



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Genetic analysis of walnut cultivars from southwest China

Citation for published version:

Wambulwa, MC, Fan, P-Z, Milne, R, Wu, Z-Y, Luo, Y-H, Wang, Y-H, Wang, H, Gao, L-M, Xiahou, Z-Y, Jin, Y-C, Ye, L-J, Xu, Z-C, Yang, Z, Li, D-Z & Liu, J 2022, 'Genetic analysis of walnut cultivars from southwest China: Implications for germplasm improvement', *Plant Diversity*, vol. 44, no. 6, pp. 530-541.
<https://doi.org/10.1016/j.pld.2021.08.005>

Digital Object Identifier (DOI):

[10.1016/j.pld.2021.08.005](https://doi.org/10.1016/j.pld.2021.08.005)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Plant Diversity

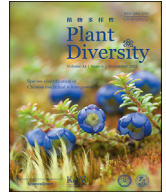
General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Research paper

Genetic analysis of walnut cultivars from southwest China: Implications for germplasm improvement



Moses C. Wambulwa ^{a, b, c, 1}, Peng-Zhen Fan ^{a, b, d, 1}, Richard Milne ^e, Zeng-Yuan Wu ^b, Ya-Huang Luo ^{a, b}, Yue-Hua Wang ^f, Hong Wang ^a, Lian-Ming Gao ^a, Zuo-Ying Xiahou ^b, Ye-Chuan Jin ^f, Lin-Jiang Ye ^a, Zu-Chang Xu ^b, Zhi-Chun Yang ^g, De-Zhu Li ^{a, b, h, *}, Jie Liu ^{a, b, **}

^a CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, Yunnan, China

^b Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, Yunnan, China

^c Department of Life Sciences, South Eastern Kenya University, 170-90200, Kitui, Kenya

^d University of the Chinese Academy of Sciences, Beijing, 100049, China

^e Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

^f School of Ecology and Environmental Science, Yunnan University, Kunming, 650091, Yunnan, China

^g Yangbi Forestry and Grassland Administration, Dali, 672500, Yunnan, China

^h College of Life Sciences, University of Chinese Academy of Sciences, Kunming, 650201, Yunnan, China

ARTICLE INFO

Article history:

Received 17 June 2021

Received in revised form

26 July 2021

Accepted 18 August 2021

Available online 2 September 2021

Keywords:

Cultivars

Genetic diversity

Juglans sigillata

Southwest China

Walnut

ABSTRACT

Walnuts are highly valued for their rich nutritional profile and wide medicinal applications. This demand has led to the intensification of breeding activities in major walnut production areas such as southwest China, in order to develop more superior cultivars. With the increasing number of cultivars, accurate identification becomes fundamental to selecting the right cultivar for grafting, industrial processing or development of new cultivars. To ensure proper identification of cultivars and understand the genetic structure of wild and cultivated material, we genotyped 362 cultivated and wild individuals of walnut trees from southwest China (with two additional populations from Xinjiang, plus three cultivars from Canada, France and Belgium) using 36 polymorphic microsatellite loci. We found relatively low indices of genetic diversity ($H_O = 0.570$, $H_E = 0.404$, $N_A = 2.345$) as well as a high level of clonality (>85% of cultivars), indicating reliance on genetically narrow sources of parental material for breeding. Our STRUCTURE and PCoA analyses generally delineated the two species, though considerable levels of introgression were also evident. More significantly, we detected a distinct genetic group of cultivated *Juglans sigillata*, which mainly comprised individuals of the popular 'Yangbidapao' landrace. Finally, a core set of 18 SSR loci was selected, which was capable of identifying 32 cultivars. In a nutshell, our results call for more utilization of genetically disparate material, including wild walnut trees, as parental sources to breed for more cultivars. The data reported herein will significantly contribute towards the genetic improvement and conservation of the walnut germplasm in southwest China.

Copyright © 2021 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author. Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, Yunnan, China

** Corresponding author. CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, Yunnan, China.

E-mail addresses: dzl@mail.kib.ac.cn (D.-Z. Li), liujie@mail.kib.ac.cn (J. Liu).

Peer review under responsibility of Editorial Office of Plant Diversity.

¹ Authors with equal contribution.

1. Introduction

Walnut trees are among the most popular nut crops, famous for their nutritional and medicinal use, as well as for their high-quality timber (Chen et al., 2019; Gupta et al., 2019). The broad biochemical profile of walnuts (Ros et al., 2018) has been associated with reduced risk of cardiovascular, respiratory and cancer-related complications (Aune et al., 2016), as well as antimicrobial activity

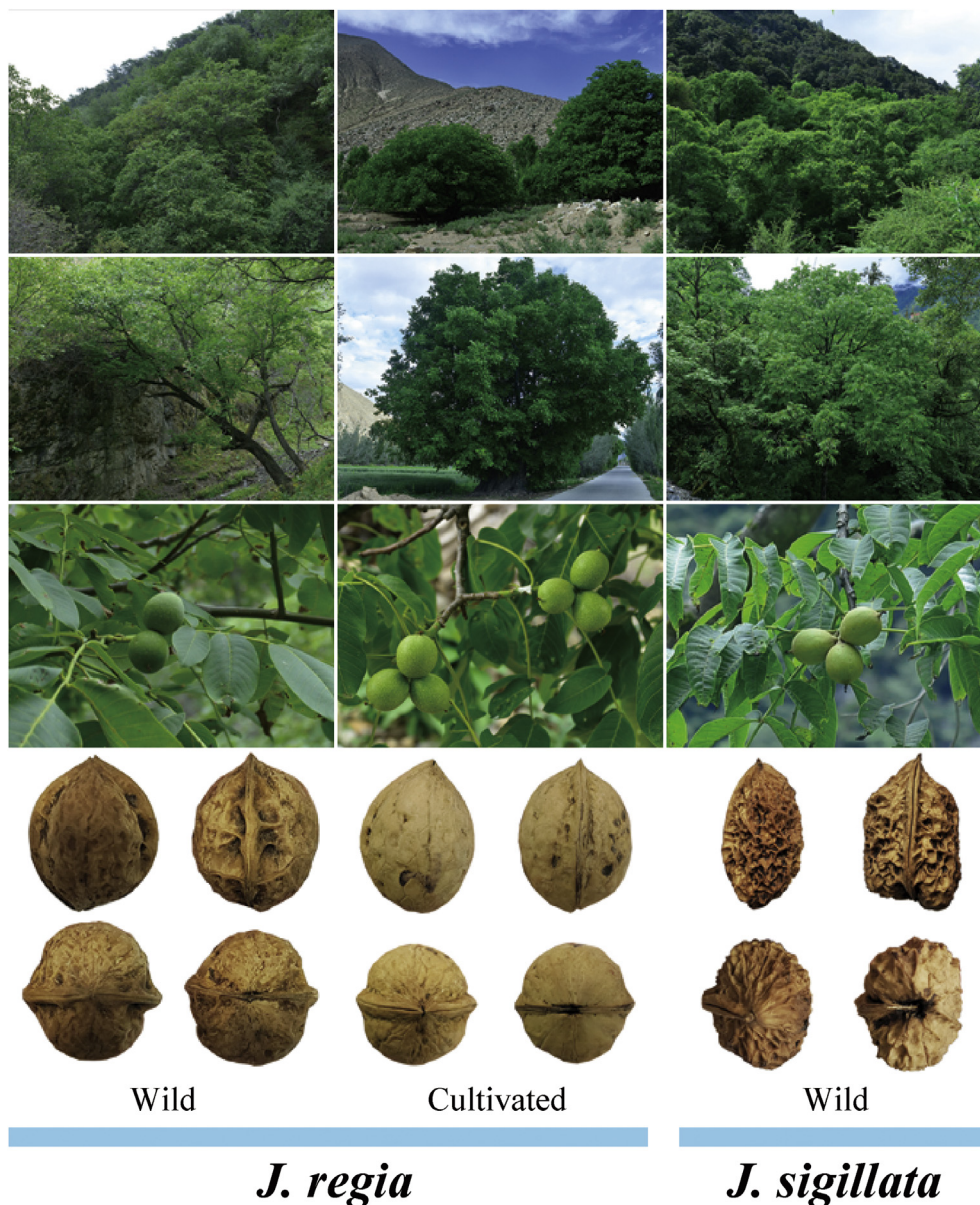


Fig. 1. Habitats, fruit and nut morphology of the wild and cultivated walnuts. In general, *Juglans regia* nuts have wrinkled nut surfaces, whereas those of *J. sigillata* have deep pits and seal-like depressions on their surfaces.

(Gyawali and Ibrahim, 2014). These dietary and therapeutic benefits have increased the global demand for walnuts, and in 2018, the global walnut production totalled ~3,662,507 metric tonnes, of which China produced ~43%, i.e. 1,586,367 metric tonnes (FAOSTAT, 2020). Within China, the southwestern region is the most important centers of walnut production, and has been shown to harbor high levels of genetic diversity (Feng et al., 2018). As a result, numerous walnut cultivars have been developed in the region, most of which are propagated through grafting.

Walnut trees belong to the genus *Juglans* (Juglandaceae), which comprises 21 diploid species ($2n = 32$) distributed across tropical and temperate zones in America, Eurasia, and the Caribbean (Freeman and Reveal, 2005; Lu et al., 1999). The genus is classified into four or three sections based both on phylogeny and morphology (Aradhya et al., 2007; Dong et al., 2017; Manning, 1978). Section *Juglans* comprises the two most economically important and widely cultivated species of walnuts, i.e. *J. regia* L.

and *J. sigillata* Dode, which are native to the mountains of central Asia (McGranahan and Leslie, 2009), and southwest China (Lu et al., 1999), respectively. In spite of the clear morphological differences between the two species (Fig. 1), some previous studies have demonstrated low genetic differentiation between *J. regia* and *J. sigillata* (Dang et al., 2019; Yuan et al., 2018), and natural hybrids are found where their ranges overlap in Yunnan Province and Xizang Autonomous Region (Dang et al., 2019; Wang et al., 2015).

