Front. Agr. Sci. Eng. 2022, 9(4): 627–641 https://doi.org/10.15302/J-FASE-2021436

RESEARCH ARTICLE

NEW INSIGHTS INTO THE PHYLOGENY AND SPECIATION OF KUMQUAT (*FORTUNELLA* SPP.) BASED ON CHLOROPLAST SNP, NUCLEAR SSR AND WHOLE-GENOME SEQUENCING

Chenqiao ZHU*[1], Peng CHEN*[1], Junli YE[1], Hang LI[1,2], Yue HUANG[1], Xiaoming YANG[1], Chuanwu CHEN[1,3], Chenglei ZHANG[1], Yuantao XU[1], Xiaoli WANG[1,4], Xiang YAN[4], Guangzhou DENG[3], Xiaolin JIANG[1], Nan WANG[1], Hongxing WANG[1], Quan SUN[1], Yun LIU[1], Di FENG[1], Min YU[1], Xietian SONG[1], Zongzhou XIE[1], Yunliu ZENG[1], Lijun CHAI[1], Qiang XU[1], Chongling DENG[3], Yunjiang CHENG[1], Xiuxin DENG (✉)[1]

1 Key Laboratory of Horticultural Plant Biology (Ministry of Education), Huazhong Agricultural University, Wuhan 430070, China.
2 Fruit Industry Technology Guidance Stand of Ganzhou, Ganzhou 341000, China.
3 Guangxi Academy of Specialty Crops/Guangxi Citrus Breeding and Cultivation Research Center of Engineering Technology, Guilin 541004, China.
4 Ganzhou Citrus Research Institute, Ganzhou 341000, China.

*These authors contribute equally to the work

KEYWORDS

Citrus, *Fortunella*, kumquat, phylogenetics

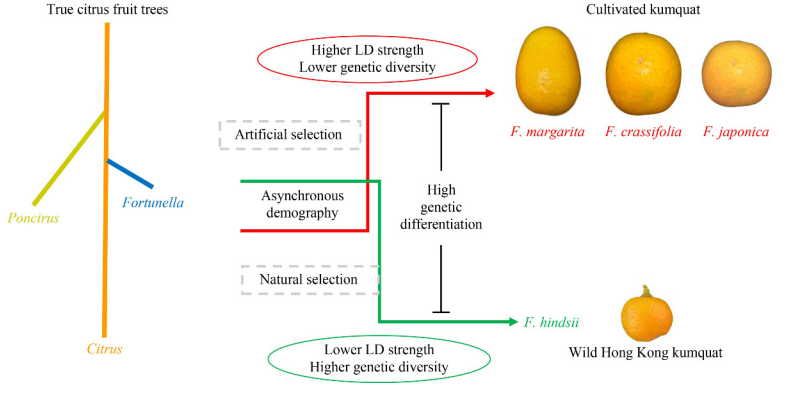
HIGHLIGHTS

- *Fortunella* genus consists of two populations: cultivated kumquat and wild Hong Kong kumquat.
- Artificial selection might involve in the origin of cultivated *Fortunella* species.
- A hypothesis for the differentiation and speciation of *Fortunella* species is proposed.

Received June 13, 2021;
Accepted November 16, 2021.

Correspondence: xxdeng@mail.hzau.edu.cn

GRAPHICAL ABSTRACT



ABSTRACT

Kumquat (*Fortunella* spp.) is a fruit and ornamental crop worldwide due to the palatable taste and high ornamental value of its fruit. Although *Fortunella* is classified into the economically important true citrus fruit tree group together with *Citrus* and *Poncirus*, few studies have been focused on its evolutionary scenario. In this study, analysis of five chloroplast loci and 47 nuclear microsatellites (nSSR) loci from 38 kumquat and 10 citrus accessions revealed the independent phylogeny of *Fortunella* among citrus taxa, and that *Fortunella* mainly comprises two populations: CUL, cultivated *Fortunella* spp. (*F. margarita*, *F. crassifolia* and *F. japonica*); and HK, wild Hong Kong kumquat

(*Fortunella hindsii*). Genomic analysis based on whole-genome SNPs indicated that the allele frequency of both pupations deviated from the neutral selection model, suggesting directional selection was a force driving their evolutions. CUL exhibited lower genomic diversity and higher linkage strength than HK, suggesting artificial selection involved in its origin. A high level of genetic differentiation (*Fst* = 0.364) was detected and obviously asynchronous demographic changes were observed between CUL and HK. Based on these results, a new hypothesis for the speciation of *Fortunella* is proposed.

© The Author(s) 2022. Published by Higher Education Press. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0)

1 INTRODUCTION

Kumquat (*Fortunella* spp.) is classified into the economically important true citrus fruit tree group together with *Citrus* and *Poncirus*, which belongs to the family of Rutaceae[1,2]. It is a common fruit crop and ornamental tree characterized by small, flavorful and brilliant fruit[3]. Given that the major edible part is the aromatic pericarp and rind, kumquat can provide more antioxidant and antimicrobial metabolites than *Citrus* spp., whose primary edible tissue is the juice vesicle[4–8]. Hence, it is also widely processed into succade (candied peel) and jam, as well as added to beverage, tea and cocktail as a natural flavor. *Fortunella* inherently possesses multiple elite agronomic traits among the citrus taxa, such as small tree size, cold and dry tolerance, short juvenility and citrus canker resistance[3,9–11]. More importantly, monoembryonic Hong Kong kumquat (*Fortunella hindsii*), which is known as a mini-citrus, has been developed as a model system for functional genomic study of citrus due to its short juvenility and sexual reproduction, which are rare features among the citrus taxa[12].

According to the classification of Swingle and Tanaka, *Fortunella* are classified based on their morphological and phenological characteristics: fewer locules (3–9) in each ovary and only two ovules in each locule, small tree size, and continuous flowering in summer[2,3,13]. Among *Fortunella*, Meiwa (*F. crassifolia*), Nagami (*F. margarita*) and Marumi (*F. japonica*) kumquat (the cultivated *Fortunella* spp.) are widely cultivated for fruit production in China, Japan, Indonesia and Malay Peninsula. Hong Kong kumquat, a wild species with the smallest hesperidium, is indigenous to southern China and mainly used for miniascape and medicine. Changshou kumquat (*Fortunella obovata*) is an ornamental cultivar in east Asia. Besides, calamondin (*Citrus madurensis*) is also regarded as a relative of *Fortunella* due to its analogous morphology[2,3,13,14].

Although phylogeny of citrus taxa has fascinated scientists for decades and is still a focus for research, *Fortunella* is considered a relative of *Citrus* and its phylogeny remains unresolved, possibly because few species (1–3) and samples (1–6) were used[15–20]. Therefore, there are still some uncertainties concerning this genus. Tanaka[13] proposed that *Fortunella* should be further divided into subgenus *Eufortunella* (*F. margarita*, *F. crassifolia* and *F. japonica*) and *Protocitrus* (*F. hindsii*) due to the primitive morphological characters of *F. hindsii*, especially the fruit organ. Swingle speculated that *F. crassifolia* might originate from the hybridization between *F. margarita* and *F. japonica* considering its intermediate fruit shape (short oblong to round) between oval (*F. margarita*) and round (*F. japonica*), or a backcrossing with *Citrus*[2]. With the rapid development of molecular biology and genomics, numerous citrus genotypes have been demonstrated to originate from hybridization or introgression, such as sweet orange (*C. sinensis*), grapefruit (*C. paradisi*), lemon (*C. limon*) and lime (*C. aurantiifolia*)[21–23], as well as calamondin (*Citrus madurensis*) and Changshou kumquat (*F. obovata*)[15–17], which poses challenges to the classic taxonomy of *Citrus* spp. Specifically, no study has comprehensively demonstrated the phylogeny and classification of *Fortunella* based on a systematic collection of various germplasm.

