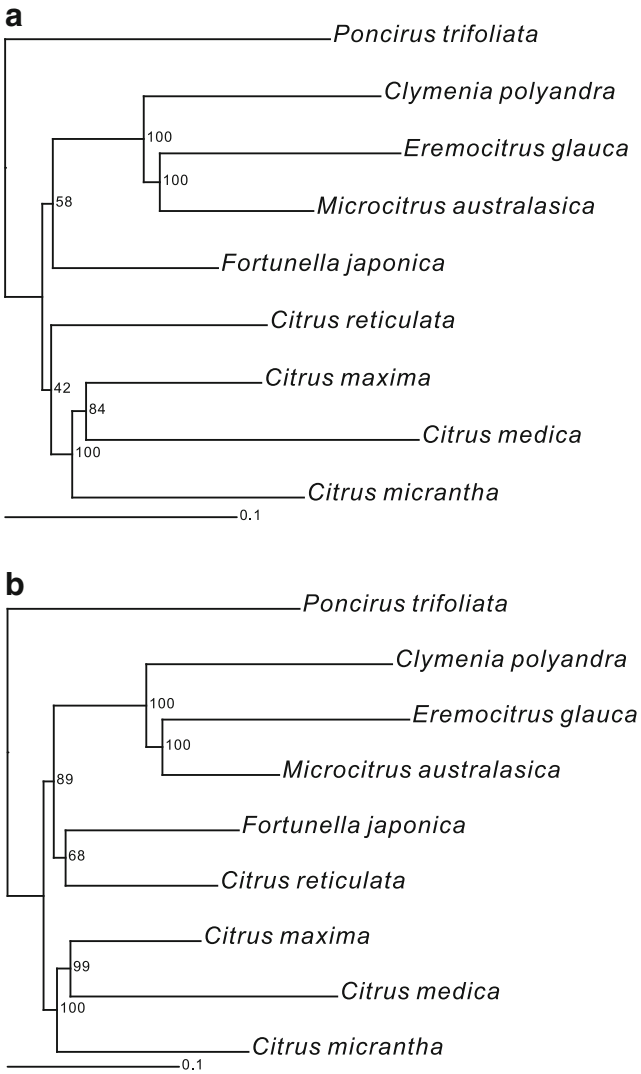

Fig. 3 Phylogenetic trees based on maximum likelihood analysis of nine species of "true citrus fruit trees." Numbers at the nodes indicate bootstrap values (% over 1000 replicates). The scale bar shows the number of substitutions per site. The phylogenetic trees were calculated based on multiple alignment created using de novo analysis (**a**) or reference genome-based analysis (**b**) of the Stacks package

but not with *Murraya koenigii*. Similar groupings were observed in other previous studies (Samuel et al. 2001; Bayer et al. 2009; Schwartz et al. 2015); however, the species used in these previous studies differed from those used in our present and previous studies (Penjor et al. 2013).

It is interesting to consider the phylogenetic positions of the genus *Murraya*. Genetic similarities between *Murraya paniculata* and *Merrillia caloxylon*, and genetic differences between *Murraya paniculata* and *Murraya koenigii* were described previously (Samuel et al. 2001; Bayer et al. 2009; Penjor et al. 2013; Oueslati et al. 2016). Morphological similarities between *Murraya paniculata* and *Merrillia caloxylon*, and morphological differences between *Murraya paniculata* and *Murraya koenigii*, were discussed in our previous study

(Penjor et al. 2013). Furthermore, previous phytochemical analysis revealed similarities between *Merrillia caloxylon* and some *Murraya* species, including *Murraya paniculata* and *Murraya exotica* (Samuel et al. 2001). In addition, genetic and karyotypic analyses showed that *Murraya koenigii* and *Murraya siamensis* were isolated from the other *Murraya* species (Guerra et al. 2000). Our present study confirms these observations, i.e., *Murraya paniculata* should be grouped with *Merrillia caloxylon*.