Although China accounts for 43% of the global walnut production (FAOSTAT, 2020), insufficient knowledge of the identity and genetic diversity of cultivars may impede future improvement efforts in the region. Moreover, different crop cultivars may be suited for different uses due to their variance in mechanical and chemical profiles, as previously demonstrated for walnut (Gharibzadeh et al., 2012) and Spanish almond cultivars (Rabadán et al., 2017). Therefore, identification and systematic analysis of walnut cultivars in southwest China may be a critical step towards efficiency in the

country's walnut processing industry. Walnut cultivars may also vary substantially in composition of nutrients such as carotenoids, fatty acids, sterols, tocopherols, volatile compounds and mineral elements such as potassium, magnesium, calcium and iron, and also in rancidity (tendency to spoil) (Abdallah et al., 2015; Bouabdallah et al., 2014; Liu et al., 2020; Martínez et al., 2006; Ojeda-Amador et al., 2018); in all of these studies, the variation was attributable to genetic variation among the cultivars, but environmental factors might also play a role (Rabadán et al., 2018). Consequently, accurate identification of walnut cultivars is important for selecting cultivars for specific uses in the walnut processing industry, including the development of new and better ones.

Walnut improvement programs worldwide have relied on hybridization to develop superior commercial varieties. For instance, Hassani et al. (2019) recently reviewed the history and current status of the improvement programs of *J. regia* in Iran, and highlighted the importance of intraspecific hybridization. In southwest China, however, co-cultivation of *J. regia* and *J. sigillata* may have led to emergence of both natural and artificial interspecific hybrids, which have been shown to exhibit relatively more superior traits than their parent species (Li, 2017). As a consequence, there has been a drastic increase in the number of hybrid accessions in China in the recent decades (Tian et al., 2009; Wu et al., 2009), with the potential consequence of misidentification and unclear pedigree records. Moreover, while morphological characters have often been useful in delineating some of these hybrids, genetic admixture among various cultivars may lead to overlap of phenotypic characteristics, hence morphological characteristics alone cannot be used to identify the interspecific hybrids. Additionally, commercial processing of walnut products, e.g., packaging as kernels or powder, often interferes with or removes diagnostic characters necessary for use in traditional plant taxonomy. Molecular markers are more reliable than morphological features, less prone to environmental plasticity, and can survive processing. Neutral genetic markers such as microsatellites have been widely applied in studying the genetic relationships among walnut germplasm and proven to be effective (Balapanov et al., 2019; Bernard et al., 2018; Dangl, 2005; Lorenzo et al., 2020; Torokeldiev et al., 2018; Yuan et al., 2018). Microsatellites have also shown considerable transferability across species of *Juglans* as well as high ubiquity in the genomes (Ebrahimi et al., 2019). Moreover, the recent development of 32 novel SSR markers for *J. sigillata* (Xu et al., 2020) provides an opportunity for accurate cultivar identification.

In the present study, we applied nuclear microsatellite markers to genotype walnut cultivars and wild trees from Yunnan Province and Xinjiang Autonomous Region of China, as well as material from Europe (Belgium and France) and North America (Canada). In the present study, we aimed to 1) assess the level of genetic variation and relationships across wild and cultivated walnuts in southwest China, 2) determine the extent of introgression of wild walnuts into the cultivated gene pool and, and 3) establish a microsatellite database for fingerprinting and identification of walnut cultivars and their hybrids in southwest China. Results of our study will rationalize the process of material selection during breeding, and promote effective utilization of walnut germplasm resources in the region.

2. Materials and methods

2.1. Sampling

A total of 362 walnut trees were collected from various locations in Yunnan Province and Xinjiang Autonomous Region of China (Table 1). Of these, 242 comprised 47 cultivars, of which 40 were *J. sigillata*, three were *J. regia* and four were putative *J. sigillata* × *J.*

regia hybrids (Table 1). The 40 *J. sigillata* cultivars included individuals of the landrace 'Yangbidapao' DP, which is popularly cultivated in Yangbi and neighboring Counties of Yunnan (e.g. Channing County). The cultivars were collected from Yangbi Walnut Germplasm Bank as well as from local orchards in the county. Though some of the cultivars were developed from breeding programs, information on their genetic pedigrees was lacking. Thus, our cultivar collection may possibly comprise clones, full, and half siblings in varying but unknown proportions. Cultivar identification was based the records at the germplasm bank, but also validated based on Pei and Lu (2011), who basically relied on morphological characters to distinguish the cultivars (Fig. 1). The collection also included three introduced cultivars of *J. regia* from Belgium, Canada and France (cultivars 120, X116 and A119, respectively). The remaining 120 individuals were from six populations, comprising four *J. sigillata* (BHLT, Liuj01, Liuj02 and TYC) and 2 *J. regia* populations (GLR and XHT). Populations BHLT, TYC and GLR were confirmed to be wild, while Liuj01, Liuj02 and XHT were populations under cultivation (Table 1). Based on the information gathered in the field, Liuj01, Liuj02 are populations of the landrace DP. For each cultivar, three to six individual trees were sampled, while the six populations had sample sizes ranging from 16 to 25 individuals (Table 1). All leaf samples intended for DNA extraction were collected and dried in silica gel. Voucher specimens were deposited at the herbarium of Kunming Institute of Botany, Chinese Academy of Sciences (KUN).

2.2. DNA extraction and SSR genotyping

Total genomic DNA was extracted from about 0.02 g of silica-dried leaf tissue based on a modified CTAB protocol (Doyle and Doyle, 1987; Liu and Gao, 2011). The concentration of all DNA samples was adjusted to 30–50 ng/μL prior to PCR procedures. A total of 36 nuclear microsatellite primer pairs were selected from literature (Xu et al., 2020) and screened for genotyping the 362 walnut trees (Table S1). All forward primers were fluorescently labelled with FAM, HEX or TAMRA dyes (Optimus Bio, Kunming, China) at the 5' end. A combination of size and colour multiplexing allowed the 36 primer pairs to be grouped into 5 multiplexes (Table S1). PCR amplification was carried out on a Veriti® 96-Well Thermo-Cycler (Applied Biosystems, Foster City, California, USA) according to Xu et al. (2020). Briefly, the 15 μL multiplex PCR comprised 0.2 μL of each primer, and 1 μL of the DNA template then the volume made up with an appropriate amount of Golden Star T6 Super PCR Mix (TsingKe Biological Technology, Beijing, China), depending on the number of primer pairs used in the mix. The following cycling conditions were used: initial denaturation at 95 °C for 5 min, 30 cycles 95 °C for 10 s, primer annealing temperature (53–61 °C; Table S1) for 30 s, 72 °C for 1 min, then a final extension at 72 °C for 5 min. Annealing temperatures for each of the five multiplexed mixes were optimized accordingly. Following successful PCR, SSR fragment sizes for each individual were determined using an ABI 3730xl automated sequencer (Applied Biosystems, Foster City, California, USA).

2.3. Data analysis

GENEMARKER v4.0 (SoftGenetics, State College, Pennsylvania, USA) was used to visualize the SSR data as diploid genotypes. For convenience in our population genetic analyses, each of the 47 cultivars examined was treated as a 'population' in the initial analysis, with all members of that cultivar assigned to it; to these were added the six wild and cultivated populations making 53 'populations' in total. Genetic diversity analyses (allele number, N_A ; observed heterozygosity, H_O ; expected heterozygosity, H_E ; and

Table 1
Details of sample location, sample size, and summary statistics of genetic diversity for 47 walnut cultivars and 6 populations averaged across 36 SSR loci.