In addition, little is known about the origin of cultivated *Fortunella* spp. and their relationship with the only wild *Fortunella* spp., Hong Kong kumquat. *Fortunella* is academically recognized to originate from China[1,2,11,13]. The history of kumquat cultivation in China can be traced back to *Special Local Flora and Fauna in Linhai* (Ying Shen, c. 250), and it was repeatedly mentioned in later Chinese literature, such as *Guang Zhi* (Yigong Guo, c. 270), *Bei Hu Lu* (Gonglu Duan, c. 870), *Gui Tian Lu* (Xiu Ouyang, 1067), *Bian Min Tu Zuan* (Fan Kuang, 1502) and *Hua Li Bai Yong* (Changzuo Weng, c. 1718)[2,3,9,11,13,14]. However, to the best of our knowledge, no primitive population of cultivated *Fortunella* spp. has been reported either in ancient literature or modern studies. Although, *F. hindsii* has been found to be widely distributed in the primitive forests of southern China from

ancient to modern times[1,2,11,13,14]. More importantly, ancient Chinese scholars clearly distinguished cultivated *Fortunella* and *F. hindsii* in the rigorous pomology of *Citrus Record* (Yanzhi Han, 1178), herbology of *Compendium of Materia Medica* (Shizhen Li, 1578) and floriculture monographs of *Flower Mirror* (Haozi Chen, 1688), respectively. These facts suggest that cultivated *Fortuenlla* spp. were not selected in modern times, which leads to ongoing controversy about their evolutionary origin. One reasonable hypothesis is that cultivated *Fortunella* spp. originated from natural crossing or backcrossing between a primitive *Fortunella* spp. (probably *F. hindsii*) and *Citrus* spp.[18]. Another hypothesis is that cultivated *Fortunella* was directly domesticated from *F. hindsii*, because the main difference in phenotype between them is in the fruit: the fruit of *F. hindsii* are smaller, seedier and thin-rinded with a bitter and spicy taste, whereas fruit of cultivated *Fortunella* spp. are larger with thicker albedo and a sweet and palatable taste. According to the local chronicles of *New Book for Southern Life* (Yi Qian, 1016), *Composition of Chicken's Ribs* (Chuo Zhuang, 1143) and *New Anecdotes in Guangdong* (Dajun Qu, 1687), Luo Fu kumquat (*F. margarita*) was first selected from wild kumquat by monks living on Mount Luofu in Guangdong Province, and served as a tribute to emperors in the period of the Tang Dynasty. However, there has been no molecular evidence supporting these hypotheses.

With the rapid improvement of the population genetic method based on germplasm collection and molecular data, a number of novel primitive species and unexpected centers of origin of modern cultivars have been discovered, not only for citrus species[24,25], but also for some rare landscape[26] and medicinal plants[27], providing instructive information for breeding improvement and genetic conservation. Therefore, this study aimed to determine the genetic nature of *Fortunella* with a systematic collection of various germplasm, and conducted comprehensive phylogenetic and population analyses based on the chloroplast loci, nuclear microsatellites (nSSR) and genomic single nucleotide polymorphism (SNP) data. The findings provide new insights into the phylogeny, classification and evolution of *Fortunella*, which may greatly facilitate further research related to this genus.

2 METHODS

2.1 Plant materials

Thirty-eight *Fortunella* accessions including cultivars, landraces, residential garden plants and hybrids were sampled from Zhejiang, Hunan, Jiangxi, Fujian, Guangdong, Guangxi and the Citrus Research Institute of Chinese Academy of Agriculture Sciences (Chongqing, China), and 10 citrus accessions including pummelo (*Citrus maxima*), citron (*C. medica*), mandarin (*C. reticulata*), sweet orange, sour orange (*C. aurantium*), lemon, lime (*C. aurantifolia*), papeda (*C. ichangensis*), trifoliate orange (*Poncirus trifoliata*), Chinese box orange (*Atalantia buxifolia*) were sampled from the Institute of Citriculture of Huazhong Agriculture University (Wuhan, China). All the samples were prepared for chloroplast analysis and nSSR genotyping (Table 1). For whole-genome resequencing, the Hong Kong kumquat accessions collected close to cultivated environment and used in the above experiment were excluded. For synonymous accessions belonging to cultivated *Fortunella* and showing genetic similarity higher than 95% in nSSR analysis, only one sample was retained and included. Finally, 15 cultivated *Fortunella* (CUL) and 15 wild Hong Kong kumquat (HK) accessions, representative of their respective populations, were prepared for next generation sequencing. The DNA was extracted from leaves following the method developed by Cheng et al.[28].

2.2 Chloroplast loci and nuclear SSR analysis

Five chloroplast loci, including two intergenic spacers (*trn*K-*mat*K and *trn*Q-*psb*K), two introns (*rpl*16 and *rps*16) and one coding sequence (*mat*K) were amplified using the primers (Table S1) designed based on the sweet orange chloroplast genome[29]. Polymerase chain reaction amplification and amplicon sequencing followed a workflow previously described by Yang et al.[30]. Raw sequence data were imported into MEGA 7.0 and trimmed for multiple alignment[31]. Finally, a matrix of 4413 bp sequence by 48 samples was obtained for the following analysis. The five chloroplast regions of each sample were linked up for the construction of a phylogenetic tree using the maximum parsimony algorithm built in MEGA 7.0 with 1000 bootstrap replicates. The raw tree was annotated by iTOL[32]. The nucleotide polymorphism was calculated by DNASP 6.0 software[33]. The haplotype network was constructed using NETWORK 10.0.1[34].

Forty-six nSSRs (Table S2) were selected from the Sweet Orange Genome data set[22] and genotyping was performed following the protocol described by Ruiz et al.[35]. The polymorphism bands were recorded as the format of Genalex 6.5 for genetic similarity, diversity and principal coordinate analysis[36]. The genetic similarity matrix was transformed to the format of MEGA 7.0 for the construction of the phylogenic dendrogram (algorithm with 1000 bootstrap replicates). By using Genalex 6.5, the data set was transformed to the format of Structure 2.3.4[37] for genetic structure analysis. The *K* value was tested from 2 to 10 with three replicates and then the best *K* was estimated by Structure Harvester[38].

Table 1 Details and utilization of the accession evaluated in this study

Accession code	Common name	Species	Location	Description	Utilization
LF	Nagami	*Fortunella margarita*	Taizhou, Zhejiang	Germplasm	1, 2, 3
QZLF	Nagami	*F. margarita*	Quzhou, Zhejiang	Germplasm	1, 2
DGLF	Nagami	*F. margarita*	Zhejiang	Landrace	1, 2, 3
JZLF	Nagami	*F. margarita*	Zhejiang	Landrace	1, 2
WZLF	Nagami	*F. margarita*	Wenzhou, Zhejiang	Landrace	1, 2
LFHY	Nagami	*F. margarita*	Taizhou, Zhejiang	Landrace	1, 2
LW	Narumi	*F. japonica*	Ningbo, Zhejiang	Germplasm	1, 2, 3
GXCP	Narumi	*F. japonica*	Liuzhou, Guangxi	Landrace	1, 2, 3
LCHP	Meiwa	*F. crassifolia*	Guilin, Guangxi	Cultivar	1, 2
HPJG	Meiwa	*F. crassifolia*	Guilin, Guangxi	Landrace	1, 2, 3
LSHY	Meiwa	*F. crassifolia*	Yongzhou, Hunan	Germplasm	1, 2
YXBZQ	Meiwa	*F. crassifolia*	Sanming, Fujian	Landrace	1, 2, 3
JX4	Meiwa	*F. crassifolia*	Ji'an, Jiangxi	Cultivar	1, 2, 3
LYJD	Meiwa	*F. crassifolia*	Changsha, Hunan	Germplasm	1, 2, 3
WZJD	Meiwa	*F. crassifolia*	Wenzhou, Zhejiang	Germplasm	1, 2, 3
LSJG	Meiwa	*F. crassifolia*	Yongzhou, Hunan	Landrace	1, 2, 3
LYJG	Meiwa	*F. crassifolia*	Changsha, Hunan	Cultivar	1, 2
NBJD	Meiwa	*F. crassifolia*	Ningbo, Zhejiang	Cultivar	1, 2
RAJG	Meiwa	*F. japonica*	Liuzhou, Guangxi	Landrace	1, 2, 3
YSJG	Meiwa	*F. crassifolia*	Guilin, Guangxi	Cultivar	1, 2, 3
G1	Meiwa	*F. crassifolia*	Guilin, Guangxi	Landrace	1, 2
G2	Meiwa	*F. crassifolia*	Guilin, Guangxi	Landrace	1, 2, 3
DJD	Hong Kong	*F. hindsii*	Fujian	Ornamental	1, 2
XJD	Hong Kong	*F. hindsii*	Fujian	Ornamental	1, 2
WTD	Hong Kong	*F. hindsii*	Jieyang, Guangdong	Wild	1, 2, 3
JD	Hong Kong	*F. hindsii*	Fujian	Ornamental	1, 2
DR01	Hong Kong	*F. hindsii*	Longyan, Fujian	Wild	1, 2, 3
DB02	Hong Kong	*F. hindsii*	Ganzhou, Jiangxi	Wild	1, 2, 3
CZ	Hong Kong	*F. hindsii*	Chenzhou, Hunan	Wild	1, 2, 3
MJS	Hong Kong	*F. hindsii*	Ningbo, Zhejiang	Wild	1, 2, 3
WGJ	Wenguangju	Hybrid	/	Ornamental	1, 2
JGZ	/	Hybrid	/	Rootstock	1, 2
SJJ	Calamondin	*Citrus madurensis*	Fujian	Cultivar	1, 2
CS	Changshou	*F. obovata*	Fujian	Ornamental	1, 2
HCYJG	Meiwa	*F. crassifolia*	Ganzhou, Jiangxi	Germplasm	1, 2, 3
XLF	Nagami	*F. margarita*	Ningbo, Zhejiang	Landrace	1, 2, 3
YXJG	Meiwa	*F. crassifolia*	Sanming, Fujian	Cultivar	1, 2
YCJD	Meiwa	*F. crassifolia*	Quanzhou, Fujian	Germplasm	1, 2
ML	Lime	*C. aurantifolia*	/	Cultivar	1, 2
XC	Sweet orange	*C. sinensis*	/	Cultivar	1, 2
ZHI	Trifoliate orange	*Poncirus*	/	Cultivar	1, 2
XY	Foshou citron	*C. medica*	/	Cultivar	1, 2

