*Systematics and Biodiversity* (2019), **0(0)**: 1–13

Taylor & Francis
Taylor & Francis Group

Research Article

# Biogeography and phylogenetic relationships of Hyrcanian wild apple using cpDNA and ITS noncoding sequences

HAMED YOUSEFZADEH[1], ALI KHODADOST[2], HAMID ABDOLLAHI[3], SYED SHUJAIT ALI[4], GREGOR KOZLOWSKI[5,6] & HAMID BINA[7]

[1]Department of Environmental Science; Faculty of Natural Resources, Tarbiat Modares University, Tehran, Iran
[2]Department of Forestry, Faculty of Natural Resources, Tarbiat Modares University, Tehran, Iran
[3]Temperate Fruits Research Center, Horticultural Sciences Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran
[4]Center for Biotechnology and Microbiology (CBM), University of Swat, KPK, Pakistan
[5]Department of Biology and Botanic Garden, University of Fribourg, Chemin du Musée 10, Fribourg, CH-1700, Switzerland
[6]Natural History Museum Fribourg (NHMF), Chemin du Musée 6, Fribourg, CH-1700, Switzerland
[7]Host/Pathogen Interaction & Biotechnology of Fruit Trees Department, Seed and Plant Improvement Institute, Karaj, Iran

(Received 6 May 2018; accepted 30 January 2019)

The Hyrcanian forest of northern Iran is considered one of the potential centres for the evolution and domestication of the genus *Malus* (Rosaceae). However, the biogeography, phylogenetic position, and taxonomic status of the Hyrcanian wild apples have never been evaluated. In our study, the nucleotide sequences of the internal transcribed spacer (ITS) and the *trn*H-*psb*A intergenic spacer region from 14 natural populations were analysed. Phylogenetic analysis based on the ITS and the Maximum-likelihood (ML) tree showed that all Hyrcanian samples were closely related to *M. orientalis* and *M. asiatica* and can be placed within section *Malus* and series *Malus*. Furthermore, based on a comparison of ITS2 secondary structures, the Hyrcanian samples were identical to *M. orientalis* and *M. sieversii*. Biogeographic scenarios constructed using Statistical Dispersal-Vicariance Analysis (S-DIVA) and the Bayesian Binary Method (BBM) indicated that the ancestor of *Malus* originated during the Eocene, ~53 million years ago (Ma), and that China played a vital role in the expansion of the range of the genus. The members of *Malus* colonized the Hyrcanian region from China during the Miocene, ~22-10 Ma.

**Key words:** Divergence time, Hyrcanian forest, *Malus orientalis*, network analysis, taxonomic status

## Introduction

Despite the significance of the domesticated apple (*Malus domestica*) as one of the most important temperate fruit crops (Robinson et al., 2001), the time and place of apple evolution, its range expansion throughout the northern hemisphere, and its species number and taxonomic divisions are still not well understood. Climatic diversity among the different habitats of apples and the intrinsic natural diversity of apples in nature due to hybridization and introgression have probably played important roles in the evolution of apples and in the Rosaceae family generally (Katayama & Uematsu, 2003; Phipps, Robertson, Smith, & Rohrer, 1990), causing ambiguity about the number of extant *Malus* species. Thus, significant disagreements still exist concerning the names assigned to the species and varieties of wild apples (Forte, Ignatov, Ponomarenko, Dorokhov, & Savelyev, 2002; Janick, 2003; Savelyeva, Boris, Kochieva, & Kudryavtsev, 2013).

The genus *Malus* Mill. comprises 25–47 species with five genetic centres: East-Asiatic, Middle-Asiatic, Caucasian, European, and North American (Zhukovsky, 1965). *Malus* is traditionally divided into six sections: *Eriolobus* (downy-lobed apples), *Docyniopsis* (docynious apples), *Sorbomalus* (mountain-ash apples), *Chloromeles* (green-fruited apples), *Gymnomeles* (berry apples), and *Malus* (true apples) (Langenfeld, 1991).

Similarly, disagreement prevails about the taxonomy of wild apples in the Transcaucasian region and Iran.

Correspondence to: Hamed Yousefzadeh. E-mail: h.yousefzadeh@modares.ac.ir; hamed.yousefzadehseraj@gmail.com

ISSN 1477-2000 print / 1478-0933 online

https://dx.doi.org/10.1080/14772000.2019.1583689

Published online 05 Jun 2019

Apple is widespread in Iran from the north (Hyrcanian forest) to the west and central parts of the country, growing at different altitudes and habitats from the coastal plains and steppes into the mountain regions (Sabeti, 1976). Apple is an ancient fruit crop in Iran (Janick, 2003), and phylogenetic studies have suggested that Iran could be a major centre of diversity for domestic apples and a very important hub for its domestication and transfer from Central Asia to the West via the Silk Road (Gharghani et al., 2009). Vavilov (1930) characterized the South Caucasian centre as a 'vast wood consisting solely of the wild progenitors of fruit trees'.

Most studies have reported *M. orientalis* as the only wild apple in the flora of Iran (Rechinger, 1964; Sabeti, 1976), but Phipps et al. (1990) mentioned the presence of *M. sieversii* as well. Gharghani et al. (2009), in a comprehensive study using wild and domestic apple germplasm from different parts of the world, demonstrated that Iranian cultivars and landraces were closely related to both *M. sieversii* from Central Asia and *M. orientalis*, which is native to Iran, Turkey, Russia, and the Caucasus region. However, the wide distribution of *M. orientalis* across different altitudes leads to a high variability in vegetative morphology as well as morphological and olfactory fruit characteristics (Fischer & Schmidt, 1938; Mansfeld & Büttner, 2001). Forsline, Aldwinckle, Dickson, Luby, and Hokanson (2003) described two subspecies of *M. orientalis*: subsp. *montana* (Uglitzk.) Likh. and subsp. *turkmenorum* (Juz.) Langenf.

DNA barcoding, as an effective tool for correct species identification, has been receiving increasing attention in recent years (Bina, Yousefzadeh, Ali, & Esmailpour, 2016; Hebert, Cywinska, & Ball, 2003; Taberlet et al., 2007; Tautz, Arctander, Minelli, Thomas, & Vogler, 2003; von Cräutlein, Korpelainen, Pietiläinen, & Rikkinen, 2011). Many loci, including *rbc*L, *mat*K, *psb*A-*trn*H, *rpo*C1, and ITS2, have been popularly used as DNA barcodes in plants worldwide. The CBOL Plant Working Group (Group et al., 2009) proposed a combination of two chloroplast loci, *mat*K + *rbc*L, as the core barcode for land plants, with *trn*H-*psb*A and the nuclear ribosomal internal transcribed spacer (ITS) as a complement. Chen et al. (2010) have shown that ITS2 as a universal barcode correctly identified 92.7% of over 6600 samples in seven phyla (angiosperms, gymnosperms, ferns, mosses, liverworts, algae, and fungi). Moreover, the applicability of ITS2 in discriminating among a wide range of plants within many plant families (e.g., Asteraceae, Rutaceae, Rosaceae, etc.) has been confirmed (Gao et al., 2010; Liu et al., 2012; Pang et al., 2011; Yao et al., 2010; Yousefzadeh, Colagar, Tabari, Sattarian, & Assadi, 2012). Thus, DNA barcoding techniques may have the ability to provide significant information regarding the systematic classification of the genus *Malus*.

To date, no research has been reported on the genetic diversity of the Hyrcanian wild apple species and its genetic relationship with the other apples of the world. Therefore, the aims of this study are (1) to investigate the taxonomic status of Hyrcanian *Malus* using a DNA barcoding approach, (2) to explore the phylogenetic relationships of the Hyrcanian members within the genus *Malus*, and more generally, (3) to elucidate the historical biogeography of the genus *Malus*.

# Materials and methods

## Plant material, DNA extraction, and ITS amplification

Leaf samples were collected from 14 apple populations covering the entire distribution of species in the Hyrcanian forest (Fig. 1; Table S1, see online supplemental material, which is available from the article's Taylor & Francis Online page at https://www.doi.org/10.1080/14772000.2019.1583689). Total DNA was extracted from fresh leaves using the method of Murray and Thompson (1980) with some modifications (Janfaza, 2016). Although studies by Liu et al. (2012) indicate that 8–10 individuals per species from the entire geographic distribution of the species analysed appear to be sufficient for plant DNA barcoding, based on the DNA barcoding database (http://www.barcodinglife.org/views/login.php), 3–8 trees from each population were selected and their ITS and trnH-psbA regions were sequenced. The primers ITS-1 and ITS-4 (White, Bruns, Lee, & Taylor, 1990) were used to amplify the complete ITS regions. The *trn*H-*psb*A forward and reverse primers designed by Tate and Simpson (2003) and Sang, Crawford, and Stuessy (1997), respectively, were used. PCR amplifications were accomplished in 20 µl reactions with the AccuPower HotStart PCR Premix kit (Bioneer, Korea). The thermal cycling profile consisted of an initial denaturation step of 360 s at 95 °C, followed by 32 cycles of 60 s at 95 °C, 45 s at 56 °C, 90 s at 72 °C, and a final extension step of 5–7 min at 72 °C.