In our previous study (Penjor et al. 2013), the topology of the branch containing seven trees (three *Clausena* species, two *Glycosmis* species, *Micromelum minutum*, and *Murraya koenigii*) differed between neighbor-joining and maximum-likelihood trees. Furthermore, the topology of these four groups differed between several previous studies (Samuel et al. 2001; Bayer et al. 2009; Penjor et al. 2013). Our present analysis shows that in all phylogenetic trees except the maximum-likelihood trees based on c80m4 and c90m4 (Supplementary Figs. 5 and 7), *Micromelum minutum* is an outgroup of the remaining six species, and *Murraya koenigii* is the outgroup of the clade containing two *Glycosmis* species and three *Clausena* species, However, low support values are present in the clade containing these seven species. Therefore, the branching order within the clade should be interpreted with caution.

In comparison with our previous study (Penjor et al. 2013), our present analysis more clearly elucidates the genetic relationships within the tribe Clauseneae. In the clade containing three *Clausena* species, *Clausena lansium* is an outgroup of the sub-clade containing *Clausena anisata* and *Clausena harmandiana*; this was not clear in our previous study (Penjor et al. 2013).

Relationships within the tribe Citreae

Classical morphology subdivided the tribe Citreae into three subtribes—Triphasiinae (minor citroid fruit trees), Citrinae (citrus fruit trees), and Balsamocitrinae (hard-shelled citroid fruit trees) (Swingle and Reece 1967). Our present molecular study clearly discriminates these three subtribes as follows: (1) each of the subtribes Triphasiinae and Citrinae forms a monophyletic group, and (2) Balsamocitrinae does not form a monophyletic group and is an outgroup of Triphasiinae and Citrinae (Supplementary Figs. 1–6), or Triphasiinae is an outgroup of Balsamocitrinae and Citrinae (Figs. 1 and 2, and Supplementary Fig. 7). Thus, our present analysis corresponds well with the morphological classification (Swingle and Reece 1967). Previous molecular analysis (Bayer et al. 2009; Penjor et al. 2013) was unable to elucidate the relationships among the three subtribes. For example, members of Balsamocitrinae, *Feronia limonia*, *Feroniella oblata*, and *Swinglea glutinosa* were nested in the subtribe Citrinae (Bayer et al. 2009; Penjor et al. 2013). In a recent study (Schwartz et al. 2015), the monophyletic nature of Triphasiinae was clearly shown; however, a member of Balsamocitrinae (*Swinglea*) and a member of Citrinae (*Citropsis*) belonged to a single clade. Thus, our present analysis corresponds well with the morphological characterization.

Classical morphology subdivided the subtribe Balsamocitrinae into three groups—wood apple, tabog (monotypic genus), and bael fruit (Swingle and Reece 1967). Our present analysis clearly separates these three groups and corresponds well with the morphological observations (Swingle and Reece 1967). In all phylogenetic trees, bael fruit is an outgroup of a branch containing the subtribe Citrinae and the other subtribes of Balsamocitrinae (wood apple and tabog). In all phylogenetic trees, except the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), this branch is subdivided into two sub-branches—one is Citrinae, and the other consists of wood apple and tabog. In the maximum likelihood and Bayesian inference trees based on c85m12 (Supplementary Figs. 2 and 4), wood apple is an outgroup of the branch containing tabog and Citrinae, and tabog is an outgroup of Citrinae.

Consistent with previous studies (Bayer et al. 2009; Penjor et al. 2013), our present study shows that *Aegle marmelos* (Bael) and *Afraegle paniculata*—known as the Bael fruit group of the subtribe Balsamocitrinae—form a single cluster. It is important to note that *Aegle marmelos* is native to India, and *Afraegle paniculata* is native to tropical Africa (Swingle and Reece 1967). Our present study further shows that *Feronia limonia* and *Feroniella oblata*, known as the wood-apple group of the subtribe Balsamocitrinae (Swingle and Reece 1967), belong to the same cluster; this observation is consistent with the previous results (Morton et al. 2003; Morton 2009; Penjor et al. 2010, 2013). *Feroniella oblata* is not nested in "true citrus fruit trees," and this observation is inconsistent with the previous finding (Bayer et al. 2009).