| Population code | Cultivar name | Species | Provenance (County, Province) | N | N _A | H _O | H _E | F |
|-----------------|------------------|---------------------|-------------------------------|-----|----------------|----------------|----------------|--------|
| 114 | Like 1 | <i>J. sigillata</i> | Yulong, YN | 5 | 2.417 | 0.639 | 0.416 | -0.477 |
| 118 | Wumeng 19 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.722 | 0.489 | 0.422 | -0.145 |
| 119 | Wumeng 1 | <i>J. sigillata</i> | Yiliang, YN | 5 | 3.222 | 0.617 | 0.535 | -0.160 |
| 121 | Baohe 7 | <i>J. sigillata</i> | Tengchong, YN | 5 | 2.361 | 0.546 | 0.390 | -0.310 |
| 122 | Lushui 1 | <i>J. sigillata</i> | Lushui, YN | 5 | 1.722 | 0.722 | 0.361 | -1.000 |
| 123 | Li 53 | <i>J. sigillata</i> | Yulong, YN | 5 | 2.889 | 0.518 | 0.470 | -0.107 |
| 124 | Ludiandapao 3 | <i>J. sigillata</i> | Ludian, YN | 5 | 1.583 | 0.556 | 0.280 | -0.986 |
| 125 | Yangza 2 | <i>J. sigillata</i> | Yiangbi, YN | 6 | 2.333 | 0.573 | 0.371 | -0.461 |
| 127 | Yangza 1 | <i>J. sigillata</i> | Yiangbi, YN | 6 | 1.639 | 0.551 | 0.289 | -0.893 |
| 128 | Xixianghetao | <i>J. sigillata</i> | Changning, YN | 5 | 1.833 | 0.494 | 0.296 | -0.611 |
| 129 | Yong 11 | <i>J. sigillata</i> | Yongsheng, YN | 4 | 2.222 | 0.556 | 0.391 | -0.359 |
| 130 | Yangza 3 | <i>J. sigillata</i> | Yiangbi, YN | 5 | 2.333 | 0.615 | 0.410 | -0.460 |
| 131 | Yangjiang 1 | <i>J. sigillata</i> | Yiangbi, YN | 5 | 1.778 | 0.644 | 0.333 | -0.923 |
| 132 | Baohe 3 | <i>J. sigillata</i> | Changning, YN | 5 | 2.167 | 0.583 | 0.379 | -0.468 |
| 135 | Huaningdabaikae | <i>J. sigillata</i> | Huaning, YN | 5 | 2.028 | 0.489 | 0.330 | -0.426 |
| 136 | Ludiandama 1 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.000 | 0.572 | 0.355 | -0.525 |
| 138 | Wei 2 | <i>J. sigillata</i> | Weixi, YN | 5 | 2.111 | 0.672 | 0.403 | -0.648 |
| 139 | Baohe 2 | <i>J. sigillata</i> | Changning, YN | 5 | 1.722 | 0.517 | 0.291 | -0.784 |
| 140 | Santai | <i>J. sigillata</i> | Dayao, YN | 5 | 2.722 | 0.558 | 0.475 | -0.189 |
| 141 | Huaningdashake | <i>J. sigillata</i> | Huaning, YN | 5 | 3.167 | 0.567 | 0.531 | -0.091 |
| 143 | Baohe 5 | <i>J. sigillata</i> | Tengchong, YN | 5 | 2.222 | 0.472 | 0.372 | -0.247 |
| 146 | Yunwanshuang 1 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.167 | 0.656 | 0.399 | -0.607 |
| 149 | Hongpilianchuan | <i>J. sigillata</i> | Huize, YN | 5 | 2.889 | 0.583 | 0.533 | -0.067 |
| 150 | Honghe 1 | <i>J. sigillata</i> | Mile, YN | 5 | 2.444 | 0.494 | 0.403 | -0.172 |
| 151 | Honghe 2 | <i>J. sigillata</i> | Jianshui, YN | 5 | 2.639 | 0.617 | 0.445 | -0.345 |
| 152 | Yunwanshuang 2 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.167 | 0.572 | 0.388 | -0.356 |
| 379 | Yangxin 1 | <i>J. sigillata</i> | Yangbi, YN | 4 | 2.306 | 0.583 | 0.436 | -0.287 |
| DP | Yangbidapao | <i>J. sigillata</i> | Yangbi, YN | 5 | 1.806 | 0.706 | 0.367 | -0.932 |
| H138 | Hushan 1 | <i>J. sigillata</i> | nd | 3 | 1.944 | 0.528 | 0.375 | -0.382 |
| L144 | Ludiandama 2 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.639 | 0.561 | 0.451 | -0.236 |
| L149 | Li 20 | <i>J. sigillata</i> | Yulong, YN | 5 | 2.944 | 0.611 | 0.536 | -0.116 |
| NQ | Niangqing | <i>J. sigillata</i> | Yangbi, YN | 5 | 1.778 | 0.600 | 0.327 | -0.813 |
| S115 | Like 2 | <i>J. sigillata</i> | Yulong, YN | 5 | 2.278 | 0.506 | 0.371 | -0.252 |
| S142 | Qingfeng 2 | <i>J. sigillata</i> | Ludian, YN | 5 | 1.833 | 0.656 | 0.348 | -0.912 |
| W116 | Wumeng 3 | <i>J. sigillata</i> | Yiliang, YN | 5 | 2.444 | 0.644 | 0.487 | -0.311 |
| X142 | Qingfeng 3 | <i>J. sigillata</i> | Zhaoyangqu, YN | 5 | 2.389 | 0.617 | 0.405 | -0.444 |
| X145 | Wumeng 16 | <i>J. sigillata</i> | Ludian, YN | 5 | 2.528 | 0.522 | 0.482 | -0.034 |
| X147 | Qingfeng 1 | <i>J. sigillata</i> | Ludian, YN | 5 | 3.000 | 0.672 | 0.507 | -0.285 |
| XP | Xiaopaohetao | <i>J. sigillata</i> | Yangbi, YN | 5 | 1.667 | 0.633 | 0.328 | -0.903 |
| YZX01 | Yangzaoxiang | <i>J. sigillata</i> | Yangbi, YN | 5 | 2.306 | 0.522 | 0.378 | -0.308 |
| 120 | – | <i>J. regia</i> | Belgium | 5 | 2.194 | 0.372 | 0.304 | -0.190 |
| A119 | na | <i>J. regia</i> | France | 5 | 2.861 | 0.494 | 0.450 | -0.060 |
| X116 | na | <i>J. regia</i> | Canada | 5 | 2.472 | 0.343 | 0.361 | 0.073 |
| 133 | Yunxin 306 | Hybrid | – | 5 | 1.889 | 0.744 | 0.403 | -0.834 |
| 134 | Yunxingaoyuan 14 | Hybrid | – | 5 | 2.444 | 0.706 | 0.447 | -0.549 |
| 137 | Yunxingaoyuan 34 | Hybrid | – | 5 | 2.722 | 0.644 | 0.529 | -0.207 |
| 144 | Yunxin 303 | Hybrid | – | 4 | 2.722 | 0.611 | 0.497 | -0.238 |
| BHLT | Wild population | <i>J. sigillata</i> | Longyang, YN | 19 | 3.667 | 0.463 | 0.499 | 0.082 |
| Liuj01 | – | <i>J. sigillata</i> | Changning, YN | 25 | 1.778 | 0.723 | 0.363 | -0.961 |
| liuj02 | – | <i>J. sigillata</i> | Yangbi, YN | 25 | 1.889 | 0.721 | 0.368 | -0.842 |
| TYC | Wild population | <i>J. sigillata</i> | Changning, YN | 20 | 3.611 | 0.455 | 0.470 | 0.034 |
| GLR | Wild population | <i>J. regia</i> | Gongliu, XJ | 25 | 2.000 | 0.279 | 0.272 | -0.020 |
| XHT | – | <i>J. regia</i> | Hetian, XJ | 16 | 2.667 | 0.432 | 0.406 | -0.049 |
| Total | | | | 198 | 2.345 | 0.570 | 0.404 | -0.376 |

Note: N, Sample size; N_A, Number of alleles; H_O, Observed heterozygosity; H_E, Expected heterozygosity; F, Fixation index. YN, Yunnan; XJ, Xinjiang; nd, no data.

fixation index, *F*), and Principal Co-ordinates Analysis (PCoA) were implemented in GenAlEx v6.503 (Peakall and Smouse, 2012) for the 53 ‘populations’. GenAlEx was also used to carry out analysis of molecular variance (AMOVA) separately on cultivated walnut, and the three wild populations. AMOVA was repeated for the entire dataset, treating cultivars vs wild walnuts as a level of separation above ‘population’, in order to examine the partitioning of genetic diversity between the cultivated and wild groups. We also carried out clonal assignment using GENODIVE v3.0 (Meirmans, 2020), a program that relies on pairwise genetic distances to group individuals into various clonal groups on the basis of a threshold

value set by the user. To arrive at the optimal threshold value, we ran trials with threshold-values ranging from 0 to 10 (0–5% of the total genetic distance; Lasso, 2008; Broeck et al., 2018). Since we had clear information that Liuj01 is a clonal population, we chose a value at which all individuals of population Liuj01 were assigned to the same clone, which was a threshold value of eight (8).

Bayesian clustering in STRUCTURE v2.3.4 (Pritchard et al., 2000) was applied to assign each individual to a given genetic cluster based on the pre-defined number of clusters (*K*). We ran 20 simulations for each value of *K* with 100,000 generations of ‘burn-in’ and 1200,000 Markov chain Monte Carlo (MCMC) generations for

increasing numbers of K subdivisions from 1 to 10. The optimal K value was determined based on the ΔK method in STRUCTURE HARVESTER (Earl and vonHoldt, 2012). Based on membership coefficients at the optimal value of K , individuals were re-assigned to their respective genetic clusters using a threshold value of 0.80. Individuals with a high probability of belonging to any of the genetic clusters ($Q \geq 0.80$) were classified as pure genetic groups, while those with low probabilities ($Q < 0.80$) were defined as 'mosaic', according to Wambulwa et al. (2016). We then recalculated genetic diversity values with all material respectively assigned to their taxonomic groups, i.e. *J. regia*, *J. sigillata* and the 'mosaic' group, to allow interspecific comparisons. AMOVA was also repeated with the entire dataset grouped into the pure (parental) groups and the 'mosaic' group. Pairwise F_{ST} values were calculated in Arlequin v3.5 (Excoffier et al., 2005) and a heat map of the F_{ST} estimates between different cultivars generated using the StAMPP R package (Pembleton et al., 2013). We used the software Populations v1.2.32 (<http://bioinformatics.org/populations/>) to construct the neighbour joining tree, and visualized it using FigTree v1.4.4 (Rambaut, 2018).

To assess the fingerprinting power of the entire SSR marker set, we extracted a data subset containing individuals with matching multilocus genotypes using the GenALEX software. From this subset, we further selected cultivars that contain at least 40% of individuals with perfectly matching genotypes across all 36 loci; the matching genotypes were considered to be identification fingerprints for the respective selected cultivars. Furthermore, for each marker we calculated the probability of identity (PI) using the formula $PI = 2(\sum Pi^2)^2 - \sum Pi^4$, where Pi is the frequency of the i th allele at a given locus, and calculated the polymorphism information content (PIC) using the online resource PICcalc (Nagy et al., 2012). We then selected a core set of SSR markers based on two criteria applied previously by Liu et al. (2017), i.e. 1) $PI < 0.2$, and 2) $PIC > 0.5$. Based on the core set of markers, we again extracted a data subset of individuals with matching multi-locus genotypes, and this was used to identify more cultivars.