## Phylogenetic, network analysis, and species delimitation methods

The ITS region sequences were manually checked by eye with Chromas ver. 2.31 (Technelysium Pty. Ltd, South Brisbane, Australia), aligned by MUSCLE and refined manually in MEGA 7 software (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). Nucleotide composition, number of variables, parsimony-

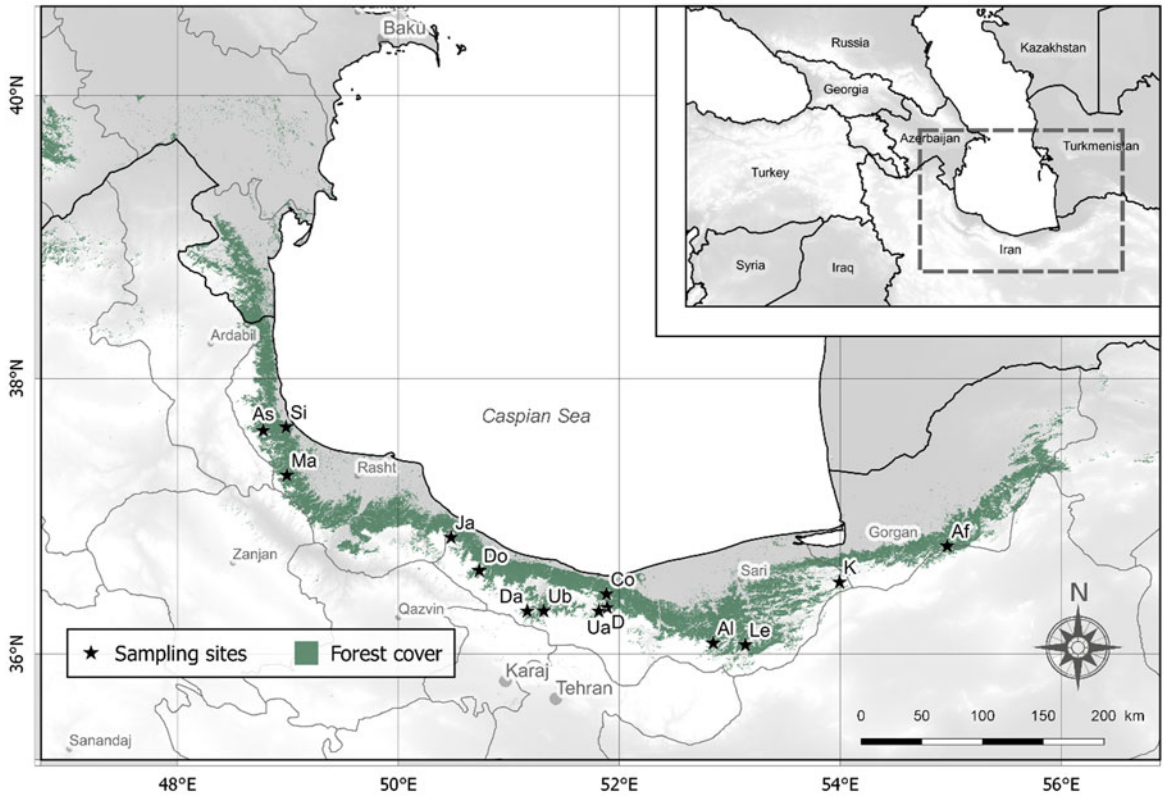


**Fig. 1.** Geographic locations of sampling sites are indicated by black asterisks.

**Table 1.** Characteristics of the aligned ITS data matrix used for phylogenetic analyses.

| | Iranian apple | | | Species apple in GenBank | | |
|---|---|---|---|---|---|---|
| | | ITS | | | ITS | |
| **Region** | ***trnH-psbA*** | **ITS1** | **ITS2** | ***trnH-psbA*** | **ITS1** | **ITS2** |
| A (%) | 42.8 | 17 | 12 | 34.2 | 17 | 12 |
| C (%) | 15.7 | 34 | 38 | 10.2 | 34.1 | 38 |
| G (%) | 10.4 | 31 | 33 | 15.8 | 31 | 32 |
| U (%) | 31.1 | 17 | 17 | 39.8 | 17 | 18 |
| Length (bp) | 372 | 591 | 218 | 372 | 500 | 272 |
| Conserved sites | 296 | 574 | 208 | 111 | 409 | 122 |
| Variable sites | 69 | 17 | 10 | 250 | 191 | 254 |
| Parsimony site | — | 5 | 2 | 178 | 104 | 72 |

informative and conserved sites for each species were calculated separately for the ITS1, 5.8S, and ITS2 regions. Maximum likelihood (ML) trees were generated using Mega 6 software with a bootstrap procedure (1,000 replications). The best-fit evolutionary model and parameters were chosen by model test based on the Akaike Information Criterion (AIC) as embedded in MEGA 6 software (Tamura et al., 2013). Mega 6 software was also used to evaluate genetic distance among the selected alignment taxa (Table S2, see supplemental material online) via the Neighbour-joining (NJ) algorithm using Kimura 2-Parameter distance; sites were weighted using gamma distribution, estimating alpha parameters with Maximum likelihood (Kimura, 1980). Secondary structure of ITS2 of the selected taxa was compared by ITS2 database (Koetschan et al., 2009; Schultz et al., 2006; Selig, Wolf, Müller, Dandekar, & Schultz, 2007). Species delimitation plug-in (SDP) (Masters, Fan, & Ross, 2011) was used to determine the taxonomic status of Iranian apple, which have recently been proposed for species delimitation by Prévot, Jordaens, Sonet, and Backeljau (2013).

### Divergence time estimate and biogeographic analysis

Likelihood ratio test (LRT) was performed in PAUP v4.0b10 (Swofford, 2002) to check whether molecular clock is appropriate for our data, however the result ($P = 0.00001$) suggests that the relax molecular clock approach is best fitted to explain the divergence time. Partition homogeneity or incongruence length difference (ILD) (Farris, Kallersjo, Kluge, & Bult, 1994) test was performed in PAUP to concatenate our data (ITS and trnH-psbA) or not. A heuristic search approach with 1,000 replicates and 100 random stepwise additions with tree bisection reconstruction (TBR) branch swapping implemented in PAUP v4.0b10 (Swofford, 2002) was followed for ILD test.

Divergence time estimates were obtained by calibrating the basal node of Pyrinae (the outgroup) to a mean age of 45 million years ago (Ma) (95% upper and lower bound are 49.8 Ma and 37 Ma) and applying the normal distribution parameter based on the previous studies of Lo Presti and Oberprieler (2009) and Semerikov, Semerikova, Polezhaeva, Kosintsev, and Lascoux (2013) with standard deviation value of 1.0. The phylogenetic trees were obtained as an output of a BEAST analysis run for 50,000,000 generations in BEAST ver. 1.6.1 (Drummond & Rambaut, 2007) using the uncorrelated lognormal relaxed clock parameter under the Yule model of speciation and the GTR + I + G model. The maximum clade credibility tree was obtained by Tree Annotator ver. 1.7.5 (Drummond, Suchard, Xie, & Rambaut, 2012) and visualized in FigTree ver. 1.4 (Rambaut, 2012).

The distribution area of *Malus* and *Pyrus* was divided into six regions based on the available samples: A (Iran), B (Western Asia), C (China), D (North America), E (Europe), and F (Eastern Asia). The biogeographic scenario was inferred by applying the event- and model-based approaches S-DIVA and BBM, which are both embedded in the RASP software (Ali, Yu, Pfosser, & Wetschnig, 2012, 2013; Yu, Harris, Blair, & He, 2015).

The uncertainties in phylogeny were overcome by loading 10,000 trees from a Markov chain Monte Carlo (MCMC) output into the RASP software. The maximum clade credibility tree and distribution file were uploaded to show the biogeographic reconstructions obtained by S-DIVA analysis. For BBM analysis, only the maximum clade credibility tree was used along with the distribution file to obtain a reconstruction. The MCMC chains were run under the JC + G (Jukes–Cantor + Gamma) model for 5,000,000 generations.