Classical morphology subdivided the subtribe Citrinae into three groups—"true citrus fruit trees," near citrus fruit trees, and primitive citrus fruit trees (Swingle and Reece 1967). Members of the near citrus fruit trees were proposed to be nearer to *Citrus* species than to the two subtribes Balsamocitrinae and Triphasiinae; this observation is supported by our present analysis. Our present study shows that members of the primitive citrus fruit trees—*Severinia buxifolia* (*Atalantia buxifolia*, commonly known as Chinese box-orange) and *Hesperethusa crenulata*—are not clustered, and each is nested in near citrus fruit trees. This observation is inconsistent with the morphological classification (Swingle and Reece 1967).

Our present study shows that *Severinia buxifolia* (*Atalantia buxifolia*) and five species of *Atalantia* (*Atalantia bilocularis*, *Atalantia ceylanica*, *Atalantia monophylla*, *Atalantia*

roxburghiana, and *Atalantia spinosa*) form a monophyletic group; this finding is consistent with our previous observation (Penjor et al. 2013) and similar to those of other previous studies (Morton et al. 2003; Bayer et al. 2009; Morton 2009; Schwartz et al. 2015). In our present study, *Hesperethusa crenulata* (*Naringi crenulata*) and three species of the genus *Citropsis* (*Citropsis gabunensis*, *Citropsis gilletiana*, and *Citropsis schweinfurthii*) form a single cluster. Similar relationships between *Hesperethusa* and *Citropsis* were observed in previous studies (Morton et al. 2003; Bayer et al. 2009; Morton 2009; Penjor et al. 2010, 2013; Schwartz et al. 2015). It is important to note that *Citropsis* is native to Africa, and *Hesperethusa crenulata* is native to Southeast Asia (Swingle and Reece 1967). Thus, although *Severinia buxifolia* and *Hesperethusa crenulata* are members of primitive citrus fruit trees (Swingle and Reece 1967), our present analysis shows that *Severinia buxifolia* is closely related to *Atalantia* species, and *Hesperethusa crenulata* is closely related to *Citropsis* species.

Our present analysis shows that, in Citrinae, the group containing three *Citropsis* species with *Hesperethusa crenulata* as an outgroup, and the group containing five *Atalantia* species and *Severinia buxifolia*, are distinct from "true citrus fruit trees"; in previous studies, it was difficult to detect these relationships (Morton et al. 2003; Bayer et al. 2009; Morton 2009; Penjor et al. 2010, 2013; Oueslati et al. 2016).

Relationships within "true citrus fruit trees"

In the present study, the most economically important group, "true citrus fruit trees," forms a monophyletic group and is clearly separated from the other species. With the exception of the relationships between *Citrus reticulata* (mandarin) and *Fortunella japonica* (round kumquat) (compare Fig. 3a and Supplementary Figs. 2, 4, and 8A with Figs. 1, 2, and 3b, and Supplementary Figs. 1, 3, 5–7, and 8B), similar topologies were obtained in all the phylogenetic trees. Further studies to determine the precise locations of *Citrus reticulata* and *Fortunella japonica* are required.