3. Results

3.1. Genetic diversity and structure

Although our SSR data had some missing data points (0.007%), the SSR loci showed relatively high polymorphism with minimal stuttering and allele drop-out. The total genetic diversity indices for our dataset were generally low ($H_O = 0.570$, $H_E = 0.404$, $N_A = 2.345$) (Table 1). The mean observed heterozygosity (H_O) across loci ranged from 0.279 (population GLR) to 0.744 (hybrid cultivar 133). The average number of alleles for the entire dataset was 2.345, and ranged from 1.583 (cultivar 124) to 3.667 (population BHLT). Analysis of molecular variance (AMOVA) for cultivated walnuts found that 60.39% of the diversity resided between cultivars, and 39.61% within them (Table S2). For wild walnuts, 52.68% of the genetic diversity was partitioned among populations. In sharp contrast, AMOVA on the entire dataset (sorted as cultivated versus wild walnuts) showed that only 13.55% of the genetic variation is partitioned between those two groups (Table S3). Following clonal assignment analysis in GENODIVE program, we found that the entire dataset of 362 individuals comprise a total of 186 clones, which were defined at a threshold value (T) of eight (8) (Fig. S1). Our analysis showed that 85.1% of the cultivar groups are clonal, though at varying degrees (Table S4). All individuals of wild populations (BHLT, GLR, and TYC), and the cultivated population (XHT) did not have any clone mates in the entire collection.

STRUCTURE analysis clearly delineated individuals of *J. regia*, *J. sigillata* as well as their hybrids (Fig. 2a). At $K = 2$, which was the best K value according to the ΔK method (Fig. 2b), individuals of *J. regia* (yellow) and *J. sigillata* (red) were generally well defined. After regrouping individuals based on Q values at $K = 2$, a total of 60 cultivars initially identified in the field as *J. sigillata* were found to be 'mosaic', while seven (cultivars 146–2, 146–3, 146–4, 146–5, S115–2, X147–1, and X147–2) were found to have the *J. regia* genetic group (Table S5). One cultivar identified in the field as *J. regia* (A119–1) was re-assigned to the *J. sigillata* group, while four cultivars thought to be hybrids were re-assigned to the pure parental groups: cultivars 137–4 and 137–5 to *J. sigillata*, and cultivars 144–1 and 144–2 to *J. regia* (Table S5). The 'mosaic' group ($Q < 0.80$) showed the highest genetic diversity ($H_O = 0.611$, $H_E = 0.573$) followed by *J. sigillata* cultivars ($H_O = 0.587$, $H_E = 0.551$), while *J. regia* cultivars had the lowest value ($H_O = 0.394$, $H_E = 0.485$) (Table S6). AMOVA on the re-grouped data showed that 33.53% of the genetic variation is partitioned among the three groups (*J. regia*, *J. sigillata*, and the 'mosaic' group) (Table 2). At $K = 3$, a unique genetic group (blue) was defined for the 'Yangbidapao' landrace comprising cultivar DP and the cultivated populations Liuj01 and Liuj02. Germplasm from this new genetic group has introgressed into several cultivars that are known to be intra-specific hybrids between cultivars DP and NQ (i.e. 125, 127, 130, 131, 132, 138, 139, 140, 379), as well as the putative *J. regia* \times *J. sigillata* hybrids (133, 134, 137, and 144) (Fig. 2a).

The PCoA clustering pattern was in conformity with the pre-assigned taxonomic classes, and was generally concordant with STRUCTURE analysis results (Fig. 3). Both the first (20.99%) and second (8.95%) axes generally separated individuals of *J. regia* and *J. sigillata*. Individuals of the 'Yangbidapao' landrace ('populations' DP, Liuj01 and Liuj02) formed a small cluster at the periphery of the PCoA plot, and were separated from most of the other *J. sigillata* material by the first axis (Fig. 3). The *J. regia* \times *J. sigillata* hybrids clustered in between the respective parental groups, namely *J. regia* and *J. sigillata*. Additionally, some individuals exhibited unexpected clustering patterns in the PCoA plot; e.g. individual A119–1, which belongs to cultivar A119 (*J. regia*) according to the germplasm bank record, clustered with *J. sigillata* material.

3.2. Genetic differentiation and relationships

The average inbreeding coefficient (F_{IS}) was -0.376 , with all 'populations' showing negative F_{IS} values except for populations BHLT and TYC and cultivar X116. Pairwise F_{ST} values were generally high between *J. regia* and *J. sigillata* cultivars (Fig. 4). Moreover, F_{ST} values were generally high between *J. regia* (GLR, XHT, A119, X116 and 120) and the rest of the 'populations', including hybrids.

The neighbour joining (NJ) tree of the 53 cultivars and wild populations showed a similar general trend, with all *J. regia* material forming a single cluster (Fig. 5). However, the *J. regia* group was a sub-cluster of a larger group which included several *J. sigillata* cultivars, namely L144, L149, X142, 124, 136, 149. Three *J. regia* \times *J. sigillata* hybrid cultivars (133, 134 and 137) clustered together with some Yangbi County cultivars (DP, 125, 127, 130, 131). The other *J. regia* \times *J. sigillata* hybrid cultivar (144) clustered with two *J. sigillata* cultivars, S115 and 114. Results of the NJ tree for the 362 individuals were consistent with the population tree (Fig. S2), except for a few individuals, which did not conform to the expected clustering patterns. For instance, individual A119–1 was identified in the field as *J. regia* but clustered with *J. sigillata* cultivars. Correspondingly, STRUCTURE analysis also showed a substantial degree of genetic admixture within cultivar A119 (Fig. 2a). Several other individuals

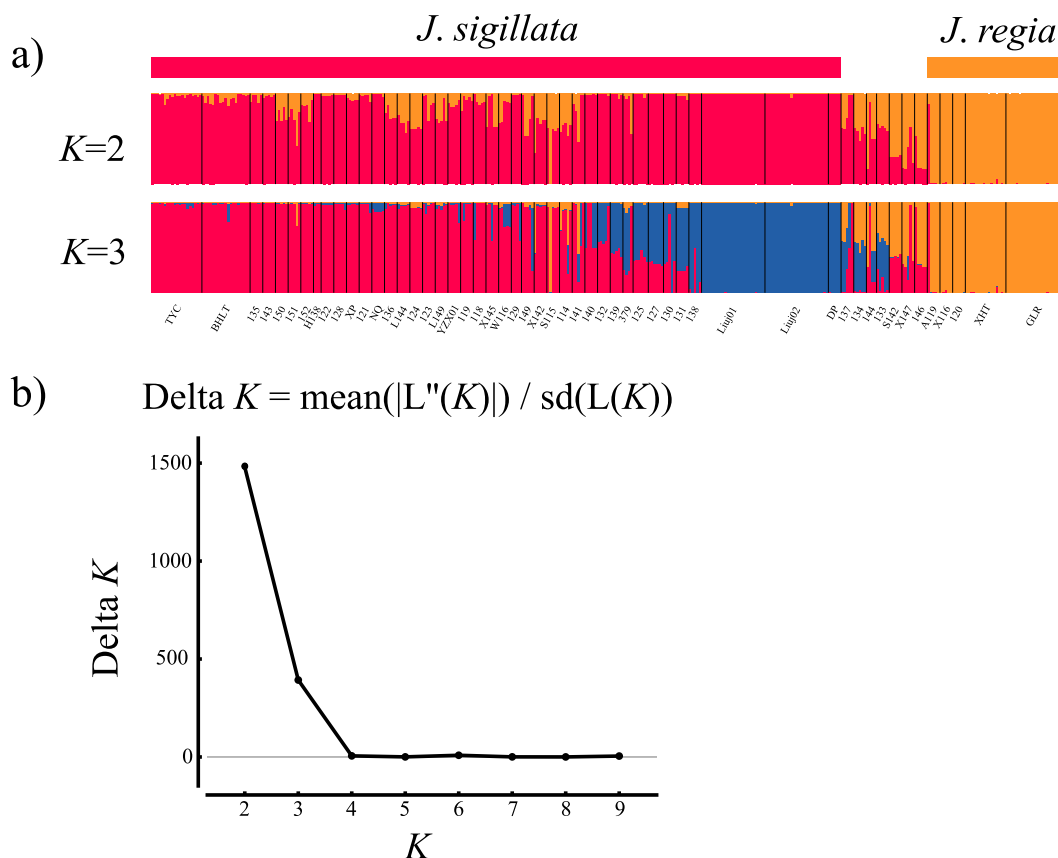


Fig. 2. Genetic clustering of the 47 cultivars and 6 wild populations based on 36 SSR loci. (a) Bayesian inference of population structure at $K = 2$ and $K = 3$ for all the 53 ‘populations’ of *J. regia* and *J. sigillata* using STRUCTURE. (b) Inference of the optimal K value using the Delta K (ΔK) method.

Table 2
Analysis of variance (AMOVA) for all the 362 samples separated into three groups: *J. regia*, *J. sigillata*, and the ‘mosaic’ group after assigning individuals to their respective groups based on a Q value threshold of 0.80 in STRUCTURE.

| Source | d.f. | Sum of squares | Mean squares | Percentage of variation (%) |
|---------------|------|----------------|--------------|-----------------------------|
| Among groups | 2 | 2004.493 | 1002.246 | 33.53 |
| Within groups | 359 | 7102.179 | 19.783 | 66.47 |
| Total | 361 | 9106.671 | | 100 |

Note: d.f., degree of freedom.