## Results

### ITS and *trn*H-*psb*A sequence characteristics and phylogenetic analysis

The nucleotide composition, total length, GC content, and sequence divergence for the amplified regions ITS2 and total ITS for Hyrcanian *Malus* and all *Malus* taxa are presented in Table 1. Phylogenetic analysis based on ITS and the ML tree showed that all species in section *Malus* were located in a distinct clade, supported by an 81% bootstrap value, and that species of the *Malus* series were completely separated from those of the *Baccata* series (Fig. 2). All Hyrcanian samples were located in section *Malus* and series *Malus* with *M. orientalis* and *M. asiatica* Nakai. Section *Sorbomalus* was not monophyletic; *M. toringoides* Hughes and *M. transitoria* C.K. Schneid. were located in the same clade as the species from section *Malus*. We also constructed phylogenetic trees of apples using the *trn*H-*psb*A region, which did not discriminate among the sections and series of *Malus* (Fig. S1, see supplemental material online).

Based on pairwise distance (K2P) among *Malus* taxa, the Hyrcanian samples had a minimum distance from *M. orientalis*. Additionally, the Hyrcanian samples showed a maximum pairwise distance from *M. hupehensis* (Pamp.) Rehder and *M. yunnanensis* C.K. Schneid. (Table 2).

### Comparison of ITS2 secondary structures

The secondary structures of the ITS2 regions in *Malus* taxa, as expected for angiosperms (Schultz, Maisel, Gerlach, Müller, & Wolf, 2005), had four helices. A comparison of secondary structure based on nucleotide composition revealed that Helix III was the longest and completely conserved, but it differed from the other helices, especially from Helix II (Table 3). Based on this comparison, the secondary structure of the Hyrcanian samples was identical with those of *M. orientalis* and *M. sieversii*. However, a certain level of variation in secondary structure was observed among different samples of *M. sieversii* and *M. orientalis*.

### Species delimitation plugin (SDP)

The result of the SDP analysis based on the MP tree phylogeny is summarized in Table 4. The highest Intra/Inter ratio was found in clade B (0.95). The highest strict and liberal values of P ID were observed for clade A (Table 4). The maximum average distance between ancestral species in a clade (MRCA-tips) was detected for clades B and E. Rosenberg's P AB values showed monophyly for these clades (>0.05 in all clades). If the

**Table 2.** Minimum and maximum pairwise genetic distance (Kimura 2-Parameter distance; sites were weighted using gamma distribution) among analysed *Malus* species.

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** | **25** | **26** | **27** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 0.02 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 0.01 | 0.02 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | 0.01 | 0.02 | 0.02 | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | 0.03 | 0.05 | 0.03 | 0.05 | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | 0.03 | 0.05 | 0.03 | 0.04 | 0.00 | | | | | | | | | | | | | | | | | | | | | | |
| 7 | 0.07 | 0.08 | 0.07 | 0.08 | 0.08 | 0.08 | | | | | | | | | | | | | | | | | | | | | |
| 8 | 0.10 | 0.10 | 0.09 | 0.11 | 0.11 | 0.11 | 0.07 | | | | | | | | | | | | | | | | | | | | |
| 9 | 0.02 | 0.04 | 0.02 | 0.03 | 0.03 | 0.03 | 0.05 | 0.08 | | | | | | | | | | | | | | | | | | | |
| 10 | 0.04 | 0.05 | 0.03 | 0.05 | 0.05 | 0.05 | 0.08 | 0.11 | 0.04 | | | | | | | | | | | | | | | | | | |
| 11 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.05 | 0.09 | 0.12 | 0.04 | 0.03 | | | | | | | | | | | | | | | | | |
| 12 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.05 | 0.08 | 0.11 | 0.05 | 0.04 | 0.04 | | | | | | | | | | | | | | | | |
| 13 | 0.03 | 0.05 | 0.04 | 0.05 | 0.06 | 0.06 | 0.10 | 0.12 | 0.05 | 0.07 | 0.07 | 0.07 | | | | | | | | | | | | | | | |
| 14 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.05 | 0.09 | 0.12 | 0.05 | 0.04 | 0.05 | 0.04 | 0.07 | | | | | | | | | | | | | | |
| 15 | 0.04 | 0.05 | 0.04 | 0.05 | 0.04 | 0.04 | 0.07 | 0.10 | 0.01 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | | | | | | | | | | | | | |
| 16 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.05 | 0.09 | 0.11 | 0.05 | 0.04 | 0.05 | 0.04 | 0.07 | 0.02 | 0.06 | | | | | | | | | | | | |
| 17 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.05 | 0.09 | 0.12 | 0.05 | 0.01 | 0.04 | 0.04 | 0.07 | 0.05 | 0.06 | 0.05 | | | | | | | | | | | |
| 18 | 0.05 | 0.07 | 0.05 | 0.06 | 0.05 | 0.06 | 0.10 | 0.12 | 0.05 | 0.05 | 0.04 | 0.05 | 0.08 | 0.05 | 0.06 | 0.05 | 0.05 | | | | | | | | | | |
| 19 | 0.04 | 0.06 | 0.03 | 0.05 | 0.05 | 0.05 | 0.09 | 0.12 | 0.05 | 0.04 | 0.05 | 0.04 | 0.07 | 0.01 | 0.06 | 0.01 | 0.05 | 0.05 | | | | | | | | | |
| 20 | 0.02 | 0.04 | 0.02 | 0.03 | 0.03 | 0.03 | 0.05 | 0.07 | 0.01 | 0.03 | 0.04 | 0.04 | 0.05 | 0.04 | 0.02 | 0.04 | 0.04 | 0.05 | 0.04 | | | | | | | | |
| 21 | 0.10 | 0.11 | 0.10 | 0.10 | 0.11 | 0.11 | 0.07 | 0.07 | 0.08 | 0.11 | 0.12 | 0.11 | 0.13 | 0.12 | 0.10 | 0.11 | 0.12 | 0.12 | 0.11 | 0.07 | | | | | | | |
| 22 | 0.05 | 0.07 | 0.05 | 0.06 | 0.06 | 0.06 | 0.09 | 0.12 | 0.06 | 0.05 | 0.05 | 0.04 | 0.08 | 0.02 | 0.07 | 0.02 | 0.06 | 0.06 | 0.02 | 0.05 | 0.12 | | | | | | |
| 23 | 0.01 | 0.02 | 0.01 | 0.02 | 0.04 | 0.04 | 0.08 | 0.10 | 0.03 | 0.04 | 0.05 | 0.05 | 0.04 | 0.05 | 0.04 | 0.05 | 0.05 | 0.06 | 0.05 | 0.03 | 0.11 | 0.06 | | | | | |
| 24 | 0.05 | 0.06 | 0.04 | 0.05 | 0.06 | 0.05 | 0.10 | 0.13 | 0.06 | 0.05 | 0.06 | 0.05 | 0.08 | 0.02 | 0.07 | 0.03 | 0.06 | 0.06 | 0.03 | 0.05 | 0.13 | 0.04 | 0.05 | | | | |
| 25 | 0.00 | 0.02 | 0.01 | 0.01 | 0.03 | 0.03 | 0.07 | 0.10 | 0.02 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.03 | 0.04 | 0.04 | 0.05 | 0.04 | 0.02 | 0.10 | 0.05 | 0.01 | 0.05 | | | |
| 26 | 0.06 | 0.07 | 0.05 | 0.07 | 0.07 | 0.07 | 0.01 | 0.06 | 0.04 | 0.07 | 0.08 | 0.07 | 0.08 | 0.08 | 0.06 | 0.07 | 0.08 | 0.09 | 0.07 | 0.03 | 0.06 | 0.08 | 0.06 | 0.08 | 0.06 | | |
| 27 | 0.04 | 0.05 | 0.04 | 0.05 | 0.05 | 0.05 | 0.08 | 0.11 | 0.02 | 0.06 | 0.06 | 0.06 | 0.07 | 0.07 | 0.03 | 0.06 | 0.06 | 0.07 | 0.06 | 0.03 | 0.11 | 0.08 | 0.05 | 0.08 | 0.04 | 0.07 | |
| 28 | 0.05 | 0.06 | 0.04 | 0.05 | 0.06 | 0.05 | 0.09 | 0.12 | 0.05 | 0.05 | 0.05 | 0.04 | 0.08 | 0.02 | 0.06 | 0.01 | 0.05 | 0.06 | 0.02 | 0.04 | 0.12 | 0.01 | 0.05 | 0.03 | 0.05 | 0.07 | 0.07 |

1-T4; 2-T5; 3-*M.orientalis*(6630899); 4-*M.transitoria*(133779100); 5-*M.toringoides*(133779121); 6-*M.asiatica*(126842417); 7-*M.baccata*(145843646); 8-*M.hupehensis*(6630904); 9-*M.coronaria*(6630925); 10-*M.doumeri*(6630930); 11-*M.florentina*(6630921); 12-*M.sylvestris*(255683063); 13-*M.fusca*(6630916); 14-*M.halliana*(6630903); 15-*M.honanensis*(6630911); 16-*M.ioensis*(6630927); 17-*M.tschonoskii*(6630928); 18-*M.kansuensis*(6630913); 19-*M.mandshurica*(6630905); 20-*M.micro*(145843647); 21-*M.yunnanensis*(6630909); 22-*M.neidzwetzkyana*(6630897); 23-*M.ombrophilia*(6630914); 24-*M.prunifolia*(6630901); 25-*M.pumila*(145843648); 26-*M.sieboldii*(6630906); 27-*M.prattii*(6630912.