In our present analysis, we mainly used nuclear markers, because the small portion of the whole genome is chloroplast or mitochondrial, and the maximum depth filtering in the pyRAD program can remove these organellar sequences. In concordance with a previous study (Schwartz et al. 2015) using nuclear and chloroplast markers, our analysis showed that *Poncirus trifoliata* (trifoliate orange) is an outgroup of the remaining "true citrus fruit trees." In previous studies using chloroplast markers (Bayer et al. 2009; Penjor et al. 2013) or whole chloroplast genome (Carbonell-Caballero et al. 2015), *Poncirus trifoliata* was not an outgroup of the remaining "true citrus fruit trees"; however, the analysis based on whole chloroplast genome indicated that the paternal parent of *Poncirus trifoliata* may be a non-citrus plant (Carbonell-Caballero et al. 2015). Therefore, paternal parentage of *Poncirus trifoliata* contributes to the results, indicating that *Poncirus trifoliata* is an outgroup of the remaining "true citrus fruit trees."

Clymenia, *Eremocitrus*, and *Microcitrus* are native to Australia/New Guinea, and the previous study discriminated these genera from Asian citrus trees, including *Citrus*, *Poncirus*, and *Fortunella* (Penjor et al. 2013). However, our previous molecular analysis (Penjor et al. 2013) shows that these Australian/New Guinean citrus trees do not form a single clade—*Clymenia* was isolated from *Eremocitrus*, *Microcitrus*, and Asian citrus trees. In our present analysis, *Clymenia*, *Eremocitrus*, and *Microcitrus* form a monophyletic clade. In other previous studies, *Clymenia*, *Eremocitrus*, and *Microcitrus* also formed a monophyletic clade, although *Citrus medica* was nested in this clade (Bayer et al. 2009; Schwartz et al. 2015). Geographic separation may explain the genetic separation between Asian and Australian/New Guinean citrus trees.

In concordance with a previous study (Schwartz et al. 2015) using nuclear and chloroplast markers, our present analysis, mainly using nuclear markers, shows that Australian/New Guinean citrus trees are not an outgroup of Asian citrus trees. In contrast, in previous studies using chloroplast markers (Bayer et al. 2009; Penjor et al. 2013) or whole chloroplast genome (Carbonell-Caballero et al. 2015), Australian/New Guinean citrus trees were an outgroup of Asian citrus trees. Species belonging to five genera (*Citrus*, *Fortunella*, *Poncirus*, *Eremocitrus*, and *Microcitrus*) can cross-hybridize (Iwamasa et al. 1988). Cross-compatible features of these genera may contribute to the inconsistency between nuclear-based and chloroplast-based results. One possible explanation for this inconsistency is chloroplast capture—the introgression of a chloroplast from one species into another (Rieseberg and Soltis 1991; Tsitrone et al. 2003). Probably, ancestors of Australia/New Guinean citrus trees captured the chloroplast genome from ancestors of Asian citrus trees through inter-species hybridization and subsequent backcrosses.

Benefits and limitations of this study

In comparison with previously published studies using small numbers of markers or short DNA sequences, RAD-Seq data produced phylogenetic trees of the subfamily Aurantioideae with higher support values, clear discriminations based on longer branch length, and elucidations of earlier branching events. Although this is a benefit of our study, there were also some limitations. We obtained several types of multiple alignments, and there were some discrepancies between different types of alignment. Currently, there is no means of generalizing the method used to create multiple alignments. As research based on large data size becomes more popular, improved techniques for analyzing these types of data will be required.

Additional limitations related to RAD-Seq have been discussed previously (e.g., Andrews et al. 2016).

Acknowledgements We thank Ms. Akiko Sakai for technical assistance. Part of this work was supported by a Grant-in-Aid for Scientific Research (26450039 to Y.N., 19405019 to M.Y) from the Japan Society for the Promotion of Science. The RAD-Seq analysis was supported by the Joint Usage/Research Program of the Center for Ecological Research, Kyoto University. We would like to thank Editage for English language editing.