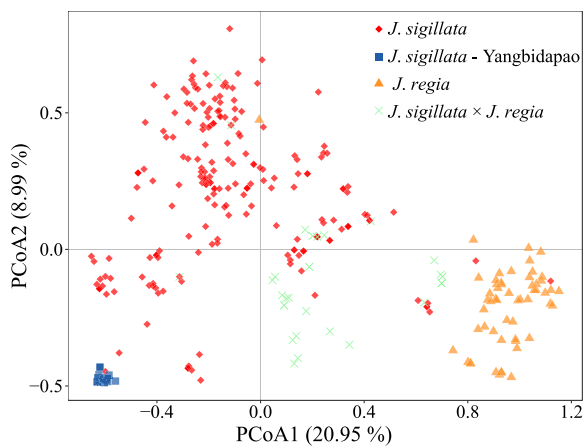


Fig. 3. Principal Co-ordinates Analysis (PCoA) of all 362 individuals (130 wild, and 232 cultivated). The colours represent the groups as follows: Red, *J. sigillata*; ‘Yangbidapao’ landraces, Orange, *J. regia*; Green, *J. regia* × *sigillata* putative hybrids.

such as 114–4, 137–5, 144–1 and 144–2, clustered with conspecific individuals but failed to cluster with other individuals belonging to the same cultivar.

The three wild populations clustered with certain cultivars in the NJ tree; the wild *J. regia* population GLR clustered with three *J. regia* cultivars (A119, X116 and 120) as well as with the cultivated population XHT. Two wild *J. sigillata* populations, BHLT and TYC both clustered with cultivars 118, 119 and 123.

The entire panel of 36 SSR markers could identify 19 cultivars (Table S7). After filtering the SSR loci based on the set thresholds of PI and PIC, we identified a set of 18 core loci (Table 3). A total of 32 cultivars showed perfect multi-locus matches at all the 18 core loci, although some individuals within certain cultivars were non-conforming at a few loci. The other 15 cultivars contained individuals that did not show consistent intra-cultivar matches, and were therefore deemed to be unidentifiable. More importantly, we did not detect any multi-locus matches between individuals of different cultivars. If any individuals within a cultivar had matching genotypes across all the core loci, then that particular genotype profile was deemed to be a standard fingerprint for the cultivar. The

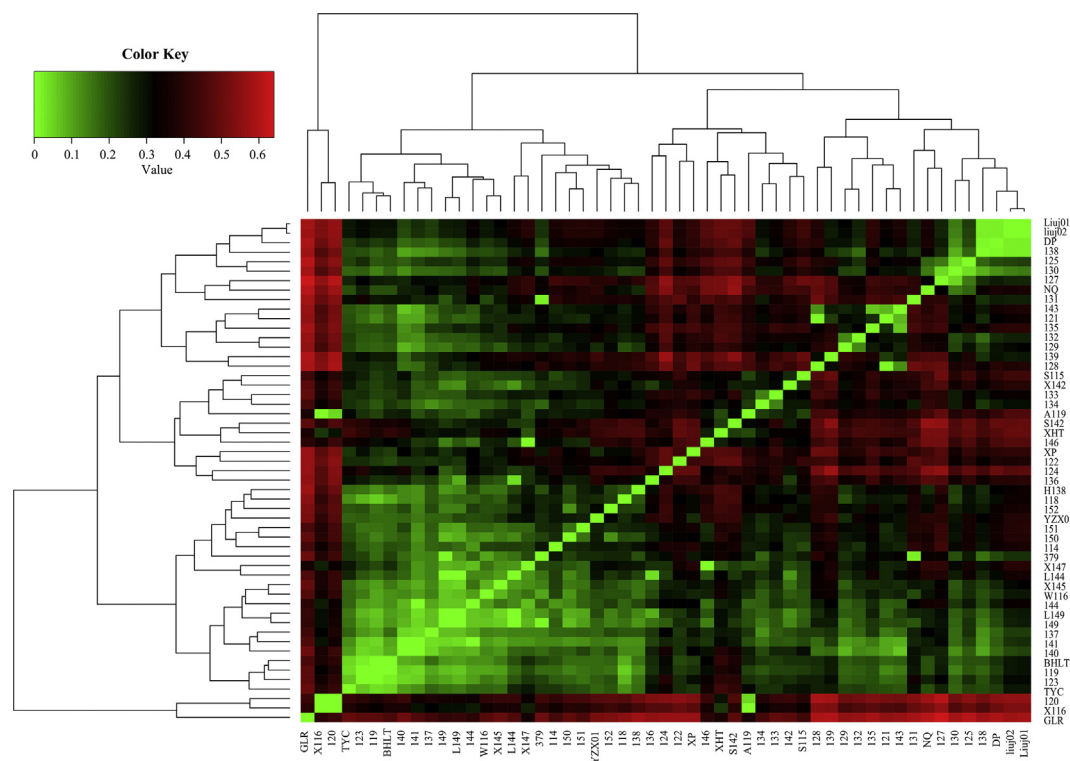


Fig. 4. Heat map depicting pairwise F_{ST} estimates between 'populations' (both cultivar groups and wild trees). The 'populations' are across the X- and Y-axis, with each square indicating the F_{ST} estimate between the two respective individuals. The magnitude of the genetic differentiation is indicated by the colour ramp.

fingerprint profiles for the 32 cultivars across the 18 core SSR loci are provided in Table S8.

4. Discussion

4.1. Low genetic diversity of the walnut germplasm

We found that walnuts from southwest China harbor a low to moderate level of genetic diversity (Table 1), which is generally comparable to previous estimates for *J. regia* from other regions (Bernard et al., 2018; Torokeldiev et al., 2018; Vischi et al., 2017). However, Ebrahimi et al. (2017) found higher genetic diversity values of *J. regia* ($N_A = 7.4$, $H_E = 0.75$, $H_O = 0.61$) which could be as a result of the wider sampling range used in that study. Similarly, a recent study of walnuts from a germplasm bank found higher genetic diversity values ($N_A = 9.45$, $H_O = 0.67$) (Balapanov et al., 2019). Such direct comparisons of genetic diversity indices, however, should be made cautiously because the diversity values may be influenced by factors such as sample size (Chai et al., 2017), marker type (Igwe et al., 2017), and even the nature of the microsatellite loci employed. The genetic diversity of domesticated species is influenced, in part, by their mating system and mode of dissemination and propagation (Ingvarsson and Dahlberg, 2019). The relatively low genetic diversity in the current study could be attributed to clonal propagation by grafting in some cultivars, as corroborated by the surprisingly high number of clones (93.48%) in the present analysis. Indeed, grafting has been widely adopted in southwest China for production of walnut cultivars. Though grafting may guarantee consistency in the desirable agronomic characteristics such as nut and kernel weight (Khadivi et al., 2019), it often leads to over-reliance on one or just a few cultivars as scion sources, thus lowering genetic diversity. Reduced genetic diversity

due to grafting has also been observed before in purple mombin (*Spondias purpurea*) (Miller and Schaal, 2006) from South America, almond (*Prunus dulcis*) from Lebanon (Hamadeh et al., 2018) and sweet chestnut (*Castanea sativa*) from Europe (Mattioli et al., 2008). We found consistently negative F_{IS} values for most of the cultivars, a trend that has previously been linked to asexual reproduction in wild cherry populations (Stoeckel et al., 2006). Based on observations and information gathered during our sample collection, most cultivars from southwest China in our sample set are propagated by grafting, hence the negative F_{IS} values, which indicate genetic homogeneity. Although vegetative propagation provides a reliable tool for passing of traits to the next generation, this method has been associated with reduced genetic diversity as a result of lack of sexual recombination (McKey et al., 2010; Miller and Gross, 2011). Such low levels of genetic diversity may compromise the evolutionary potential of a crop species, hence reducing its resilience and adaptability to abiotic challenges in the wake of climate change (Henry, 2019), as well as biotic challenges like the severe pathogen and pest pressures currently facing the walnut industry (Li et al., 2017; Seybold et al., 2019).

4.2. Introgression and implications in germplasm improvement

Although STRUCTURE analysis confirmed the hybrid nature of the putative *J. regia* × *J. sigillata* hybrids (cultivars 133, 134, 137 and 144) (Fig. 2), these four hybrid cultivars did not cluster together in the NJ tree (Fig. 5), indicating the use of diverse parental materials for cultivar development. This inference is supported by the high H_O values of the putative *J. regia* × *J. sigillata* hybrids detected in our analysis. The NJ analysis indicated that *J. regia* × *J. sigillata* hybrids (other than cultivar 144) are genetically closer to the *J. sigillata* parent, and were likely developed from the "Yangbidapao"