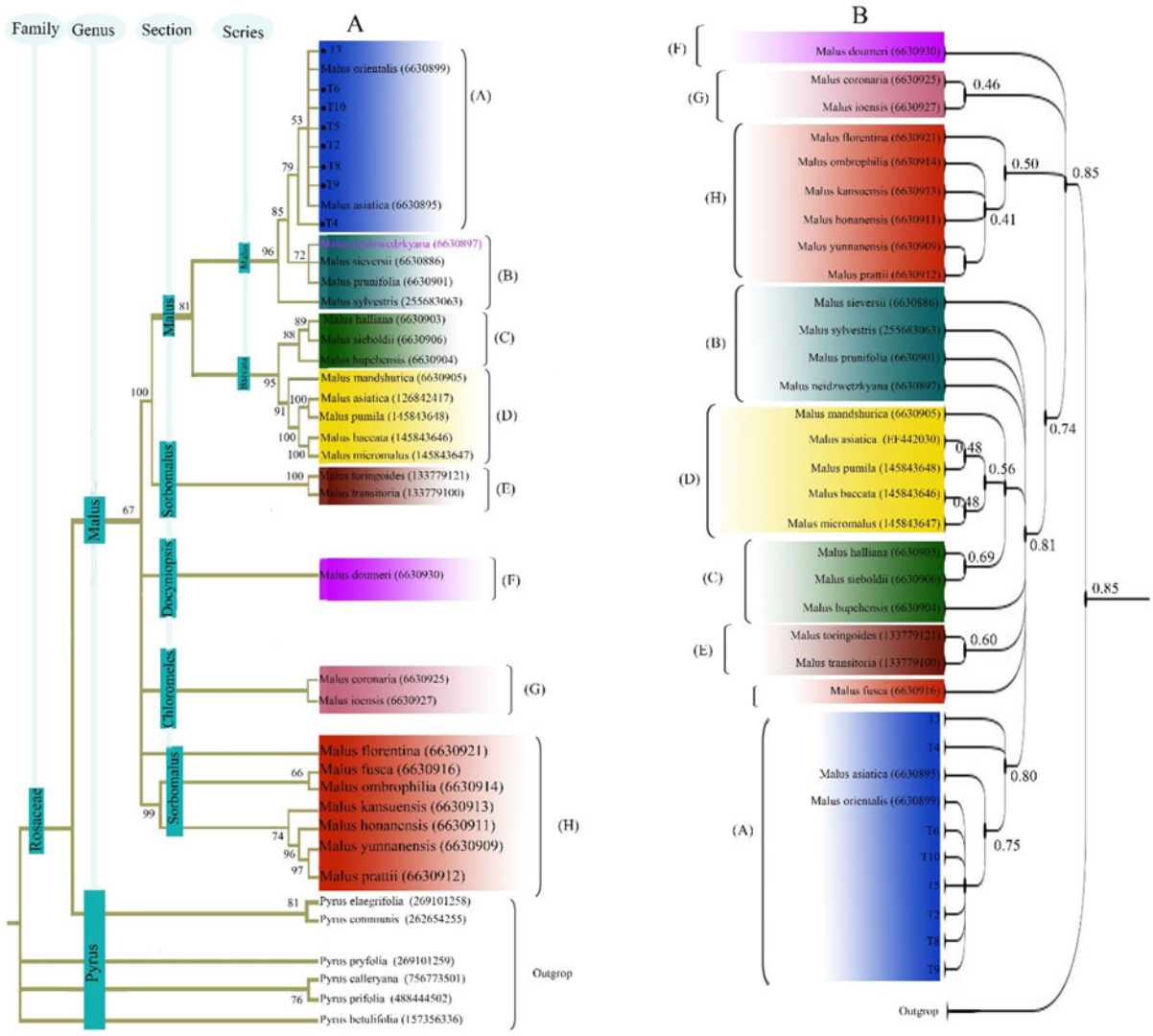


**Fig. 2.** Consensus tree based on Maximum likelihood trees (1) and Bayesian inferences (2) produced with an analysis of ITS2 sequences of this study (circle) together with some species of the genus *Malus* from GenBank and the genus *Pyrus* as outgroup. Bootstrap values are reported on branches if higher than 50%.

value of P (RD) is more than 0.05, the SDP method supports the clade. Accordingly, the SDP method supported all clades diagnosed in the MP tree, but the Bayesian tree did not support the MP result (Fig. 2).

## Network analysis

Based on a network analysis, the Iranian *Malus* samples were located in one group with *M. orientalis*, *M. asiatica*, *M. prunifolia*, *M. sieversii*, *M. sylvestris*, and *M. niedzwetzkyana*. This group, which shares four mutations at positions 13, 85, 117, and 118, was the newest group in this network after *M. toringoides* and *M. transitoria*. Iranian samples and *M. orientalis* were separated from *M. asiatica* by two mutations at positions 593 and 464 (Fig. 3).

## Biogeography of the genus *Malus*

The LRT result ($P = 0.00001$) suggests that relax molecular clock approach is best fitted to explain the divergence time. ILD test result ($P = 0.97$) favours the concatenation of ITS and trnH-psbA datasets.

The biogeographic scenario constructed with the S-DIVA and BBM analyses indicates that the current distribution pattern of *Malus* and *Pyrus* is a result of numerous dispersal and vicariance events. S-DIVA postulates 40 dispersal, 14 vicariance, and one extinction event, whereas the BBM analysis indicates 49 dispersal and 11 vicariance events.

According to the ancestral reconstruction (at node I) by S-DIVA, the ancestors of *Malus* and *Pyrus* originated in China (C) or China + Europe (C + E) during

**Table 3.** Comparison of secondary structure of ITS2 of species of the genus *Malus* obtained from ITS2 database.

| Species | Accession number | E-value | Helix 1 | Helix 2 | Helix 3 | Helix 4 |
|---|---|---|---|---|---|---|
| ***Malus orientalis*** | **6630899** | 1.1e-81 | **100** | **100** | **100** | 100 |
| ***M.sieversii*** | 6630888 | 1.1e-81 | **100** | **100** | **100** | 100 |
| ***M.orientalis*** | 6630894 | 2.9e-80 | 100 | 91 | 100 | 88 |
| ***M.sieversii*** | 6630890 | 3.6e-80 | 100 | 100 | 100 | 88 |
| ***M.niedzwetzkyana*** | 6630897 | 2.9e-80 | 100 | 91 | 100 | 88 |
| ***M.asiatica*** | 6630895 | 1.1e-76 | 100 | 100 | 100 | 88 |
| ***M.fusca*** | 6630915 | 1.1e-76 | 100 | 91 | 100 | 88 |
| ***M.torigoides*** | 133779162 | 2.1e-76 | 100 | 91 | 100 | 88 |
| ***M.transitoria*** | 133779104 | 6.9e-76 | 94 | 91 | 100 | 88 |
| ***M.prunifolia*** | 6630901 | 8.7e-76 | 100 | 91 | 100 | 88 |
| ***M.halliana*** | 6630903 | 1.8e-74 | 100 | 91 | 100 | 80 |
| ***M.angustifolia*** | 6630924 | 2.3e-74 | 94 | 90 | 100 | 80 |
| ***M.coronaria*** | 6630926 | 2.3e-74 | 94 | 90 | 100 | 80 |
| ***M.sieboldii*** | 6630906 | 4.7e-74 | 100 | 91 | 100 | 88 |
| ***M. × domestica*** | 6630885 | 9.3e-80 | 100 | 91 | 100 | 80 |
| ***M. × domestica*** | 6630879 | 1.2e-79 | 100 | 91 | 100 | 100 |
| ***M. × domestica*** | 6630883 | 1.2e-79 | 100 | 91 | 100 | 88 |