(*Continued*)

Accession code	Common name	Species	Location	Description	Utilization
YCC	Papeda	*C. ichangensis*	/	Wild	1, 2
GXMY	Pummelo	*C. maxima*	/	Cultivar	1, 2
MSYJ	Mandarin	*C. reticulata*	/	Cultivar	1, 2
HKC	Box orange	*Atalantia*	/	Relatives	1, 2
NM	Lemon	*C. limon*	/	Cultivar	1, 2
SC	Sour orange	*C. aurantium*	/	Cultivar	1, 2
CFX	Hong Kong	*F. hindsii*	Ganzhou, Jiangxi	Wild	3
JIEX	Hong Kong	*F. hindsii*	Jieyang, Guangdong	Wild	3
DYS02	Hong Kong	*F. hindsii*	Sanming, Fujian	Wild	3
LH08	Hong Kong	*F. hindsii*	Xiamen, Fujian	Wild	3
DYT01	Hong Kong	*F. hindsii*	Longyan, Fujian	Wild	3
ZX9	Hong Kong	*F. hindsii*	Wenzhou, Zhejiang	Wild	3
LY43	Hong Kong	*F. hindsii*	Ningbo, Zhejiang	Wild	3
HC27	Hong Kong	*F. hindsii*	Shaoguan, Guangdong	Wild	3
XMS	Hong Kong	*F. hindsii*	Guangzhou, Guangdong	Wild	3
JLS	Hong Kong	*F. hindsii*	Ganzhou, Jiangxi	Wild	3

Note: Utilization 1, chloroplast sequencing; 2, nSSR; and 3, resequencing.

2.3 Resequencing work flow and population genomic analysis

Ten microgramme of high-quality genomic DNA of each sample was prepared for the construction of an NGS library. The paired-end sequencing libraries with an average insert size of ~300 bp were constructed and then sequenced by using Illumina Hiseq 2500 platform with an average depth of about thirtyfold genome coverage. The raw paired-end reads were removed with the adapter and quality filtered using Trimmomatic 0.33[39] with an option of SLIDINGWINDOW: 4:15 MINLEN:36 HEADCROP:5. The clean reads were mapped to the mini-citrus reference genome V1.0[12] using BWA (0.7.12)[40] with default parameters. The SAM (sequence alignment map) files were transformed to BAM (binary alignment map) files[41] using SAMtools (1.3.1) with the q parameter set to 30, and then sorted and duplication removed with the default parameters. AddOrReplaceReadGroups procedure in Picard was performed to add a Read Group to BAM files. Population-based SNP calling was performed using SAMtools and the raw SNPs were flittered with the criteria of QUAL < 30.0 || MQ < 40.0 || DP < 5.0 using the Bcftools tool. The SNPs were annotated by using SnpEff (4.3T)[42].

Principal component analysis was performed by using GCTA (1.92.3)[43]. The population structure was estimated using ADMIXTURE (1.3.0)[44] and the *K* value was tested from 2 to 6. All the *Fst* indexes were calculated using VCFtools[45], and the high differentiation genome region was screened by a criterion of mean *Fst* > 0.3 for a 10-kb window according to the statistical distribution of global *Fst*. The linkage disequilibrium (LD) decay was calculated using PLINK (1.90)[46]. The *Pi* values were calculated by using Variscan (2.0.3)[47] with the parameters WidthSW set to 20,000 and JumpSW set to 10,000. Population demography analysis was performed using the pairwise sequentially Markovian coalescent model[48]; the paired-end clean reads were transformed to psmcfa format using the fq2psmcfa script. The mean generation time was set at 4 years for CUL and 2 years for HK. The mutation rate was assumed as 2.2×10^{-8} substitutions per site per generation as described by Wang et al.[24].

3 RESULTS

3.1 Phylogenetic analysis of *Fortunella* based on chloroplast loci

Among the 38 *Fortunella* accessions, 25 polymorphic sites (*Np*) and 11 chloroplast haplotypes (*Nh*) were identified from five chloroplast loci with a haplotype diversity (*Hd*) of 0.69, a

nucleotide diversity of (*Pi*) 7.3×10^{-4} and an average number of nucleotide difference (*Nk*) of 3.14. The locus *trn*K-*mat*K was the most polymorphic one (with 12 *Np* and 5 *Nh*), suggesting the application potential of this locus for germplasm barcoding in the future; whereas *trn*Q-*psb*K was the most conserved locus with only one polymorphic site. With the data of 10 citrus accessions added to the above data set, the *Np* and *Nh* value increased markedly to 156 and 11, respectively, resulting a *Hd* of 0.80, a *Pi* of 3.0×10^{-3} and a *Nk* of 12.77 (Table S3). These results indicated higher chloroplast conservativeness in *Fortunella* than in *Citrus*. Eight SNPs showed diagnostic value for *Fortunella*/*Citrus*, which may serve as useful markers for offspring identification in cytomixis or crossing breeding between these two genera (Table S4).

Distance-based clustering by neighbor joining revealed six main clades (Fig. 1(a)) among the 48 accessions, with Chinese box orange in clade I (black; located at the basal), citron in clade II (yellow), papeda, wild mandarin and lime in clade III (light green), *Poncirus* in clade IV (dark green), pummelo, lemon sweet orange and sour orange in clade V (olive), and all the 38 *Fortunella* accessions in clade VI (red). *Fortunella* spp. were clearly separated from citron, mandarin, pummelo and papeda, indicating an independent phylogeny of *Fortunella* in the true citrus fruit tree group. The overall tree topology indicates that *Fortunella* has a closer phylogenic relationship with *Citrus* than with *Poncirus*. Within the *Fortunell*a clade, no obvious hierarchical structure was observed and all the accessions clustered with very low genetic differences to each other, indicating the monophyletic origin of the *Fortunella* lineage. All the four known hybrid accessions, SJJ (calamondin), CS (Changshou kumquat, *F. obovata*), WGJ (Wenguangju) and JGZ (a rootstock), were clustered within the *Fortunella* clade, indicating that their female parent should be *Fortunella*.

To further clarify the cytoplasmic evolution within *Fortunella*, a haplotype network was constructed using the median-joining algorithm (Fig. 1(b)). Notably, 11 haplotypes of *Fortunella* fell into a single branch, which is divergent from the eight *Citrus* haplotypes, demonstrating the independent chloroplast origin of *Fortunella*. The 11 haplotypes could be divided into two distinct groups: CUL with five haplotypes (H_1 to H_5) and HK with six haplotypes (H_6 to H_11), suggesting a further dichotomous differentiation after the origination of the common *Fortunella* ancestor. Among the 11 haplotypes, H_1 is