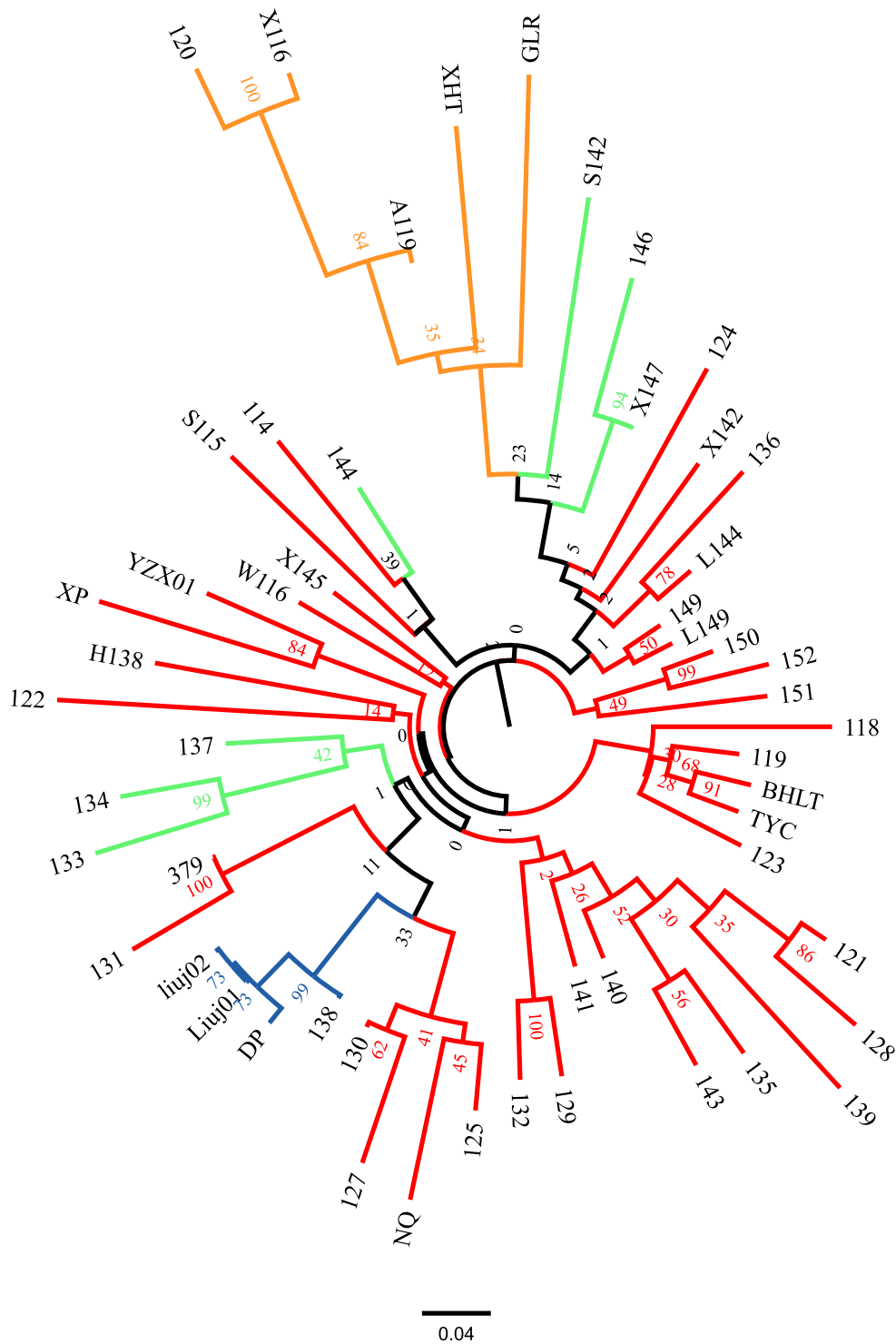


Fig. 5. Neighbour-joining tree of the 53 ‘populations’ from the two *Juglans* species based on Nei’s genetic distance. The colours correspond to groups identified in PCoA in Fig. 2.

landrace, which is popularly cultivated in west Yunnan (e.g. Yangbi and Channing Counties) for its economic and cultural significance. The re-grouped dataset based on STRUCTURE results at $K = 2$ also suggested that a total of 60 cultivars that were initially thought to be pure *J. sigillata* material are likely artificial interspecific hybrids between the two walnut species. This observation indicates that there could be more *J. regia* × *J. sigillata* hybrids in the locally conserved germplasm than is documented, and this can be a potential source of confusion both for farmers and breeders.

Furthermore, STRUCTURE analysis at $K = 3$ and PCoA revealed a distinct genetic group that generally comprised the four *J. regia* × *J. sigillata* putative hybrids (cultivars 133, 134, 137, and 144), the intraspecific DP × NQ hybrids (cultivars 125, 127, 130, 131, 138, 139, 140, and 379) and the ‘Yangbidapao’ landraces (DP, Liuj01 and Liuj02). Although the optimal number of clusters was $K = 2$, we argue that this new group (blue at $K = 3$) represents a distinct genetic unit, and may have substantial evolutionary significance. Cultivars of this distinct gene pool certainly represent genetically

Table 3
Eighteen (18) core set of SSR markers selected based on the Probability of Identity and Polymorphism Information Content.

| Locus | Probability of Identity (PI) | Polymorphism Information Content (PIC) |
|-----------|------------------------------|--|
| JS12 | 0.142254 | 0.6449 |
| JS24 | 0.12283 | 0.6693 |
| JS03 | 0.156423 | 0.626 |
| JS07 | 0.129882 | 0.6656 |
| JR11 | 0.164741 | 0.6149 |
| JS19 | 0.078461 | 0.7489 |
| JS041 | 0.095226 | 0.7192 |
| JS05 | 0.099624 | 0.716 |
| BFU-jr38 | 0.128345 | 0.6561 |
| BFU-jr277 | 0.069056 | 0.7703 |
| JS15 | 0.196285 | 0.5765 |
| JS14 | 0.15773 | 0.6227 |
| JS22 | 0.085149 | 0.7421 |
| JS25 | 0.148138 | 0.6404 |
| JR04 | 0.181196 | 0.5843 |
| JS13 | 0.120937 | 0.6796 |
| ZMZ35 | 0.152727 | 0.6267 |
| SSR18 | 0.173492 | 0.6065 |

rich resources that should be conserved both *in* and *ex situ* for future utilization by breeders.

Despite the obvious phenotypic differences between *J. regia* and *J. sigillata* (Fig. 1) (Lu et al., 1999), several population genetic studies have indicated that these two species have no genetic barriers and that introgression has occurred between them (Dang et al., 2019; Wang et al., 2015; Yuan et al., 2018). Consequently, interspecific hybrids (both natural and artificial) between the two species are possible, and in fact, the two species might be indistinguishable based on SSR markers at local scale in northwest Yunnan (Gunn et al., 2010). Consistent with these previous studies, our STRUCTURE and PCoA results unveiled significant levels of introgression between *J. regia* and *J. sigillata* gene pools (Figs. 2 and 3), which together have been cultivated in China for thousands of years. The introgression could be as a result of artificial breeding between the two species, to combine their favorable traits. Despite this evidence of interspecific hybrids, the significantly high F_{ST} values between *J. sigillata* cultivars and the wild *J. regia* population GLR could be indicative of underutilization of wild *J. regia* genetic resources in breeding and improvement of *J. sigillata* cultivars.

Wild relatives of cultivated plants are important reservoirs of genes for improving commercial varieties, particularly in the context of climate change (Zhang and Batley, 2019). Our F_{ST} and NJ results (Figs. 4 and 5) indicated that the wild gene pools might have been utilized as parental sources in the breeding of some cultivars. For instance, the three cultivars 118, 119 and 123 might have been directly derived from wild *J. sigillata*, and hence might represent a unique gene pool that should be explored to enrich the walnut germplasm in Yunnan. Though cultivars A119, X116 and 120 clustered with the wild *J. regia* population (GLR), it is not certain whether they were derived directly from wild *J. regia*, as they could have originated via other *J. regia* cultivars not examined here. We detected only limited genetic differentiation (13.55% of the total; Table S4) between cultivated walnuts in Yunnan Province, and their wild progenitors. However, a separate AMOVA analysis of wild populations showed that a higher proportion of the genetic variation resides among than within populations, signifying genetic differentiation, and underscoring the role of ecological barriers and other vicariant factors that usually shape the genetic structure of natural plant communities in southwest China (Jia et al., 2020; Liu et al., 2013). This relatively high genetic differentiation among wild populations of walnuts should be tapped and utilized to increase genetic heterogeneity among cultivated walnuts in Yunnan. Introgression of wild genotypes into cultivars has been demonstrated in

many species (e.g. Dutra et al., 2018; He et al., 2019; Janzen et al., 2018), further underscoring the importance of wild genetic resources in improvement of walnut cultivars. Hence although wild walnut trees have been substantially utilized in germplasm improvement in the region, there is far more that can be gained from them, and wild populations should therefore be protected through such interventions as *in situ* conservation and outlawing the unsustainable use of walnut trees.

4.3. Molecular identification of walnut

Despite the evidence for introgression discussed above, our structure analysis showed that the two walnut species (*J. regia* and *J. sigillata*) can be distinguished based on SSR markers. Our finding is, however, discordant with some previous SSR-based studies, which reported minimal genetic differentiation between *J. regia* and *J. sigillata* (e.g. Dang et al., 2019; Gunn et al., 2010; Wang and Pei, 2008). This inconsistency could be ascribed to differences in sampling design, as well as the use of fewer and/or non-neutral (EST-) SSR markers, which could have limited the discrimination power of these previous studies. At the level of cultivars, the entire set of markers could only identify 19 cultivars, and this low number could be attributed to the high PI and low PIC values for some of the loci. To increase the number of cultivars that can be identified, we established a key for identifying 32 cultivars based on a core selection of 18 SSR loci. These fingerprints are a useful resource for cultivar identification as well as for informing the choice of parental material during breeding. For ease of accessing and confirming the SSR genotypes, the alleles can be converted into two-dimensional QR codes using a website (<http://cli.im/>); see for example QR code for cultivar DP (Fig. 6). Our analysis uncovered several cases of potential cultivar misidentifications. The non-conformity of the seven cultivars identified in the field as *J. sigillata*, and cultivar A119-1 thought to be *J. regia* indicates possible erroneous identifications (due to phenotypic plasticity), though interspecific introgression could also explain the disharmony. Failure by some individuals within a given cultivar to show consistent matches at all the 18 core loci also points to possible cases of misidentification. For example, of the six individuals of cultivar 127, only four (127-2, 127-3, 127-4 and 127-5) showed consistent matches across the entire set of core markers. Individuals 127-1 and 127-6 were possibly either erroneously identified as cultivar 127 during collection or were collected from scions that originated from a different cultivar. This discordance could be as a result of the



Fig. 6. QR code for rapid FOR access of genotyping information for cultivar DP. Scanning the code gives the following information: Name of cultivar, Species, Cultivation region and the biallelic multilocus fingerprint at the 18 core loci.

reliance on morphological characteristics to identify cultivars. The utility of morphological characters can be limited by their phenotypic plasticity. For cultivated walnuts, morphological identification may be hampered by the frequent interspecific introgression between *J. regia* and *J. sigillata*. In addition, introgression often blurs the known traditional boundaries between species distribution ranges, potentially resulting in misidentifications. Such cultivar mis-assignments have been detected in *ex situ* germplasm collections for *J. ailantifolia* (Hoban and Romero-Severson, 2011), and tea (Wambulwa et al., 2016) among other crop species. Intra-cultivar mis-assignments (e.g. for individuals 114–4, 137–5, 144–1 and 144–2) were probably due to errors when recording sample identities. Correction of such misidentifications will certainly be useful in the management and effective utilization of walnut genetic resources, both for conservational and agricultural purposes.