**Table 4.** Summary statistics reported by the Species Delimitation plugin for ITS in each putative species.

| Clade | Closest Species | Intra | Intra/Inter | P ID (Strict) | P ID (Liberal) | Av (MRCA-tips) | Rosenberg's P AB | Rodrigo's P(RD) |
|---|---|---|---|---|---|---|---|---|
| A | B | 0.007 | 0.26 | 0.90 (0.83, 0.96) | 0.97 (0.29, 1) | 0.0038 | 0.01 | **1** |
| A | C | 0.023 | 0.34 | 0.86 (0.79, 0.93) | 0.96 (0.91, 1) | 0.0127 | 0.01 | **1** |
| A | D | 0.022 | 0.45 | 0.83 (0.76, 0.89) | 0.95 (0.91, 0.99) | 0.0115 | 0.01 | **<0.05** |
| A | E | 0.022 | 0.06 | 0.77 (0.70, 0.84) | 0.93 (0.98, 0.99) | 0.0115 | 0.01 | **<0.05** |
| B | C | 0.096 | 0.95 | 0.38 (0.72, 0.49) | 0.72 (0.66, 0.79) | 0.397 | 0.01 | **1** |
| B | D | 0.69 | 0.81 | 0.84 (0.38, 0.59) | 0.80 (0.74, 0.87) | 0.397 | 0.01 | **1** |
| B | E | 0.096 | 0.95 | 0.38 (0.72, 0.49) | 0.72 (0.66, 0.79) | 0.397 | 0.01 | **1** |
| C | D | 0.015 | 0.21 | 0.48 (0.33, 0.63) | 0.85 (0.7, 1) | 0.0075 | 0.33 | **<0.05** |
| C | E | 0.015 | 0.26 | 0.46 (0.30, 0.61) | 0.82 (0.67, 0.97) | 0.0075 | 0.33 | **<0.05** |
| E | D | 0.005 | 0.1 | 0.72 (0.55, 0.90) | 0.95 (0.80, 1) | 0.0026 | 1/6E-6 | **<0.05** |

**Intra/Inter** – ratio of Intra (genetic differentiation among members of a putative species) to **Inter** (genetic differentiation between the members of a putative species and the members of the closest putative species), **P ID(Strict)** – mean (95% confidence interval) probability of correctly identifying an unknown member of a given clade using the criterion that it must fall within, but not sister to, the species clade in a tree, **Rosenberg's P AB** – probability of reciprocal monophyly under a random coalescent model and **Rodrigo's P(RD)** – probability that a clade has the observed degree of distinctiveness due to random coalescent processes (Masters et al., 2011).

the Eocene, ~53 Ma (95% HDP: 70.0–38.0), as shown in Fig. 4 and Table 5. The favoured ancestral inference at node III, representing the crown node of *Malus*, is C, with a marginal probability value of 87.72%. The other ancestral area at this node is China + North America (CD) with marginal probability values of 12%. The two possible ancestral ranges (C and CD) at this node probably indicate that China is the ancestral area of *Malus*. The posterior probability value (PP) for this node is 1.00, indicating strong support.

C, CD, and CE are the ancestral ranges at node IV with 69%, 20%, and 8% marginal probability values respectively, and the PP value for this node is 0.93. The ancestral reconstruction at this node and subsequent nodes suggests expansion of *Malus* from China (C) to North America (D), Eastern Asia (F), Europe (E), and West Asia (B). The subsequent nodes also indicate a vicariance event between China (C) and North America (D).

The ancestral range at nodes V and VI is China (C) with PP values of 1.00 and 0.70, respectively. Dispersal events from C to B, D and F are indicated by the biogeographic inferences. The ancestral ranges at node VI indicate a dispersal event from China (C) to Iran (A). Thus, *Malus* entered Iran through China.

Ancestral range at node VII is Iran (A) with 87.50% marginal probability value and PP value for this node is 0.60. The ancestral reconstructions at this and subsequent nodes suggest two dispersal events from Iran (A) to West Asia (B) and China (C).

Most of the dispersal events (23) occurred from China to other areas, as shown in Table 6. North America and Iran also played roles in the dispersal of

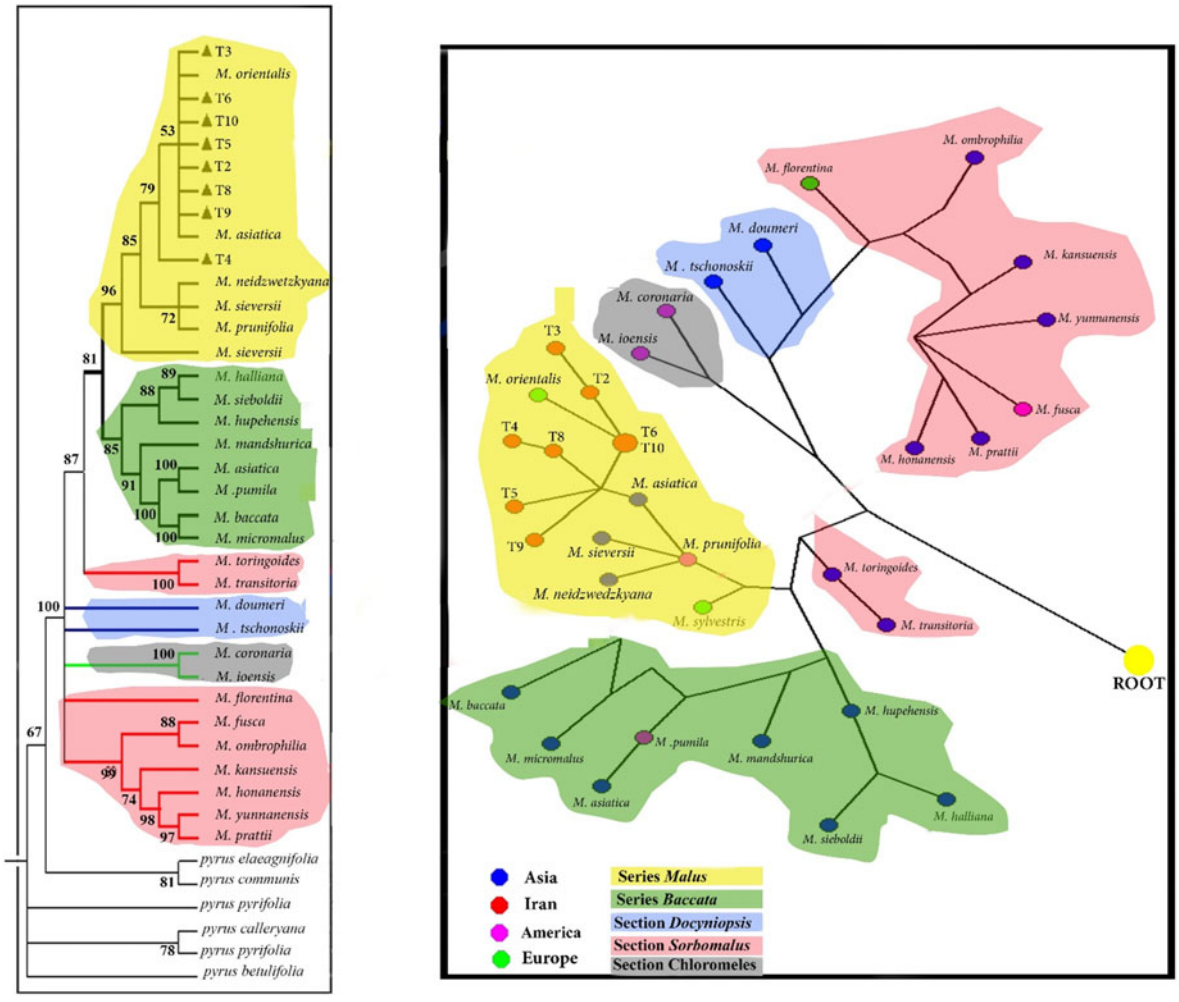


**Fig. 3.** Phylogeny tree Maximum likelihood of *Malus* based on ITS region (left) and network of apples based on ITS sequences (right).

*Malus*, as indicated by eight and two dispersal events that occurred from these areas, respectively.

Bayesian Binary Method (BBM) analysis indicates that the ancestors of *Malus* originated in China, as indicated by the ancestral reconstruction at node III with 94% marginal probability value (Fig. 4). The BBM analysis suggests that dispersal occurred from C to A, B, D, E, and F, as indicated by the ancestral reconstruction at nodes IV, V, VI, and VII. Table 6 indicates that most dispersal events occurred from C, D, and A with 19, 5, and 2 dispersals, respectively.