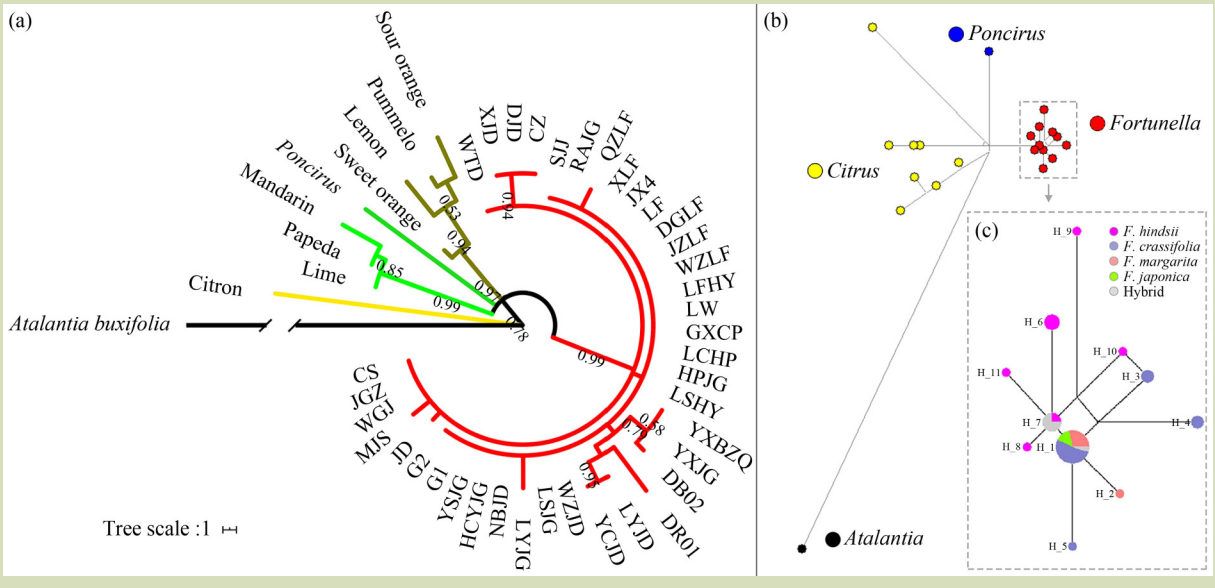


Fig. 1 Phylogenetic tree and haplotype network of *Fortunella* based on five chloroplast loci. (a) Phylogenetic tree of 38 *Fortunella* and 10 citrus accessions. Clades of the tree are highlighted by different colors. Clade I, black, Chinese box orange (*Atalantia buxifolia*); Clade II, yellow, citron (*Citrus medica*); Clade III, light green, lime (*C. aurantiifolia*), papeda (*C. ichangensis*) and wild mandarin (*C. reticulata*); Clade IV, dark green, trifoliate orange (*Poncirus trifoliata*); Clade V, olive, sweet orange (*C. sinensis*), lemon (*C. limon*), pummelo (*C. maxima*) and sour orange (*C. aurantium*); and Clade VI, red, kumquat (*Fortunella* spp.). (b) Haplotype network of 38 *Fortunella* and 10 citrus accessions. Each dot on the network presents a type of haplotype. The genera are: blue, *Poncirus*; yellow, *Citrus*; black, *Atalantia*; and red, *Fortunella*. (c) Composition of the 11 *Fortunella* haplotypes. Species are highlighted as: purple, Hong Kong kumquat (*F. hindsii*); blue, Meiwa kumquat (*F. crassifolia*); salmon, Nagami kumquat (*F. margarita*); green, Marumi kumquat (*F. japonica*); and gray, hybrid kumquat.

the most frequent one shared by 21 accessions, followed by H_7 (four accessions), H_6 (two accessions), H_3 (two accessions) and H_4 (two accessions) (Table S5). Three hybrid accessions (Wenguangju, JGZ and Changshou) shared haplotype H_7 with one HK accession, whereas the most valuable kumquat hybrid cultivar, calamondin, shared haplotype H_1 with other 20 edible CUL accessions, indicating that their female parents are different.

3.2 Genetic analysis of *Fortunella* based on nSSR markers

To further dissect the population structure of *Fortunella*, 47 nSSR loci were amplified and analyzed among the 38 kumquat and 10 citrus accessions. Two hundred and four alleles were detected among the 38 kumquat accessions. On average, the number of alleles (*Na*) and effective number of alleles (*Ne*) was 4.34 and 2.27, respectively. The allele number varied between 2 (for locus C13, D04B, A03, A21, A24, A18, E27, E28 and E30) and 8 (E6) (Table S6). The average Shannon's information index (*I*) and expected heterozygosity (*He*) was calculated as 0.92 and 0.49, respectively. E1 was the most informative locus with an *I* value of 1.58, and E30 was the least informative one (*I* = 0.39). Most of the loci (28 out of 47) showed *He* values higher than 0.5. These results indicated that *Fortunella* has higher nuclear diversity than chloroplast diversity, and this data set is more powerful for the dissection of the population structure. With the addition of the 10 citrus accessions, 325 alleles were detected. On average, the *Na* and *Ne* was 6.91 and 2.94, respectively. The *Na* varied between 2 (E27) and 15 (B26) (Table S7). The *I* and *He* was calculated as 1.27 and 0.48, respectively. B26 was the most informative locus with a high *I* value of 1.99, and E27 was the least informative one (*I* = 0.26).

The principal coordinate analysis based on *Nei's* genetic distance revealed the genetic divergence between *Fortunella* and *Citrus* accessions as well as within *Fortunella* (Fig. 2(a)). The first two principal coordinates accounted for 26.3% and 11.9% of the total genetic variance, respectively. There were two distinct groups (in green and orange dashed areas) on the positive *X* and *Y* axis formed by the 34 *Fortunella* accessions, and both were significantly differentiated from the 10 citrus accessions (in a gray dashed area); whereas the four *Fortunella* hybrid accessions formed a group intermediate between *Fortunella* and *Citrus*.

The phylogenic dendrogram constructed based on the genetic similarity matrix included the hybrid accessions (Fig. 2(b)). The overall tree topology was consistent with that of the phylogenetic tree for chloroplast and that in previous studies

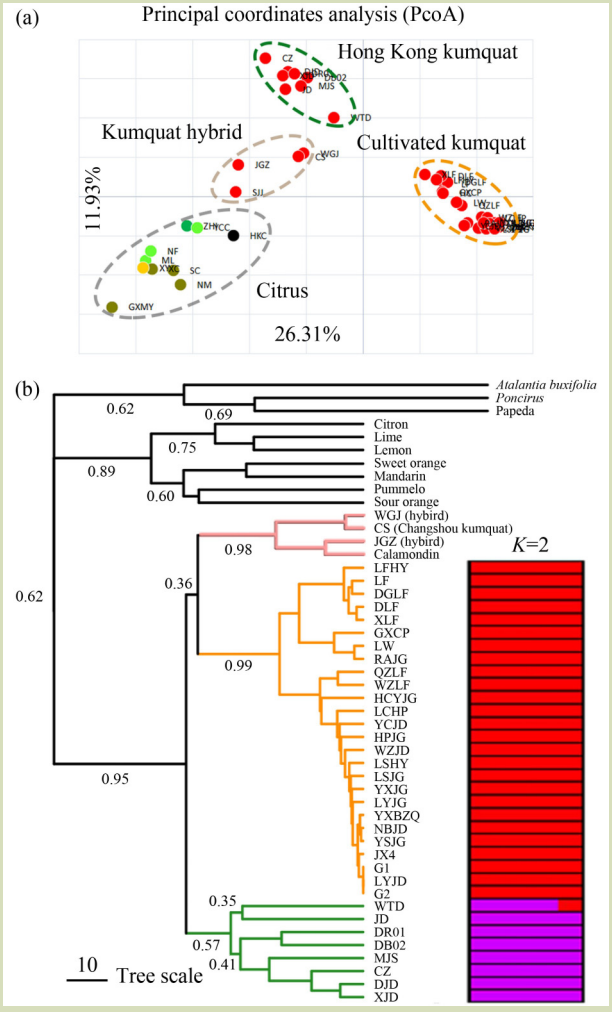


Fig. 2 Genetic structure of 38 kumquat and 10 citrus accessions based on 47 nSSR loci. (a)Principal component analysis of 38 kumquat and 10 citrus accessions. Accessions are presented by different colors (as in Fig. 1(a)). Clade I, black, Chinese box orange (*Atalantia buxifolia*); Clade II, yellow, citron (*Citrus medica*); Clade III, light green, lime (*C. aurantiifolia*), papeda (*C. ichangensis*) and wild mandarin (*C. reticulata*); Clade IV, dark green, trifoliate orange (*Poncirus trifoliata*); Clade V, olive, sweet orange (*C. sinensis*), lemon (*C. limon*), pummelo (*C. maxima*) and sour orange (*C. aurantium*); Clade VI, red, kumquat (*Fortunella* spp.). (b) Phylogenetic tree and population structure of 38 kumquat and 10 citrus accessions based on nSSR data. On the left side, phylogenetic tree was constructed using distance-based UPGMA method; clades of known kumquat hybrids (WGJ, CS, JGZ and Calamondin), cultivated *Fortunella* spp. (*F. margarita*, *F. crassifolia* and *F. japonica*) and Hong Kong kumquat (*F. hindsii*) accessions are shown in pink, orange and green, respectively. On the right side, each accession is represented by a horizontal stacked bar of genetic components with the proportion shown in color for *K* = 2 estimated by *STRUCTURE*.

based on nSSR markers, indicating that *Fortunella* is closer to *Citrus* but distant from *Poncirus*. The *Fortunella* accessions were well organized into three clusters (hybrids, CUL and HK), which was consistent with the principal coordinate analysis presented in Fig. 2(a). The first cluster included four hybrid kumquat accessions with a genetic similarity (GS) of 80.7%; the second cluster comprised 26 CUL accessions with a GS of 83.4%; and the third cluster had eight HK accessions with a GS of 76.8%. In the CUL clade, hierarchical structures were discovered, which basically corresponded to *F. margarita*, *F. japonica* and *F. crassifolia*.