Author contributions Y.N., N.K., R.M., and M.Y. designed the research; T.M., N.K., R.M., and M.Y. preserved the plants; T.M. extracted the DNA; A.N., M.H., and H.K. prepared the RAD-Seq library; Y.N. performed the bioinformatic analysis; Y.N., N.K., R.M., and M.Y. analyzed the results; Y.N. wrote the manuscript; and N.K., R.M., and M.Y. revised the manuscript; all authors approved the final version to be published.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet 17(2):81–92. https://doi.org/10.1038/nrg.2015.28

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3(10):e3376. https://doi.org/10.1371/journal.pone.0003376

Bayer RJ, Mabberley DJ, Morton C, Miller CH, Sharma IK, Pfeil BE, Rich S, Hitchcock R, Sykes S (2009) A molecular phylogeny of the orange subfamily (Rutaceae: Aurantioideae) using nine cpDNA sequences. Am J Bot 96(3):668–685. https://doi.org/10.3732/ajb.0800341

Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J (2015) A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. Mol Biol Evol 32(8):2015–2035. https://doi.org/10.1093/molbev/msv082

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) *Stacks*: building and genotyping loci *de novo* from short-read sequences. G3 (Bethesda) 1(3):171–182. https://doi.org/10.1534/g3.111.000240

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22(11):3124–3140. https://doi.org/10.1111/mec.12354

Curk F, Ancillo G, Garcia-Lor A, Luro F, Perrier X, Jacquemoud-Collet JP, Navarro L, Ollitrault P (2014) Next generation haplotyping to decipher nuclear genomic interspecific admixture in *Citrus* species: analysis of chromosome 2. BMC Genet 15(1):152. https://doi.org/10.1186/s12863-014-0152-1

Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet JP, 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(5):e0125628. https://doi.org/10.1371/journal.pone.0125628

Eaton DA (2014) PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. Bioinformatics 30(13):1844–1849. https://doi.org/10.1093/bioinformatics/btu121

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(1):49–61. https://doi.org/10.1007/s11295-010-0314-x

Garcia-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, Navarro L, Ollitrault P (2013) A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the "true citrus fruit trees" group (Citrinae, Rutaceae) and the origin of cultivated species. Ann Bot 111(1):1–19. https://doi.org/10.1093/aob/mcs227

Guerra M, Dos Santos KG, Barros E, Silva AE, Ehrendorfer F (2000) Heterochromatin banding patterns in Rutaceae-Aurantioideae—a case of parallel chromosomal evolution. Am J Bot 87(5):735–747. https://doi.org/10.2307/2656860

Iwamasa M, Nito N, Ling JT (1988) Intra- and intergeneric hybridization in the orange subfamily, Aurantioideae. In: Goren R, Mendel K (eds) Proceedings of International Citrus Congress, vol 1. Balaband, Rehovot, Israel and Margrat Publishers, Weikersheim, pp 123–130

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874. https://doi.org/10.1093/molbev/msw054

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923

Morton CM, Grant M, Blackmore S (2003) Phylogenetic relationships of the Aurantioideae inferred from chloroplast DNA sequence data. Am J Bot 90(10):1463–1469. https://doi.org/10.3732/ajb.90.10.1463

Morton CM (2009) Phylogenetic relationships of the Aurantioideae (Rutaceae) based on the nuclear ribosomal DNA ITS region and three noncoding chloroplast DNA regions, *atpB-rbcL* spacer, *rps16*, and *trnL-trnF*. Org Divers Evol 9(1):52–68. https://doi.org/10.1016/j.ode.2008.11.001

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(8):1155–1166. https://doi.org/10.1007/s001220051419

Oueslati A, Ollitrault F, Baraket G, Salhi-Hannachi A, Navarro L, Ollitrault P (2016) Towards a molecular taxonomic key of the Aurantioideae subfamily using chloroplastic SNP diagnostic markers of the main clades genotyped by competitive allele-specific PCR. BMC Genet 17(1):118. https://doi.org/10.1186/s12863-016-0426-x

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS One 7(5):e37135. https://doi.org/10.1371/journal.pone.0037135