# Discussion

## Taxonomic status of the *Malus* populations in the Hyrcanian forest

The present study is the first to consider the molecular taxonomic status and phylogenetic relationships of wild *Malus* populations from the Hyrcanian region, one of the most important centres of diversification and evolution of crop plants. Based on the three approaches used in this study to determine the taxonomic status of the Hyrcanian apple populations (ITS phylogenetic tree, comparison of secondary structure of ITS2, and network analysis), the taxa investigated are closely related to *M. orientalis*, *M. asiatica*, and *M. sieversii*. The Nei genetic distance also confirmed the topology of the reconstructed phylogenetic tree. These results correspond to previous reports of the presence of *M. orientalis* and *M. sieversii* in the Iranian flora (Browicz, 1972). Based on the network analysis, all apple taxa distributed in the Caucasian and European regions (*M. orientalis*, *M. sylvestris*, *M. sieversii*, *M. niedzwetzkyana*, and *M. asiatica*) originated from *M. prunifolia*. Additionally, Browicz (1972) reported that *M. sylvestris* is older than *M. orientalis*. Our network analysis also confirmed this proposition. However, the close relationship among

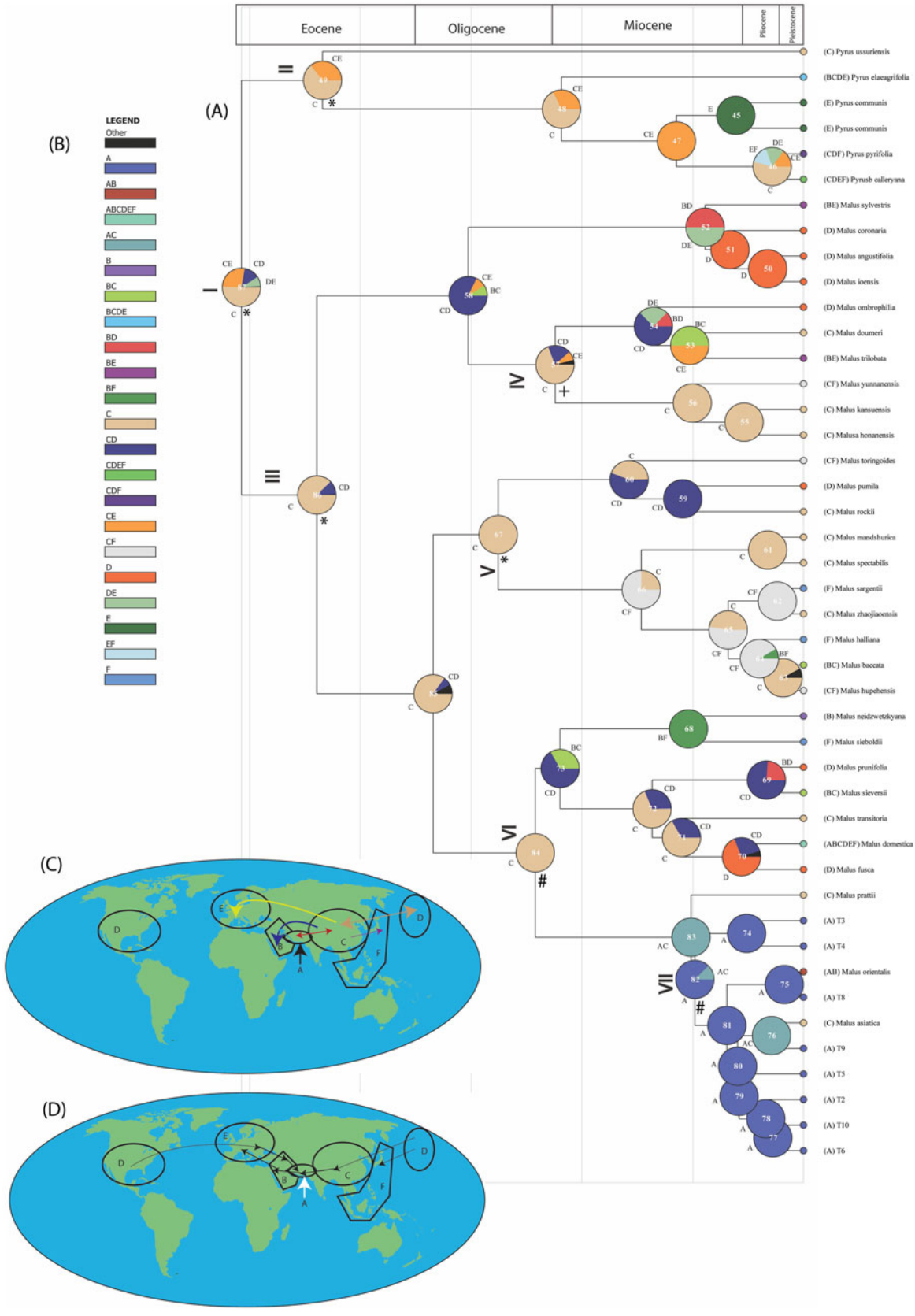
(A)
(B)
(C)
(D)
Eocene
Oligocene
Miocene
Pliocene
Pleistocene
LEGEND
Other
A
AB
ABCDEF
AC
B
BC
BCDE
BD
BE
BF
C
CD
CDEF
CDF
CE
CF
D
DE
E
EF
F
(C) Pyrus ussuriensis
(BCDE) Pyrus elaeagrifolia
(E) Pyrus communis
(E) Pyrus communis
(CDF) Pyrus pyrifolia
(CDEF) Pyrusb calleryana
(BE) Malus sylvestris
(D) Malus coronaria
(D) Malus angustifolia
(D) Malus ioensis
(D) Malus ombrophilia
(C) Malus doumeri
(BE) Malus trilobata
(CF) Malus yunnanensis
(C) Malus kansuensis
(C) Malusa honanensis
(CF) Malus toringoides
(D) Malus pumila
(C) Malus rockii
(C) Malus mandshurica
(C) Malus spectabilis
(F) Malus sargentii
(C) Malus zhaojiaoensis
(F) Malus halliana
(BC) Malus baccata
(CF) Malus hupehensis
(B) Malus neidzwetzkyana
(F) Malus sieboldii
(D) Malus prunifolia
(BC) Malus sieversii
(C) Malus transitoria
(ABCDEF) Malus domestica
(D) Malus fusca
(C) Malus prattii
(A) T3
(A) T4
(AB) Malus orientalis
(A) T8
(C) Malus asiatica
(A) T9
(A) T5
(A) T2
(A) T10
(A) T6

**Table 5.** Age estimation of nodes of divergence time clade based on S-DIVA and BBM.

| | Age estimation (Ma) | | | S-DIVA | | BBM | | |
|---|---|---|---|---|---|---|---|---|
| Nodes | Mean | 95% HPD lower | 95% HPDupper | AR | MP (%) | AR | MP | Support (PP) |
| I | 53.00 | 42.0 | 70.0 | C/CE | 50/27 | C | 95.7 | 1.00 |
| II | 43.44 | 37.0 | 49.8 | C/CE | 64/36 | C | 85.3 | 1.00 |
| III | 44.00 | 27.8 | 60.3 | C | 87.72 | C | 94.0 | 1.00 |
| IV | 22.41 | 14.2 | 35.0 | C/CD | 69/19 | C | 89.1 | 0.93 |
| V | 27.56 | 16.5 | 34.2 | C | 100 | C | 84.0 | 1.00 |
| VI | 21.96 | 24.2 | 38.0 | C | 100 | C | 90.7 | 0.70 |
| VII | 09.90 | 04.5 | 17.9 | A | 87.50 | A | 90.7 | 0.65 |

**Table 6.** Dispersal details of different distribution area based on S-DIVA and BBM.

| S-DIVA | | | | BBM | | | |
|---|---|---|---|---|---|---|---|
| Distribution Range | Dispersal from | Dispersal to | Within | Distribution Range | Dispersal from | Dispersal to | Within |
| A | 2.00 | 2.00 | 8 | A | 2.00 | 2.00 | 8 |
| B | 0.00 | 8.00 | 0 | B | 1.00 | 8.00 | 0 |
| C | 23.00 | 2.00 | 17 | C | 19.00 | 1.00 | 18 |
| D | 8.00 | 9.00 | 4 | D | 5.00 | 6.00 | 5 |
| E | 1.00 | 6.00 | 2 | E | 1.00 | 4.00 | 3 |
| F | 0.00 | 8.00 | 2 | F | 1.00 | 7.00 | 2 |

Caucasian taxa (*M. orientalis*, *M. sylvestris*, *M. sieversii*, *M. niedzwetzkyana*, and *M. asiatica*) based on this research is congruent with Harris, Robinson, and Juniper (2002), which concluded that the morphological diagnosis of these species is extremely difficult.