The above genotyping data were further used to investigate the genetic structure of the 34 true *Fortunella* accessions, with the exclusion of the four hybrids. Evanno's test indicated a sharp signal at $K = 2$ ($\Delta K = 530.0$), implying that two gene pools (in red and purple bars) were involved in the evolution of modern *Fortunella* (Fig. 2(b) and Fig. S1). All the 26 CUL accessions only showed genetic components derived from the red ancestor. Seven out of the eight Hong Kong kumquat accessions showed single genetic components derived from the purple ancestor, while the remaining one (WTD) exhibited a mixture of genetic components, with 79.8% of HK and 20.2% of CUL.

The nSSR analysis combined with chloroplast analysis demonstrated the independent phylogeny of *Fortunella* among citrus taxa and indicated the monophyletic origin of all *Fortunella* spp.. Furthermore, these results also implied the subdivision of *Fortunella* into two lineages corresponding to CUL and HK.

3.3 Comparative genomic analysis between cultivated *Fortunella* and wild Hong Kong kumquat populations

Given the high morphological similarity[2,13] and obviously different fruit phenotype (Fig. 3(a,b) between cultivated *Fortunella* and wild Hong Kong kumquat, a final data set consisting of 5,104,141 high-quality SNPs (Table S8) genotyped from 15 CUL accessions from the main production areas (population CUL) and 15 HK accessions from primitive forests was obtained by whole-genome sequencing to reveal the genetic relationship between these two populations (Table 1).

To estimate the most likely ancestral model between CUL and HK, we predefined the number of ancestral numbers (K) from two to six, and evaluated the confidence by cross-validation (CV) (Fig. S2). The minimum CV error (0.43) was observed at $K = 2$, which clearly confirmed that *Fortunella* comprises the two populations (Fig. 3(c)). Intriguingly, three HK accessions (DB02, JLS and HC27) from Jiulianshan (Ganzhou, Jiangxi Province) showed admixed genetic background, suggesting they were subjected to recent introgressions from CUL or shared ancestral variations with CUL. In addition, to determine the genetic structure of CUL, the genetic structures at $K = 3$ and 4 (CV error = 0.49 and 0.50, respectively) were also plotted. Unexpectedly, at $K = 3$, two subgroups were identified in HK, with some accessions admixed between them. At $K = 4$, CUL diverged into two subpopulations: *F. margarita* and *F. crassifolia*. It is also out of expectation that all the *F. japonica* accessions showed an admixture background, indicating that *F. japonica* instead of *F. crassifolia* has a hybrid background. This result challenges the hypothesis proposed by Swingle that *F. crassifolia* is a hybrid of *F. margarita* and *F. japonica*.

In the genomic diversity analysis, the segregating sites, mutation number (*Eta*), singleton number (*Eta_E*) of HK were obviously higher than those of CUL (Table 2), indicating a higher level of allelic variation in HK. The *Pi* and *Theta* of HK (0.23 and 0.26) were nearly twofold those of CUL (0.12 and 0.10), indicating a higher general genomic diversity of HK than CUL. Notably, the neutral test statistics (*Tajima's D*, *Fu & Li D** and *Fu & Li F**) of both CUL and HK were deviated from zero, but distributed in opposite polarities, indicating that directional selection might have occurred in their evolution history but in opposite directions (domestication and natural selection). The lineage disequilibrium (LD) strength of the two populations was further compared (Fig. S3). The LD of CUL (orange) decayed to half at ~20 kb, while that of HK (green) decayed to half at ~10 kb, indicating the LD strength of CUL is generally higher than that of HK. Collectively, according to the lower genetic diversity and stronger LD strength of CUL, it can be speculated that artificial selection might have been involved in its origin; while for HK, given its higher genetic diversity and weaker LD strength, it can be inferred that natural selection might have been the key driving force for its evolution.

3.4 Genetic differentiation and demographic history analyses between cultivated *Fortunella* and wild Hong Kong kumquat populations

To investigate the level of genetic differentiation between CUL and HK, the *Fst* between CUL and each geographic group of HK was calculated (Table 3). The *Fst* between CUL and HK was 0.364, which is a relatively high level of genetic differentiation for perennial tree species[49,50]. Although the three CUL species showed close genetic relationship in the chloroplast and nSSR analysis, the genetic differentiation level

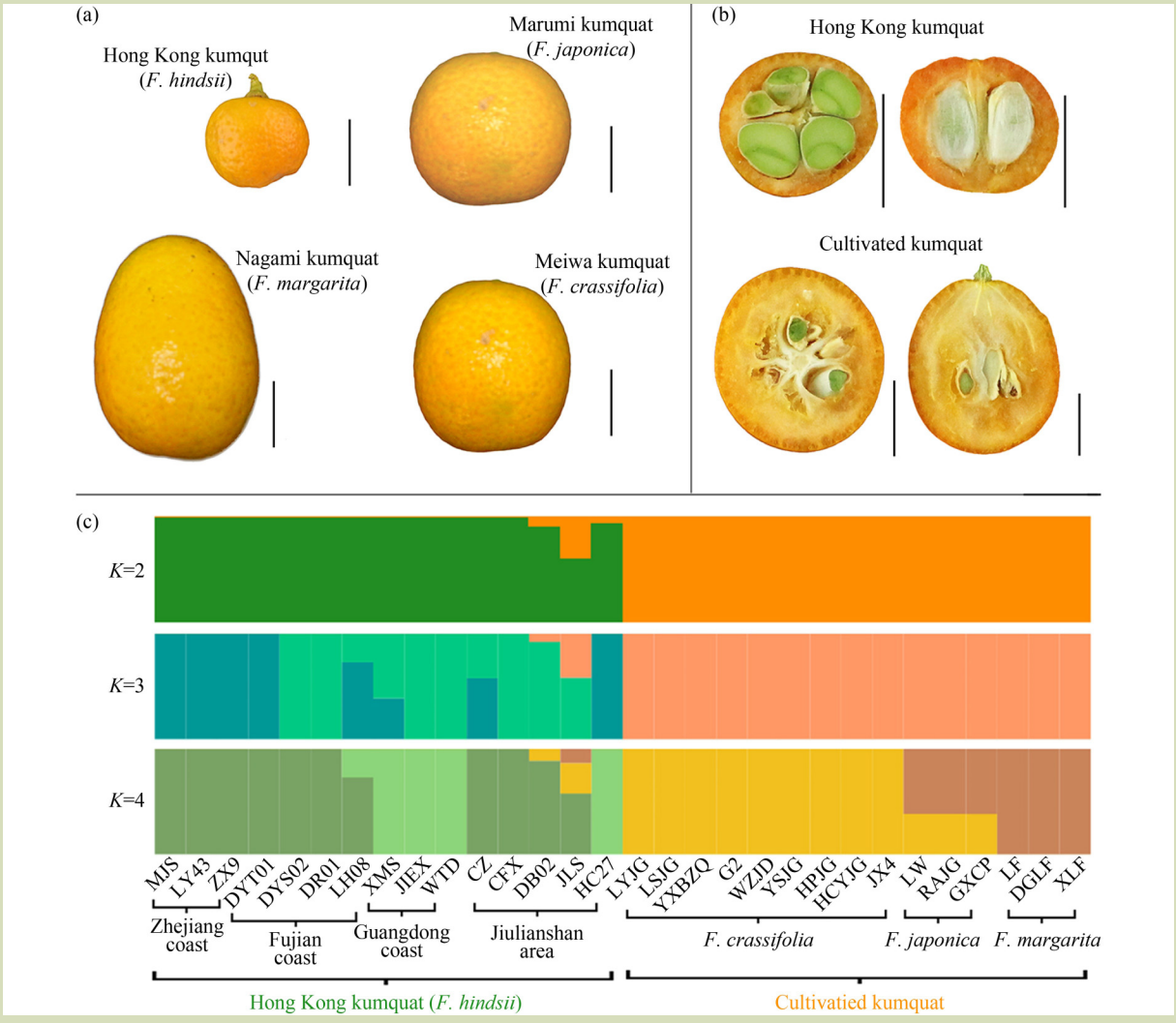


Fig. 3 Phenotype and population structure of cultivated *Fortunella* spp. (CUL) and wild Hong Kong kumquat (HK). (a) Fruit phenotypes of the four *Fortunella* spp. (b) Cross and longitudinal sections of CUL and HK; CUL has larger fruit organ with thickened and sweet albedo, whereas HK has smaller fruit with thin and acerb peel. (c) Population structure among 15 CUL and 15 HK accessions based on whole-genomic 5,104,141 SNPs. Each accession is represented by a vertical stacked column of genetic components with the proportion shown in color for K = 2, 3 and 4 estimated by ADMIXTURE.

between each pair is higher than that of the any pair of pummelo, citron, mandarin and papeda[23,24,51], indicating they should be designated to three different species. Among the three CUL species, the highest *Fst* was detected between *F. margarita* and *F. crassifolia*, which again supports the hybrid origin of *F. japonica*.