To our knowledge, the present study represents the first genetic analysis of walnut cultivars from southwest China. Although the disparities in the technical aspects among previous studies present some difficulties in making meaningful comparisons, our results suggested that the genetic diversity of cultivated walnuts in Yunnan is rather low, likely due to reliance on grafting as the propagation method. The present study also demonstrates how breeders in the region can harness the substantial genetic differentiation among wild populations to the ultimate benefit of walnut improvement programs. For future work, we recommend the inclusion of more *J. regia* individuals and the use of high-density markers such as SNPs from genomic data. In addition, more populations of wild and cultivated walnuts need to be sampled, in order to overcome the weaknesses of the present study. Notwithstanding the limitations of our study, our data confirmed the efficiency and discrimination power of nuclear microsatellites for identification of the various walnut cultivars.

Author contributions

ZYW, DZL and JL designed the research and acquired funding. JL, YHL, and ZCY collected the materials. ZYX, YCJ, ZCX, LJY and PZF

performed the experiments. MCW, LJY, YHL, and PZF carried out data analysis. MCW, RM, and JL wrote the first draft of the manuscript. All authors have contributed to interpretation of the results and editing of the manuscript.

Data availability

The SSR data generated in the current study have been archived at https://pan.baidu.com/s/1ETTWTbhUCb_7c9Nu3xKA0g (extraction code: 3ye1).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could potentially influence the results reported in this article.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (31770367, 41971071), Top-notch Young Talents Project of Yunnan Provincial “Ten Thousand Talents Program” (YNWR-QNBj-2018-146), the Key Research Program of Frontier Sciences, CAS (ZDBS-LY-7001), and Natural Science Foundation of Yunnan (2017FB027). Zeng-Yuan Wu was supported by CAS’ Youth Innovation Promotion Association (2019385), the Biological Resources Program, Chinese Academy of Sciences (KFJ-BRP-017-XX). Moses Wambulwa was supported by the Postdoctoral International Exchange Program of the Office of China Postdoctoral Council, the Postdoctoral Targeted Funding and Postdoctoral Research Fund of Yunnan Province. Molecular experiments were performed at the Laboratory of Molecular Biology at the Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences. We are grateful to Mr. Xue-Wen Liu, Tao Liu, Dr. Tao Wu and other volunteers for their help with field work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pld.2021.08.005>.

References

- Abdallah, I.B., Tlili, N., Martinez-Force, E., et al., 2015. Content of carotenoids, tocopherols, sterols, triterpenic and aliphatic alcohols, and volatile compounds in six walnuts (*Juglans regia* L.) varieties. *Food Chem.* 173, 972–978. <https://doi.org/10.1016/j.foodchem.2014.10.095>.
- Aradhya, M.K., Potter, D., Gao, F., et al., 2007. Molecular phylogeny of *Juglans* (Juglandaceae): a biogeographic perspective. *Tree Genet. Genomes* 3, 363–378. <https://doi.org/10.1007/s11295-006-0078-5>.
- Aune, D., Keum, N., Giovannucci, E., et al., 2016. Nut consumption and risk of cardiovascular disease, total cancer, all-cause and causespecific mortality: a systematic review and dose-response meta-analysis of prospective studies. *BMC Medicine* 14, 207. <https://doi.org/10.1186/s12916-016-0730-3>.
- Balapanov, I., Suprun, I., Stepanov, I., et al., 2019. Comparative analysis Crimean, Moldavian and Kuban Persian walnut collections genetic variability by SSR-markers. *Sci. Hortic.* 253, 322–326. <https://doi.org/10.1016/j.scienta.2019.04.014>.
- Bernard, A., Barreneche, T., Lheureux, F., et al., 2018. Analysis of genetic diversity and structure in a worldwide walnut (*Juglans regia* L.) germplasm using SSR markers. *PLoS ONE* 13, e0208021. <https://doi.org/10.1371/journal.pone.0208021>.
- Bouabdallah, I., Bouali, I., Martinez-Force, E., et al., 2014. Composition of fatty acids, triacylglycerols and polar compounds of different walnut varieties (*Juglans regia* L.) from Tunisia. *Nat. Prod. Res.* 28, 1826–1833. <https://doi.org/10.1080/14786419.2014.950573>.
- Brock, A.V., Cox, K., Melosik, I., et al., 2018. Genetic diversity loss and homogenization in urban trees: the case of *Tilia* × *europaea* in Belgium and The Netherlands. *Biodivers. Conserv.* 27, 3777–3792.
- Chai, X., Dong, R., Liu, W., et al., 2017. Optimizing sample size to assess the genetic diversity in common vetch (*Vicia sativa* L.) populations using Start Codon Targeted (SCoT) markers. *Molecules* 22, 567. <https://doi.org/10.3390/molecules22040567>.