Based on our network and biogeographic analyses, *M. orientalis* derives from the Iranian populations situated in the western Hyrcanian forest (e.g., Masal and Asalem). More generally, the high number of variable sites in the ITS2 region (180 variable sites) and the high consensus of this DNA region with morphological classifications indicate that this region is a suitable barcode for apple taxonomy. Of course, due to hybridization events (Coart, Van Glabeke, De Loose, Larsen, & Roldán, 2006; Wagner et al., 2014) or incomplete lineage sorting (Micheletti et al., 2011), population genetic approaches using nuclear genetic markers such as microsatellites or SNPs can help reach a more accurate conclusion (Lumley & Sperling, 2011).

## Taxonomic division of the genus *Malus*

The classification of the genus *Malus* proposed by Forsline et al. (2003) and based on morphology is confirmed by our ITS2 phylogenetic tree. Moreover, based on our network analysis results, the two series of section *Malus* (series *Baccata* and *Malus*) were clearly separated as two groups by 10 mutations, despite very common hybridization and introgression among species from these series (Phipps et al., 1990).

The two species *M. sieboldii* and *M. hupehensis* are clearly separated from the other species of section *Malus* by a large number of mutations. This separation is consistent with the distinct nature of *M. sieboldii*, as shown by its large genome (it is the only pentaploid species in section *Malus*). Furthermore, *M. sieboldii* has been described as a hybrid species (Moore & Ballington, 1990), and based on our network results, it is likely that *M. hupehensis* is one of the parents of *M. sieboldii*.

**Fig. 4.** (A) Divergence time estimations and ancestral state reconstructions based on combined data matrix (ITS + *trnH-psbA* regions) estimated by statistical Dispersal–Vicariance analysis (S-DIVA) overlaid onto the maximum clade credibility chronogram from BEAST. Nodes I to VII are discussed in the text. *Indicates above 0.95 posterior probability (PP) values, + indicates above 0.85 PP values and # suggests below 0.85 PP values. (B) The colours in the legend show ancestral area at each node. (C) Dispersal routes from China (C) to different areas are shown on the world map. (D) Dispersal routes from North America to various areas. The areas shown on the world map are (A) Iran, (B) Western Asia (C) China, (D) North America, (E) Europe, and (F) Eastern Asia.

Interestingly, our results do not support the monophyly of section *Sorbomalus* (Phipps et al., 1990) because the two species *M. transitoria* and *M. toringoides* are clearly closer to section *Malus*.

### Biogeography of the genus *Malus*

Biogeographic reconstruction using S-DIVA analysis suggests that the ancestor of *Pyrus* and *Malus* originated in China (C) or China + Europe (C + E) (Fig. 4) with 50%, and 27% marginal probability support, respectively. BBM analysis also suggests China as an ancestral area with 95.7% marginal probability support. The ancestor of *Malus* originated during the Eocene, ~44 Ma (95% HDP: 60.3–27.8).

S-DIVA analysis suggests two main dispersal routes of *Malus*. China (C) is the primary centre of diversity and dispersal and North America is the secondary centre of diversity. Both S-DIVA and BBM analysis indicates 23 and 19 dispersals from China to all other areas respectively. Similarly S-DIVA analysis suggests 8 dispersal events and BBM analysis indicates 5 dispersals from North America. Both these analyses suggest 2 dispersal events from Iran to China and West Asia.

Nine radiations to North America occurred from China during the Oligocene and Miocene, between 30 and 12 Ma. Migrations between the Old and New Worlds in diverse plant and animal groups probably occurred by either the North Atlantic Land Bridge (NALB) or Beringia (Tiffney and Manchester, 2001). Six Transoceanic dispersals occurred from China (C) to Eastern Asia (F), and all these radiations occurred during the Miocene, between 26 Ma and 4 Ma.

The members of *Malus* colonized Iran from China ~24–10 Ma, during the Miocene (Fig. 4). Later, radiations occurred from Iran to China and Western Asia during the Pliocene and Pleistocene respectively.

China played a vital role in the expansion of the distribution range of *Malus* because most of the radiations (23 and 19 as suggested by S-DIVA and BBM analyses, respectively) occurred from China. The radiations from China were multidirectional, as they occurred toward Europe, North America, Eastern Asia and Western Asia, but dispersal to Iran occurred via Western Asia. The radiations from North America to the Old World probably occurred via two routes. One dispersal occurred from North America to China and two dispersals occurred from North America to Europe.

## Conclusions

Northern Iran (Hyrcanian forest) is one of the most diverse areas of forest species such as apple trees, where the special conditions of this habitat have increased the probability of formation of various micro-varieties, and a wide variety of species has been created. The taxonomic status of the Hyrcanian apple populations revealed that the taxa investigated are closely related to *M. orientalis*, *M. asiatica*, and *M. sieversii* and all apple taxa distributed in the Caucasian and European regions originated from *M. prunifolia*. Biogeographic reconstruction suggests that members of *Malus* colonized Iran from China through Western Asia ~22–18 Ma, during the Miocene, and China played a vital role in the expansion of the distribution range of *Malus*. The similarity of molecular taxonomy with traditional classification (morphology) indicates that the ITS2 region is a suitable barcode for apple taxonomy.

## Acknowledgements

We thank Yann Fragniere for helping with figure preparation.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This research was funded by the Iran National Science Foundation [INSF, Grant Number: 97015113] and Tarbiat Modares University.

## Supplemental data

Supplemental data for this article can be accessed here: http://dx.doi.org/10.1080/14772000.2019.1583689.

## References

Ali, S. S., Pfosser, M., Wetschnig, W., Martínez-Azorín, M., Crespo, M. B., & Yu, Y. (2013). Out of Africa: Miocene dispersal, vicariance, and extinction within Hyacinthaceae subfamily Urgineoideae. *Journal of Integrative Plant Biology*, *55*, 950–964.

Ali, S. S., Yu, Y., Pfosser, M., & Wetschnig, W. (2012). Inferences of biogeographical histories within subfamily Hyacinthoideae using S-DIVA and Bayesian binary MCMC analysis implemented in RASP (Reconstruct Ancestral State in Phylogenies). *Annals of Botany*, *109*, 95–107. doi: 10.1093/aob/mcr274.

Bina, H., Yousefzadeh, H., Ali, S. S., & Esmailpour, M. (2016). Phylogenetic relationships, molecular taxonomy, biogeography of Betula, with emphasis on phylogenetic position of Iranian populations. *Tree Genetics & Genomes*, *12*, 84.

Browicz, K. (1972). *Hippocastanaceae. Flora Iranica, (no. 92)*. Graz, Austria: Akademische Druck.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., … Luo, K. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *Public Library of Science One*, *5*, e8613. doi:10.1371/journal.pone.0008613.

Coart, E., Van Glabeke, S., De Loose, M., Larsen, A. S., & Roldán, R. I. (2006). Chloroplast diversity in the genus Malus: New insights into the relationship between the European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.). *Molecular Ecology*, *15*, 2171–2182. doi:10.1111/j.1365-294X.2006.02924.x.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BioMed Central Evolutionary Biology*, *7*, 214. doi:10.1186/1471-2148-7-214.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973. doi:10.1093/molbev/mss075.

Farris, J. S., Kallersjo, M., Kluge, A. G., & Bult, C. (1994). Testing significance of incongruence. *Cladistics*, *10*, 315–319. doi:10.1111/j.1096-0031.1994.tb00181.x.

Fischer, A., & Schmidt, M. (1938). Wilde Kern-und Steinobstarten, ihre Heimat und ihre Bedeutung für die Entstehung der Kultursorten und die Züchtung. *Der Züchter*, *10*, 157–167. doi:10.1007/BF01812347.

Forsline, P., Aldwinckle, H., Dickson, E., Luby, J., & Hokanson, S. (2003). Collection, maintenance, characterization, and utilization of wild apples of Central Asia. *Horticulture Review*, *29*, 1–62. doi:10.1002/9780470650868.ch1.

Forte, A. V., Ignatov, A. N., Ponomarenko, V. V., Dorokhov, D. B., & Savelyev, N. I. (2002). Phylogeny of the *Malus* (apple tree) species, inferred from the morphological traits and molecular DNA analysis. *Russian Journal of Genetics*, *38*, 1150–1161.

Gao, T., Yao, H., Song, J., Zhu, Y., Liu, C., & Chen, S. (2010). Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. *BioMed Central Evolutionary Biology*, *10*, 324.

Gharghani, A., Zamani, Z., Talaie, A., Oraguzie, N. C., Fatahi, R., Hajnajari, H., … Gardiner, S. E. (2009). Genetic identity and relationships of Iranian apple (*Malus*× *domestica* Borkh.) cultivars and landraces, wild Malus species and representative old apple cultivars based on simple sequence repeat (SSR) marker analysis. *Genetic Resources and Crop Evolution*, *56*, 829–842. doi:10.1007/s10722-008-9404-0.

Group, C. P. W., Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., & Fazekas, A. J. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, *106*, 12794–12797.