To trace the potential domestication clues between CUL and HK, the highly differentiated genomic regions were screened by the criterion of *Fst* > 0.3 according to the statistical distribution of global *Fst* (median = 0.3) (Fig. S4). In total, 138 blocks on 79 contigs were identified (Table S10), which contained 747 protein coding genes (supplementary data set). Gene ontology analysis of these genes (Fig. S5 and supplementary data set) showed that acylglycerol acyltransferase activities (GO:0019432 and GO:0046463) and glyceride biosynthesis processes (GO:0019432 and GO:0046463) were highly enriched, which might be related to the high drought and cold tolerance of CUL[3,11,52]. We further manually annotated the 747 genes and their adjacent regions, and 36 genes involved in the tricarboxylic acid cycle were identified (Table S11).

To trace back the demographic history of CUL and HK, the

Table 2 Genomic diversity of cultivated *Fortunella* (CUL) and wild Hong Kong kumquat (HK) populations

Statistics	CUL	HK
Number of segregating sites	3,737,798	9,140,932
Total number of mutations	3,762,401	9,269,923
Number of singletons	706,304	3,862,794
Pi	0.12	0.23
Theta	0.10	0.26
Tajima_D	0.61	−0.46
FuLi_Dstar	0.75	−0.46
FuLi_Fstar	0.82	−0.53

Note: *Pi*, nucleotide diversity statistic; *Theta*, Watterson's estimator Theta; *Tajima_D*, Tajima's D test statistic; *FuLi_Dstar*, *Fu* & *Li's D** statistic; *FuLi_Fstar*, *Fu* & *Li's F** statistic.

Table 3 Genomic divergence between cultivated *Fortunella* (CUL) and wild Hong Kong kumquat (HK) populations

Comparison set	*Fst*
CUL vs HK	0.364
CUL vs HK (Jiulianshan)	0.361
CUL vs HK (Guangdong coast)	0.438
CUL vs HK (Fujian coast)	0.456
CUL vs HK (Zhejiang coast)	0.469
F. margarita vs *F. japonica*	0.338
F. margarita vs *F. crassifolia*	0.345
F. japonica vs *F. crassifolia*	0.327

Note: The resequencing data (15 cultivated *Fortunella* and 15 wild Hong Kong kumquat accessions) are archived in NCBI under BioProject PRJNA736109. The matchup between accession code in paper and data code in archive please see Table S9.

pairwise sequentially Markovian coalescent model was used to infer the fluctuations in the effective population size (*Ne*) over time. As shown in Fig. 4(a), obviously asynchronous *Ne* curves were detected for CUL and HK. CUL first exhibited a decline in *Ne* (known as a bottleneck) during ~0.7–1.2 mya, which might be associated with climatic variations in the Quaternary glacial period (QGP; ~0.02–3.0 mya); whereas HK later experienced a similar bottleneck during ~0.3–0.6 mya. These results suggested that a niche or geographic isolation between CUL and HK had been established during or before QGP. Therefore, the earlier bottleneck of CUL implied its higher latitude or altitude distribution than HK. As shown in the population structure analysis, there was very limited gene flow between CUL and HK, suggesting that the distribution of CUL during QGP was likely to be of higher latitude. After the bottleneck, both CUL and HK underwent *Ne* fluctuation, which still showed asynchronous trends, suggesting their different spatiotemporal distributions during the interglaciation.

4 DISCUSSION

4.1 Phylogeny and classification of *Fortunella*

To the best of our knowledge, this is the first comprehensive study focusing on the phylogeny of *Fortunella*. According to the high chloroplast conservation and low haplotype diversity of *Fortunella*, as well as the paralleling phylogeny of *Fortunella* to *Citrus* and the intrinsically different flowering seasons of *Fortunella* (from summer to autumn) and *Citrus* spp. (in spring)[2,13], it can be speculated that after differentiation from the common ancestor of the true citrus fruit tree group, the *Fortunella* lineage underwent a relatively independent evolutionary trajectory, which was in agreement with the previous phylogenic studies of *Citrus* spp. based on chloroplast and nSSR data[15,19,20,53]. Thus, it can be confirmed that the crossing event between *Fortunella* and *Citrus* was not involved

in the origin of CUL. Since CUL and HK accessions were closely clustered in the same clade without obvious hierarchical structure in haplotype network of chloroplast, it seems not reasonable to further classify *Fortunella* genus into subgenus *Protocitrus* and *Eufortunella* as proposed by Tanaka[13]. The genetic structure analysis demonstrated that *F. japonica* instead of *F. crassifolia* has a hybridization genetic background, rejecting the hypothesis that *F. crassifolia* is a natural hybrid between *F. margarita* and *F. japonica*[2]. Each pair of cultivated *Fortunella* species showed a relatively high level of genetic differentiation, indicating that each of them are justifiably ranked as a species and rejecting the concept of a *F. margarita* complex[18]. According to the results of the present study, especially the asynchronous demographic changes between HK and CUL, we could modify the hypothesis proposed by Yasuda et al.[18] as follows. *F. hindsii* is a surviving ancestor for other *Fortunella* spp. and modern cultivated *Fortunella* might have derived from numerous mutations and selections involving *F. hindsii* or other extinct *Fortunella* ancestors. However, it remains unclear whether there is a direct domestication relationship between CUL and HK. Wild collection with larger scale, fine annotation of the *Fortunella* genomes, genetic mapping of key genes involved in the different fruit phenotypes between CUL and HK, and related gene function researches may comprehensively provide further answer to this question.

4.2 Geographic origin of *Fortunella*

Although *Fortunella* is considered to have originate in China[2,3,9,11,13,14], no solid molecular evidence has been reported. Here, the demographic history analysis suggested that the ancient distribution of CUL should be closer to the north than HK. The current distribution of wild HK is mainly in mountainous and coast area of southern China[54]. These facts provide the first molecular evidence for the continental origin of cultivated *Fortunella*, which still needs fossil evidence for validation. Furthermore, because Nanling Mountains (24°–26° N, 110°–115° E) is the northern border of wild *F. hindsii* distribution and has been proven to be the centers of origin of citrus species, such as *C. ichangensis* and *C. reticulata*[24,25], we speculate that Nanling might be the main geographic barrier for the gene flow between primitive CUL and HK during QGP. Since the admixed genetic background of HK-Jiulianshan (belonging to Nanling) population has been detected, further wild investigation and germplasm collection in this area is necessary to determine whether the genetic introgression is caused by natural pollination from the kumquat gardens nearby, or primitive populations of *F. margarita*, *F. crassifolia* and *F. japonica* still survive in the glacial refuge in Nanling.