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. Evol Trend Plant 5:65–84

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12):1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Sakaguchi S, Sugino T, Tsumura Y, Ito M, Crisp MD, Bowman DMJS, Nagano AJ, Honjo MN, Yasugi M, Kudoh H, Matsuki Y, Suyama Y, Isagi Y (2015) High-throughput linkage mapping of Australian white cypress pine (*Callitris glaucophylla*) and map transferability to related species. Tree Genet Genomes 11(6):121. https://doi.org/10.1007/s11295-015-0944-0

Samuel R, Ehrendorfer F, Chase MW, Greger H (2001) Phylogenetic analyses of Aurantioideae (Rutaceae) based on non-coding plastid DNA sequences and phytochemical features. Plant Biol 3(1):77–87. https://doi.org/10.1055/s-2001-11747

Schwartz T, Nylinder S, Ramadugu C, Antonelli A, Pfeil BE (2015) The origin of oranges: a multi-locus phylogeny of Rutaceae subfamily

Aurantioideae. Syst Bot 40(4):1053–1062. https://doi.org/10.1600/036364415X690067

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Swingle WT, Reece PC (1967) The botany of *Citrus* and its wild relatives in the orange subfamily. In: Reuther W, Webber HJ, Bachelor LD (eds) The citrus industry, vol 1. University of California, Berkeley, pp 190–430

Penjor T, Anai T, Nagano Y, Matsumoto R, Yamamoto M (2010) Phylogenetic relationships of *Citrus* and its relatives based on *rbcL* gene sequences. Tree Genet Genomes 6(6):931–939. https://doi.org/10.1007/s11295-010-0302-1

Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, Nagano Y (2013) Phylogenetic relationships of *Citrus* and its relatives based on *matK* gene sequences. PLoS One 8(4):e62574. https://doi.org/10.1371/journal.pone.0062574

Penjor T, Mimura T, Matsumoto R, Yamamoto M, Nagano Y (2014) Characterization of limes (*Citrus aurantifolia*) grown in Bhutan and Indonesia using high-throughput sequencing. Sci Rep 4(1):4853. https://doi.org/10.1038/srep04853

Penjor T, Mimura T, Kotoda N, Matsumoto R, Nagano AJ, Honjo MN, Kudoh H, Yamamoto M, Nagano Y (2016) RAD-Seq analysis of typical and minor *Citrus* accessions, including Bhutanese varieties. Breeding Sci 66(5):797–807. https://doi.org/10.1270/jsbbs.16059

Tsitrone A, Kirkpatrick M, Levin DA (2003) A model for chloroplast capture. Evolution 57(8):1776–1782. https://doi.org/10.1554/02-746

Wang X, Xu Y, Zhang S, Cao L, Huang Y, Cheng J, Wu G, Tian S, Chen C, Liu Y, Yu H, Yang X, Lan H, Wang N, Wang L, Xu J, Jiang X, Xie Z, Tan M, Larkin RM, Chen LL, Ma BG, Ruan Y, Deng X, Xu Q (2017) Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. Nat Genet 49(5):765–772. https://doi.org/10.1038/ng.3839

Xu Q, Chen LL, Ruan X, Chen D, Zhu A, Chen C, Bertrand D, Jiao WB, Hao BH, Lyon MP, Chen J, Gao S, Xing F, Lan H, Chang JW, Ge X, Lei Y, Hu Q, Miao Y, Wang L, Xiao S, Biswas MK, Zeng W, Guo F, Cao H, Yang X, Xu XW, Cheng YJ, Xu J, Liu JH, Luo OJ, Tang Z, Guo WW, Kuang H, Zhang HY, Roose ML, Nagarajan N, Deng XX, Ruan Y (2012) The draft genome of sweet orange (*Citrus sinensis*). Nat Genet 45(1):59–66. https://doi.org/10.1038/ng.2472

Data archiving statement

Sequences are available at the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index_e.shtml; Accession no. DRA005954). The *Citrus* species data were used in another study, and the accession number is DRA004200 (Penjor et al. 2016).