Harris, S. A., Robinson, J. P., & Juniper, B. E. (2002). Genetic clues to the origin of the apple. *Trends in Genetics: TIG*, *18*, 426–430.

Hebert, P. D., Cywinska, A., & Ball, S. L. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*, 313–321. doi:10.1098/rspb.2002.2218.

Janfaza, S. (2016). *Biogeography, taxonomi status and genetic diversity of Castanea sativa in Iran*. Mazandaran: Sari Agricultural Sciences and Natural Resources University.

Janick, J. (Ed.). (2003). *Horticultural reviews: Wild apple and fruit trees of central Asia* (Vol. 31). Hoboken, NJ: John Wiley and Sons.

Katayama, H., & Uematsu, C. (2003). Comparative analysis of chloroplast DNA in *Pyrus* species: Physical map and gene localization. *Theoretical and Applied Genetics*, *106*, 303–310. doi:10.1007/s00122-002-1003-4.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*, 111–120. doi:10.1007/BF01731581.

Koetschan, C., Förster, F., Keller, A., Schleicher, T., Ruderisch, B., Schwarz, R., … Schultz, J. (2009). The ITS2 database III—sequences and structures for phylogeny. *Nucleic Acids Research*, *38*, 275–279.

Liu, Z., Zeng, X., Yang, D., Chu, G., Yuan, Z., & Chen, S. (2012). Applying DNA barcodes for identification of plant species in the family Araliaceae. *Gene*, *499*, 76–80. doi:10.1016/j.gene.2012.02.016.

Lo Presti, R. M., & Oberprieler, C. (2009). Evolutionary history, biogeography and eco-climatological differentiation of the genus *Anthemis* L. (Compositae, Anthemideae) in the circum-Mediterranean area. *Journal of Biogeography*, *36*, 1313–1332. doi:10.1111/j.1365-2699.2009.02121.x.

Lumley, L. M., & Sperling, F. A. (2011). Utility of microsatellites and mitochondrial DNA for species delimitation in the spruce budworm (*Choristoneura fumiferana*) species complex (Lepidoptera: Tortricidae). *Molecular Phylogenetics and Evolution*, *58*, 232–243. doi:10.1016/j.ympev.2010.11.023.

Mansfeld, R., & Büttner, R. (2001). *Mansfeld's encyclopedia of agricultural and horticultural crops: (except ornamentals)* (Vol. 1). Germany: Springer Science & Business Media.

Masters, B. C., Fan, V., & Ross, H. A. (2011). Species delimitation–a geneious plugin for the exploration of species boundaries. *Molecular Ecology Resources*, *11*, 154–157. doi:10.1111/j.1755-0998.2010.02896.x.

Micheletti, D., Troggio, M., Salamini, F., Viola, R., Velasco, R., & Salvi, S. (2011). On the evolutionary history of the domesticated apple. *Nature Genetics*, *43*, 1043–1044. doi:10.1038/ng.983.

Moore, J. N., & Ballington, J. R. (1990). *Genetic resources of temperate fruit and nut crops* (Vol. 1). Wageningen, The Netherlands: International Society for Horticultural Science.

Murray, M. G., & Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, *8*, 4321–4326. doi:10.1093/nar/8.19.4321.

Pang, X., Song, J., Zhu, Y., Xu, H., Huang, L., & Chen, S. (2011). Applying plant DNA barcodes for Rosaceae species identification. *Cladistics*, *27*, 165–170. doi:10.1111/j.1096-0031.2010.00328.x.

Phipps, J. B., Robertson, K. R., Smith, P. G., & Rohrer, J. R. (1990). A checklist of the subfamily Maloideae (Rosaceae). *Canadian Journal of Botany*, *68*, 2209–2269. doi:10.1139/b90-288.

Prévot, V., Jordaens, K., Sonet, G., & Backeljau, T. (2013). Exploring species level taxonomy and species delimitation methods in the facultatively self-fertilizing land snail genus *Rumina* (Gastropoda: Pulmonata). *Public Library of Science One*, *8*, e60736. doi:10.1371/journal.pone.0060736.

Rambaut, A. (2012). FigTree v. 1.4. Molecular evolution, phylogenetics and epidemiology. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/ (aaccessed 6 February 2019).

Rechinger, K. H. (1964). *Flora Iranica*. Graz, Austria: Akademische Druck-und Verlagsanstalt.

Robinson, J. P., Harris, S. A., & Juniper, B. E. (2001). Taxonomy of the genus Malus Mill. (Rosaceae) with emphasis on the cultivated apple, Malus domestica Borkh. *Plant Systematics and Evolution*, *226*(1–2), 35–58.

Sabeti, H. (1976). *Forests, trees and shrubs of Iran*. Tehran, Iran: Ministry of Agriculture and Natural Resources.

Sang, T., Crawford, D. J., & Stuessy, T. F. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *American Journal of Botany*, *84*, 1120–1136. doi:10.2307/2446155.

Savelyeva, E. N., Boris, K. V., Kochieva, E. Z., & Kudryavtsev, A. M. (2013). Analysis of sequences of ITS1 internal transcribed spacer and 5.8 S ribosome gene of Malus species. *Russian Journal of Genetics*, *49*, 1175–1182. doi:10.1134/S1022795413110148.

Schultz, J., Maisel, S., Gerlach, D., Müller, T., & Wolf, M. (2005). A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA*, *11*, 361–364. doi:10.1261/rna.7204505.

Schultz, J., Müller, T., Achtziger, M., Seibel, P. N., Dandekar, T., & Wolf, M. (2006). The internal transcribed spacer 2 database: A web server for (not only) low level phylogenetic analyses. *Nucleic Acids Research*, *34*, 704–707.

Selig, C., Wolf, M., Müller, T., Dandekar, T., & Schultz, J. (2007). The ITS2 Database II: Homology modelling RNA structure for molecular systematics. *Nucleic Acids Research*, *36*, 377–380.

Semerikov, V. L., Semerikova, S. A., Polezhaeva, M. A., Kosintsev, P. A., & Lascoux, M. (2013). Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): A range-wide analysis of cytoplasmic markers. *Molecular Ecology*, *22*, 4958–4971. doi:10.1111/mec.12433.

Swofford, D. L. (2002). *PAUP: Phylogenetic analysis using parsimony (and other methods), V. 4.0. b10*. Sunderland, Massachusetts: Sinauer Associates.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*, e14–e14.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. doi:10.1093/molbev/mst197.

Tate, J. A., & Simpson, B. B. (2003). Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany*, *28*(4), 723–737.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., & Vogler, A. P. (2003). A plea for DNA taxonomy. *Trends in Ecology and Evolution*, *18*, 70–74. doi:10.1016/S0169-5347(02)00041-1.

Tiffney, B. H., & Manchester, S. R. (2001). The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences*, *162*(S6), S3–S17.

Vavilov, N. I. (1930). Wild progenitors of the fruit trees of Turkistan and the Caucasus and the problem of the origin of fruit trees. *Report and Proceeding 9th Internationl Horticulture Congress*, 271–86.

von Cräutlein, M., Korpelainen, H., Pietiläinen, M., & Rikkinen, J. (2011). DNA barcoding: A tool for improved taxon identification and detection of species diversity. *Biodiversity and Conservation*, *20*, 373–389. doi:10.1007/s10531-010-9964-0.

Wagner, I., Maurer, W. D., Lemmen, P., Schmitt, H. P., Wagner, M., Binder, M., & Patzak, P. (2014). Hybridization and genetic diversity in Wild Apple (*Malus sylvestris* (L.) Mill.) from various regions in Germany and from Luxembourg. *Silvae Genetica*, *63*, 81–93. doi:10.1515/sg-2014-0012.

White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. L. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR protocols: A guide to methods and applications*, (pp. 315–322). New York: Academic Press, Inc.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., … Chen, S. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *Public Library of Science One*, *5*, e13102. doi:10.1371/journal.pone.0013102.

Yousefzadeh, H., Colagar, A. H., Tabari, M., Sattarian, A., & Assadi, M. (2012). Utility of ITS region sequence and structure for molecular identification of *Tilia* species from Hyrcanian forests, Iran. *Plant Systematics and Evolution*, *298*, 947–961. doi:10.1007/s00606-012-0604-x.

Yu, Y., Harris, A. J., Blair, C., & He, X. (2015). RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular Phylogenetics and Evolution*, *87*, 46–49. doi:10.1016/j.ympev.2015.03.008.

Zhukovsky, P. M. (1965). Main gene centres of cultivated plants and their wild relatives within the territory of the USSR. *Euphytica*, *14*, 177–188. doi:10.1007/BF00038985.

**Associate Editor: Nadia Bystriakova**