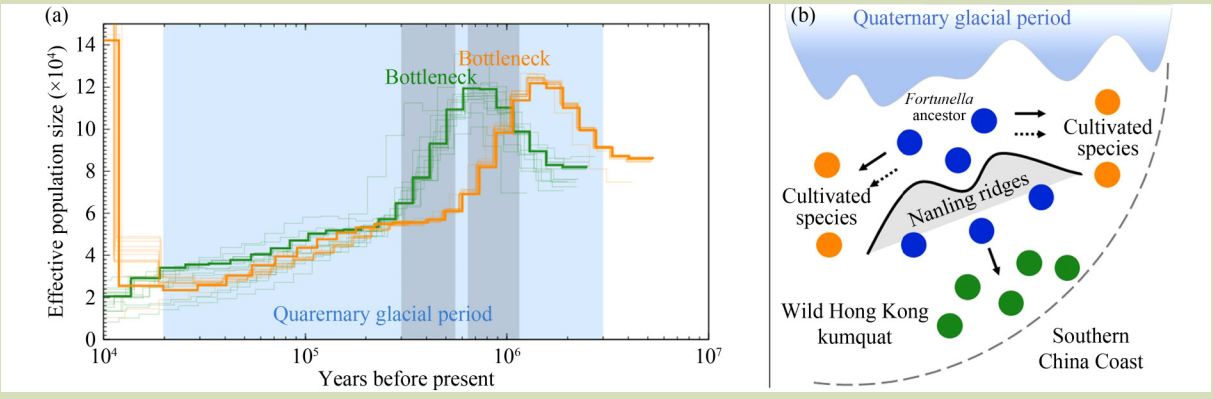


Fig. 4 Demographic history and speciation hypothesis of *Fortunella*. (a) Demographic history of cultivated *Fortunella* (CUL) and wild Hong Kong kumquat (HK) populations. Effective population size of the CUL (orange curve) and HK (green curve) were reconstructed by using the pairwise sequentially Markovian coalescent model. Quaternary glacial period is marked by blue background. Obvious population bottlenecks are marked by gray background. (b) Speciation hypothesis of *Fortunella* spp. Blue, orange and green dots represent the common ancestor of *Fortunella*, cultivated *Fortunella* spp. (*F. margartita*, *F. crassifolia* and *F. japonica*) and Hong Kong kumquat (*F. hindsii)* populations, respecitvely. The peaked line represents geographical barrier; straight arrow represents natural selection; dotted arrow indicates artificial selection. The northern population experienced earlier and severer climatic changes during Quaternary glacial period (QGP) compared to the southern population. During QGP, the northern and southern populations were gradually isolated from each other by geographical barrier (probably Nanling Mountains) and underwent adaptive evolution separately. With the southward migration of humans, modern cultivated *Fortunella* spp. were selected from the northern population.

4.3 Hypothesis for the speciation and evolution of *Fortunella*

Based on the results of this work and previous studies[12,23,55], we propose a new hypothesis about the evolutionary history of *Fortunella* (Fig. 4(b)). After differentiation from the *Citrus* lineage (~5–6 mya), the ancestor of *Fortunella* evolved into an independent lineage widely distributed in central and southern China. Along with the progression of QGP, the northern and southern populations of *Fortunella* were gradually isolated from each other (possibly by Nanling mountains). The northern *Fortunella* population was confronted with earlier and more severe natural selection (cold and dry), and thus experienced an earlier QGP bottleneck, which resulted in adaptive evolution such as thickened albedo with enrichment of sugar and secondary metabolites to protect the seeds from freezing. However, the southern population encountered moderate and later natural selection, and thus experienced later bottleneck and maintained the phenotype of primitive fruit. Along with the southward migration of humans[56,57], a few individuals of the northern population were selected and cultivated, and thus survive till the present as Luo Fu or Nagami (*F. margarita*), Jin Dan or Meiwa (*F. crassifolia*) and Luo Wen or Marumi (*F. japonica*). However, the southern population mainly underwent continuous natural selection and was discovered successively by ancient Chinese horticulturalists and modern western scholars, and named as Shan Jin Gan and Hong Kong kumquat, respectively.

5 CONCLUSIONS

In this work, by phylogenetic analysis based on chloroplast and nSSR data and population genomic analysis based on SNP data, we provide some new insights into the phylogeny, classification, and historical demography of *Fortunella*. First, *Fortunella* has an independent phylogeny among the true citrus fruit trees, and comprises two main populations corresponding to cultivated *Fortunella* spp. and Hong Kong kumquat. *F. japonica* instead of *F. crassifolia* has a hybrid origin. Artificial selection might involve in the evolution of cultivated *Fortunella* spp. instead of crossing between *Fortunella* and *Citrus*. A new hypothesis about the speciation of *Fortunella* has been proposed based on the results of the present study. Future research could focus on the domestication relationship between *F. hindsii* and cultivated *Fortunella*. These germplasms, data, results and perspectives would not only serve as useful resources for genetic improvement of kumquat and citrus, but also contribute to further evolutionary studies of citrus taxa in the future.

Supplementary materials

The online version of this article at https://doi.org/10.15302/J-FASE-2021436 contains supplementary materials (Figs. S1–S5; Tables S1–S11; supplementary data set). Resequencing data are available in the CNCB-NGDC database (BioProject PRJCA005010) and NCBI database (BioProject PRJNA736109).

Acknowledgements

This work was funded by the National Key Research and Development Program of China (2018YFD1000106), the National Natural Science Foundation of China (31630065), and Special Project for External Science and Technology Cooperation of Science and Technology Department of Yunnan Province (202003AD150014). We sincerely thank Prof. Jian Li (Department of Agriculture of Fujian Province, China), Prof. Jian'guo Xu (Citrus Research Institute of Zhejiang Province, China), Prof. Shiping Zhu (Citrus Research Institute of CAAS, China), Prof. Wu Wen (Guangdong Academy of Agricultural Sciences, China), Zhenhua Tan (Agricultural Bureau of Yizhang, China), Prof. Guohua Li (Meizhou Academy of Agricultural Sciences, China), and Prof. Min Zhang (Zhejiang A&F University, China) for their assistance in sample collection.

Compliance with ethics guidelines

Chenqiao Zhu, Peng Chen, Junli Ye, Hang Li, Yue Huang, Xiaoming Yang, Chuanwu Chen, Chenglei Zhang, Yuantao Xu, Xiaoli Wang, Xiang Yan, Guangzhou Deng, Xiaolin Jiang, Nan Wang, Hongxing Wang, Quan Sun, Yun Liu, Di Feng, Min Yu, Xietian Song, Zongzhou Xie, Yunliu Zeng, Lijun Chai, Qiang Xu, Chongling Deng, Yunjiang Cheng, and Xiuxin Deng declare that they have no conflicts of interest or financial conflicts to disclose. This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Huang C. Flora of China. Beijing: *Science Press*, 1997 (in Chinese)
2. Swingle W T, Reece P C. The botany of *Citrus* and its wild relatives. In: Reuther W, Webber H J, Batchelor L D, eds. The Citrus Industry. Berkeley: *University of California,* 1967, 190–430
3. Zhou K, Ye M. Fruit Trees in China. Citrus Volume. Beijing: *China Forestry Publishing House*, 2009 (in Chinese)
4. Barreca D, Bellocco E, Laganà G, Ginestra G, Bisignano C. Biochemical and antimicrobial activity of phloretin and its glycosilated derivatives present in apple and kumquat. *Food Chemistry*, 2014, **160**: 292–297
5. Sadek E S, Makris D P, Kefalas P. Polyphenolic composition and antioxidant characteristics of kumquat (*Fortunella margarita*) peel fractions. *Plant Foods for Human Nutrition*, 2009, **64**(4): 297–302
6. Wang Y W, Zeng W C, Xu P Y, Lan Y J, Zhu R X, Zhong K, Huang Y N, Gao H. Chemical composition and antimicrobial activity of the essential oil of kumquat (*Fortunella crassifolia* Swingle) peel. *International Journal of Molecular Sciences*, 2012, **13**(3): 3382–3393
7. Nagahama K, Eto N, Shimojo T, Kondoh T, Nakahara K, Sakakibara Y, Fukui K, Suiko M. Effect of kumquat (*Fortunella crassifolia*) pericarp on natural killer cell activity *in vitro* and *in vivo*. *Bioscience, Biotechnology, and Biochemistry*, 2015, **79**(8): 1327–1336
8. Lou S N, Ho C T. Phenolic compounds and biological activities of small-size citrus: kumquat and calamondin. *Journal of Food and Drug Analysis*, 2017, **25**(1): 162–175
9. Deng X, Peng S. Citrus Science. Beijing: *China Agriculture Press*, 2013 (in Chinese)
10. Fu X Z, Gong X Q, Zhang Y X, Wang Y, Liu J H. Different transcriptional response to *Xanthomonas citri* subsp. *citri* between kumquat and sweet orange with contrasting canker tolerance. *PLoS One*, 2012, **7**(7): e41790
11. Ye Y M. Kumquat germplasms in China. *China Seed Industry*, 1983, **4**: 2–5 (in Chinese)
12. Zhu C, Zheng X, Huang Y, Ye J, Chen P, Zhang C, Zhao F, Xie Z, Zhang S, Wang N, Li H, Wang L, Tang X, Chai L, Xu Q, Deng X. Genome sequencing and CRISPR/Cas9 gene editing of an early flowering mini-citrus (*Fortunella hindsii*). *Plant Biotechnology Journal*, 2019, **17**(11): 2199–2210
13. Tanaka T. Species problem in *Citrus*: a critical study of wild and cultivated unites of citrus, based upon field studies in their native homes (*Revisio Aurantiacearum* IX). *Japanese Society for the Promotion of Science*, 1954: 1–141
14. Deng X, Peng C, Chen Z, Deng Z, Xu J, Li J, Liu Y, Tang X, Zhong G. Citrus Varieties in China. Beijing: *China Agriculture Press*, 2008 (in Chinese)
15. Barkley N A, Roose M L, Krueger R R, Federici C T. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theoretische und Angewandte Genetik*, 2006, **112**(8): 1519–1531
16. Cheng Y, de Vicente M C, Meng H, Guo W, Tao N, Deng X. A set of primers for analyzing chloroplast DNA diversity in *Citrus* and related genera. *Tree Physiology*, 2005, **25**(6): 661–672
17. Pang X M, Hu C G, Deng X X. Phylogenetic relationships within *Citrus* and its related genera as inferred from AFLP markers. *Genetic Resources and Crop Evolution*, 2007, **54**(2): 429–436
18. Yasuda K, Yahata M, Komatsu H, Kunitake H. Phylogeny and classification of *Fortunella* (Aurantioideae) inferred from DNA polymorphisms. *Bulletin of the Faculty of Agriculture Miyazaki University*, 2010, **56**: 103–110
19. Garcia-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, Navarro L, Ollitrault P. 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. *Annals of Botany*, 2013, **111**(1): 1–19
20. Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Molecular Biology and Evolution*, 2015, **32**(8): 2015–2035
21. Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P. Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. *Annals of Botany*, 2016, **117**(4): 565–583
22. Xu Q, Chen L L, Ruan X, Chen D, Zhu A, Chen C, Bertrand D, Jiao W B, Hao B H, Lyon M P, Chen J, Gao S, Xing F, Lan H, Chang J W, Ge X, Lei Y, Hu Q, Miao Y, Wang L, Xiao S, Biswas M K, Zeng W, Guo F, Cao H, Yang X, Xu X W, Cheng Y J, Xu J, Liu J H, Luo O J, Tang Z, Guo W W, Kuang H, Zhang H Y, Roose M L, Nagarajan N, Deng X X, Ruan Y. The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics*, 2013, **45**(1): 59–66
23. Wu G A, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, Perrier X, Ruiz M, Scalabrin S, Terol J, Takita M A, Labadie K, Poulain J, Couloux A, Jabbari K, Cattonaro F, Del Fabbro C, Pinosio S, Zuccolo A, Chapman J, Grimwood J, Tadeo F R, Estornell L H, Muñoz-Sanz J V, Ibanez V, Herrero-Ortega A, Aleza P, Pérez-Pérez J, Ramón D, Brunel D, Luro F, Chen C, Farmerie W G, Desany B, Kodira C, Mohiuddin M, Harkins T, Fredrikson K, Burns P, Lomsadze A, Borodovsky M, Reforgiato G, Freitas-Astúa J, Quetier F, Navarro L, Roose M, Wincker P, Schmutz J, Morgante M, Machado M A, Talon M, Jaillon O, Ollitrault P, Gmitter F, Rokhsar D. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication.

Nature Biotechnology, 2014, **32**(7): 656–662

24. Wang L, He F, Huang Y, He J, Yang S, Zeng J, Deng C, Jiang X, Fang Y, Wen S, Xu R, Yu H, Yang X, Zhong G, Chen C, Yan X, Zhou C, Zhang H, Xie Z, Larkin R M, Deng X, Xu Q. Genome of wild mandarin and domestication history of mandarin. *Molecular Plant*, 2018, **11**(8): 1024–1037
25. Yang X M, Li H, Yu H W, Chai L J, Xu Q, Deng X X. Molecular phylogeography and population evolution analysis of *Citrus ichangensis* (Rutaceae). *Tree Genetics & Genomes*, 2017, **13**(1): 29
26. Yang X, Zhou T, Su X, Wang G, Zhang X, Guo Q, Cao F. Structural characterization and comparative analysis of the chloroplast genome of *Ginkgo biloba* and other gymnosperms. *Journal of Forestry Research*, 2020, **32**(2): 765–778
27. Wang Y, Zhou T, Li D, Zhang X, Yu W, Cai J, Wang G, Guo Q, Yang X, Cao F. The genetic diversity and population structure of *Sophora alopecuroides* (Faboideae) as determined by microsatellite markers developed from transcriptome. *PLoS One*, 2019, **14**(12): e0226100
28. Cheng Y J, Guo W W, Deng X X. Molecular characterization of cytoplasmic and nuclear genomes in phenotypically abnormal Valencia orange (*Citrus sinensis*) + Meiwa kumquat (*Fortunella crassifolia*) intergeneric somatic hybrids. *Plant Cell Reports*, 2003, **21**(5): 445–451
29. Bausher M G, Singh N D, Lee S B, Jansen R K, Daniell H. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biology*, 2006, **6**(1): 21
30. Yang X, Li H, Liang M, Xu Q, Chai L, Deng X. Genetic diversity and phylogenetic relationships of citron (*Citrus medica* L.) and its relatives in southwest China. *Tree Genetics & Genomes*, 2015, **11**(6): 129
31. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 2016, **33**(7): 1870–1874
32. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, 2016, **44**(W1): W242–W245
33. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, Guirao-Rico S, Librado P, Ramos-Onsins S E, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*, 2017, **34**(12): 3299–3302
34. Polzin T, Daneshmand S V. On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters*, 2003, **31**(1): 12–20
35. Ruiz C, Paz Breto M, Asíns M J. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. *Euphytica*, 2000, **112**(1): 89–94
36. Peakall R, Smouse P E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 2012, **28**(19): 2537–2539
37. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 2005, **14**(8): 2611–2620
38. Earl D A, Vonholdt B M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 2012, **4**(2): 359–361
39. Bolger A M, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014, **30**(15): 2114–2120
40. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009, **25**(14): 1754–1760
41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*, 2009, **25**(16): 2078–2079
42. Cingolani P, Platts A, Wang L, Coon M, Nguyen T, Wang L, Land S J, Lu X, Ruden D M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 2012, **6**(2): 80–92
43. Yang J, Lee S H, Goddard M E, Visscher P M. GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 2011, **88**(1): 76–82
44. Alexander D H, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 2011, **12**(1): 246
45. Danecek P, Auton A, Abecasis G, Albers C A, Banks E, DePristo M A, Handsaker R E, Lunter G, Marth G T, Sherry S T, McVean G, Durbin R. The variant call format and VCFtools. *Bioinformatics*, 2011, **27**(15): 2156–2158
46. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A R, Bender D, Maller J, Sklar P, de Bakker P I W, Daly M J, Sham P C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 2007, **81**(3): 559–575
47. Vilella A J, Blanco-Garcia A, Hutter S, Rozas J. VariScan: analysis of evolutionary patterns from large-scale DNA sequence polymorphism data. *Bioinformatics*, 2005, **21**(11): 2791–2793
48. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature*, 2011, **475**(7357): 493–496
49. Meirmans P G, Hedrick P W. Assessing population structure: F(ST) and related measures. *Molecular Ecology Resources*, 2011, **11**(1): 5–18
50. Weir B S, Cockerham C C. Estimating F-statistics for the analysis of population structure. *Evolution*, 1984, **38**(6): 1358–1370
51. 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 R M, Chen L L, Ma B G, Ruan Y, Deng X, Xu Q. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nature Genetics*, 2017, **49**(5): 765–772

52. Lin D, Wu F. Distribution and variety of kumquat in China. *South China Fruits*, 1987, **1**(1): 3–5 (in Chinese)
53. Garcia-Lor A, Luro F, Ollitrault P, Navarro L. Genetic diversity and population structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers. *Tree Genetics & Genomes*, 2015, **11**(6): 123
54. Chen P. The investigation and genetic diversity evaluation of wild Hong Kong kumquat (*Fotunella hindsii* swingle) in China. Dissertation for the Master's Degree. Wuhan: *Huazhong Agricultural University*, 2011 (in Chinese)
55. Wu G A, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, Domingo C, Tadeo F R, Carbonell-Caballero J, Alonso R, Curk F, Du D, Ollitrault P, Roose M L, Dopazo J, Gmitter F G, Rokhsar D S, Talon M. Genomics of the origin and evolution of *Citrus*. *Genomics of the origin and evolution of Citrus*, 2018, **554**(7692): 311–316
56. Atahan P, Itzstein-Davey F, Taylor D, Dodson J, Qin J, Zheng H, Brooks A. Holocene-aged sedimentary records of environmental changes and early agriculture in the lower Yangtze, China. *Quaternary Science Reviews*, 2008, **27**(5–6): 556–570
57. Zhang J Z, Chen C F, Yang Y Z. Origins and early development of agriculture in China. Journal of National Museum of China, 2014, **1**: 6–16 (in Chinese)