- Chen, F., Chen, J., Wang, Z., et al., 2019. Genomics: cracking the mysteries of walnuts. *J. Genet.* 98, 33. <https://doi.org/10.1007/s12041-019-1084-3>.
- Dang, D., Yue, M., Zhang, M., et al., 2019. Gene introgression among closely related species in sympatric populations: a case study of three walnut (*Juglans*) species. *Forests* 10, 965. <https://doi.org/10.3390/f10110965>.
- Dangl, G.S., 2005. Characterization of 14 microsatellite markers for genetic analysis and cultivar identification of walnut. *J. Am. Soc. Hortic. Sci.* 130, 348–354. <https://doi.org/10.21273/JASHS.130.3.348>.
- Dong, W., Xu, C., Li, W., et al., 2017. Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences. *Front. Plant Sci.* 8, 1148. <https://doi.org/10.3389/fpls.2017.01148>.
- Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bulletin* 19, 11–15. <https://worldveg.tind.io/record/33886/>.
- Dutra, W., Guerra, Y., Ramos, J., et al., 2018. Introgression of wild alleles into the tetraploid peanut crop to improve water use efficiency, earliness and yield. *PLoS ONE* 13, e0198776. <https://doi.org/10.1371/journal.pone.0198776>.
- Earl, D.A., vonHoldt, B.M., 2012. Structure HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>.
- Ebrahimi, A., Mathur, S., Lawson, S.S., et al., 2019. Microsatellite borders and micro-sequence conservation in *Juglans*. *Sci. Rep.* 9, 3748. <https://doi.org/10.1038/s41598-019-39793-z>.
- Ebrahimi, A., Zarei, A., McKenna, J.R., et al., 2017. Genetic diversity of Persian walnut (*Juglans regia*) in the cold-temperate zone of the United States and Europe. *Sci. Hortic.* 220, 36–41. <https://doi.org/10.1016/j.scienta.2017.03.030>.
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform.* 1, 47–50. <https://doi.org/10.1177/117693430500100003>.
- FAOSTAT, 2020. FAOSTAT: FAO Statistics Division 2020. <https://www.fao.org/faostat/en/#data/QC>. (Accessed 14 September 2020).
- Feng, X., Zhou, H., Zulfiqar, S., et al., 2018. The phylogeographic history of common walnut in China. *Front. Plant Sci.* 9, 1399. <https://doi.org/10.3389/fpls.2018.01399>.
- Freeman, C.C., Reveal, J.L., 2005. *Flora of North America*. Missouri Botanical Garden, St. Louis, USA.
- Gharibzadeh, S.M.T., Mousavi, S.M., Hamed, M., et al., 2012. Comparative analysis of new Persian walnut cultivars: nut/kernel geometrical, gravimetric, frictional and mechanical attributes and kernel chemical composition. *Sci. Hortic.* 135, 202–209. <https://doi.org/10.1016/j.scienta.2011.11.030>.
- Gunn, B.F., Aradhya, M., Salick, J.M., et al., 2010. Genetic variation in walnuts (*Juglans regia* and *J. sigillata*; Juglandaceae): species distinctions, human impacts, and the conservation of agrobiodiversity in Yunnan, China. *Am. J. Bot.* 97, 660–671. <https://doi.org/10.3732/ajb.0900114>.
- Gupta, A., Behl, T., Panichayupakaranan, P., 2019. A review of phytochemistry and pharmacology profile of *Juglans regia*. *Obes. Med.* 16, 100142. <https://doi.org/10.1016/j.obmed.2019.100142>.
- Gyawali, R., Ibrahim, S.A., 2014. Natural products as antimicrobial agents. *Food Contr.* 46, 412–429. <https://doi.org/10.1016/j.foodcont.2014.05.047>.
- Hamadeh, B., Chalak, L., d'Eeckenbrugge, G.C., et al., 2018. Evolution of almond genetic diversity and farmer practices in Lebanon: impacts of the diffusion of a graft-propagated cultivar in a traditional system based on seed-propagation. *BMC Plant Biol.* 18, 155. <https://doi.org/10.1186/s12870-018-1372-8>.
- Hassani, D., Sarikhani, S., Dastjerdi, R., et al., 2019. Situation and recent trends on cultivation and breeding of Persian walnut in Iran. *Sci. Hortic.* 270, 109369. <https://doi.org/10.1016/j.scienta.2020.109369>.
- He, F., Pasam, R., Shi, F., et al., 2019. Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat. Genet.* 51, 896–904. <https://doi.org/10.1038/s41588-019-0382-2>.
- Henry, R.J., 2019. Innovations in plant genetics adapting agriculture to climate change. *Curr. Opin. Plant Biol.* 13, 1–6. <https://doi.org/10.1016/j.pbi.2019.11.004>.
- Hoban, S., Romero-Severson, J., 2011. Homonymy, synonymy and hybrid misassignments in butternut (*Juglans cinerea*) and Japanese walnut (*Juglans ailanatifolia*) nut cultivars. *Genet. Resour. Crop Evol.* 59, 1397–1405. <https://doi.org/10.1007/s10722-011-9767-5>.
- Igwe, D.O., Afuwka, C.A., Ubi, B.E., et al., 2017. Assessment of genetic diversity in *Vigna unguiculata* L. (Walp) accessions using inter-simple sequence repeat (ISSR) and start codon targeted (SCoT) polymorphic markers. *BMC Genetics* 18, 98. <https://doi.org/10.1186/s12863-017-0567-6>.
- Ingvarsson, P.K., Dahlberg, H., 2019. The effects of clonal forestry on genetic diversity in wild and domesticated stands of forest trees. *Scand. J. For. Res.* 34, 370–379. <https://doi.org/10.1080/02827581.2018.1469665>.
- Janzen, G.M., Wang, L., Hufford, M.B., 2018. The extent of adaptive wild introgression in crops. *New Phytol.* 221, 1279–1288. <https://doi.org/10.1111/nph.15457>.
- Jia, Y., Milne, R., Zhu, J., et al., 2020. Evolutionary legacy of a forest plantation tree species (*Pinus armandii*): implications for widespread afforestation. *Evol. Appl.* 13, 2646–2662. <https://doi.org/10.1111/eva.13064>.
- Khadivi, A., Montazeran, A., Rezaei, M., et al., 2019. The pomological characterization of walnut (*Juglans regia* L.) to select the superior genotypes – an opportunity for genetic improvement. *Sci. Hortic.* 248, 29–33. <https://doi.org/10.1016/j.scienta.2018.12.054>.
- Lasso, E., 2008. The importance of setting the right genetic distance threshold for identification of clones using amplified fragment length polymorphism: a case study with five species in the tropical plant genus *Piper*. *Mol. Ecol. Resour.* 8, 74–82.
- Li, L., 2017. *Growth and propagation characteristics of Juglans interspecific hybrid varieties*. Chinese Academy of Forestry, Beijing, China.
- Li, X., He, P., Zhu, T., et al., 2017. Characterization of a new pathogen in walnut twig blight. *Microbiol. China* 44, 1339–1348. <https://doi.org/10.13344/j.microbiol.china.160946>.
- Liu, J., Gao, L., 2011. Comparative analysis of three different methods of total DNA extraction used in *Taxus*. *Guihaia* 31, 244–249. <https://www.cabdirect.org/cabdirect/abstract/20113266180>.
- Liu, B., Liang, J., Zhao, D., et al., 2020. Morphological and compositional analysis of two walnut (*Juglans regia* L.) cultivars growing in China. *Plant Foods Hum. Nutr.* 75, 116–123. <https://doi.org/10.1007/s11130-019-00794-y>.
- Liu, S., Liu, H., Wu, A., et al., 2017. Construction of fingerprinting for tea plant (*Camellia sinensis*) accessions using new genomic SSR markers. *Mol. Breed.* 37, 93. <https://doi.org/10.1007/s11032-017-0692-y>.
- Liu, J., Möller, M., Provan, J., et al., 2013. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytol.* 199, 1093–1108. <https://doi.org/10.1111/nph.12336>.
- Lorenzo, S., Giovanna, S., Michelangelo, M., et al., 2020. Cocoa beans and liquor fingerprinting: a real case involving SSR profiling of CCN51 and “Nacional” varieties. *Food Contr.* 118, 107392. <https://doi.org/10.1016/j.foodcont.2020.107392>.
- Lu, A., Stone, D., Grauke, L., 1999. *Juglandaceae*. In: Wu, Z.Y., Peter, R.H. (Eds.), *Flora of China*. Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis, Missouri, pp. 277–285.
- Manning, W., 1978. The classification within the Juglandaceae. *Ann. Mo. Bot. Gard.* 65, 1058–1087. <https://doi.org/10.2307/2398782>.
- Martínez, M.L., Mattea, M.A., Maestri, D.M., 2006. Varietal and crop year effects on lipid composition of walnut (*Juglans regia*) genotypes. *J. Am. Oil Chem. Soc.* 83, 791–796. <https://doi.org/10.1007/s11746-006-5016-z>.
- Mattioni, C., Cherubini, M., Micheli, E., et al., 2008. Role of domestication in shaping *Castanea sativa* genetic variation in Europe. *Tree Genet. Genomes* 4, 563–574. <https://doi.org/10.1007/s11295-008-0132-6>.
- McGrath, G., Leslie, C., 2009. Breeding walnuts (*Juglans regia*). In: Priyadarshan, P.M., Jain, S.M. (Eds.), *Breeding Plantation Tree Crops: Temperate Species*. Springer, New York, pp. 249–273. https://doi.org/10.1007/978-0-387-71203-1_8.
- McKey, D., Elias, M., Pujol, B., et al., 2010. The evolutionary ecology of clonally propagated domesticated plants. *New Phytol.* 186, 318–332. <https://doi.org/10.1111/j.1469-8137.2010.03210.x>.
- Meirmans, P.G., 2020. GENODIVE version 3.0: easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol. Ecol. Resour.* 20, 1126–1131. <https://doi.org/10.1111/1755-0998.13145>.
- Miller, A., Gross, B., 2011. From forest to field: perennial fruit crop domestication. *Am. J. Bot.* 98, 1389–1414. <https://doi.org/10.3732/ajb.1000522>.
- Miller, A.J., Schaal, B.A., 2006. Domestication and the distribution of genetic variation in wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* L. (Anacardiaceae). *Mol. Ecol.* 15, 1467–1480. <https://doi.org/10.1111/j.1365-294X.2006.02834.x>.
- Nagy, S., Poczai, P., Cernak, I., Gorji, A.M., Hegedus, G., Taller, J., 2012. PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. *Biochem. Genet.* 50, 670–672. <https://doi.org/10.1007/s10528-012-9509-1>.
- Ojeda-Amador, R.M., Salvador, M.D., Gómez-Alonso, S., et al., 2018. Characterization of virgin walnut oils and their residual cakes produced from different varieties. *Food Res. Int.* 108, 396–404. <https://doi.org/10.1016/j.foodres.2018.03.066>.
- Peakall, R., Smouse, P.E., 2012. GenAlix 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28, 2537–2539. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>.
- Pei, D., Lu, X.Z., 2011. *Walnut Germplasm Resources in China*. China Forestry Publishing House, Beijing, China.
- Pembleton, L.W., Cogan, N.O.L., Forster, J.W., 2013. STAMP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Resour.* 13, 946–952. <https://doi.org/10.1111/1755-0998.12129>.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. <https://www.genetics.org/content/155/2/945>.
- Rabadán, A., Álvarez-Ortí, M., Gómez, R., et al., 2017. Suitability of Spanish almond cultivars for the industrial production of almond oil and defatted flour. *Sci. Hortic.* 225, 539–546. <https://doi.org/10.1016/j.scienta.2017.07.051>.
- Rabadán, A., Pardo, J.E., Pardo-Giménez, A., et al., 2018. Effect of genotype and crop year on the nutritional value of walnut virgin oil and defatted flour. *Sci. Total Environ.* 634, 1092–1099. <https://doi.org/10.1016/j.scitotenv.2018.04.090>.
- Rambaut, A., 2018. FigTree v1.4.4. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. <http://tree.bio.ed.ac.uk/software/figtree/>. (Accessed 15 July 2020). Accessed.
- Ros, E., Izquierdo-Pulido, M., Sala-Vila, A., 2018. Beneficial effects of walnut consumption on human health: role of micronutrients. *Curr. Opin. Clin. Nutr. Metab. Care* 21, 498–504. <https://doi.org/10.1097/MCO.0000000000000508>.

- Seybold, S.J., Klingeman III, W.E., Hishinuma, S.M., et al., 2019. Status and impact of walnut twig beetle in urban forest, orchard, and native forest ecosystems. *J. For.* 117, 152–163. <https://doi.org/10.1093/jofore/fvy081>.
- Stoeckel, S., Grange, J., Fernández-Manjarres, J., et al., 2006. Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Mol. Ecol.* 15, 2109–2118. <https://doi.org/10.1111/j.1365-294x.2006.02926.x>.
- Tian, J., Wu, Y., Wang, Y., Han, F., 2009. Development and prospects of the walnut industry in China. *ISHS Acta Hortic.* 861. <https://doi.org/10.17660/ActaHortic.2010.861.2>. VI International Walnut Symposium.
- Torokeldiev, N., Ziehe, M., Gailing, O., et al., 2018. Genetic diversity and structure of natural *Juglans regia* L. populations in the southern Kyrgyz Republic revealed by nuclear SSR and EST-SSR markers. *Tree Genet. Genomes* 15, 5. <https://doi.org/10.1007/s11295-018-1311-8>.
- Vischi, M., Chiabà, C., Raranciuc, S., et al., 2017. Genetic diversity of walnut (*Juglans Regia* L.) in the eastern Italian alps. *Forests* 8, 81. <https://doi.org/10.3390/f8030081>.
- Wambulwa, M., Meegahakumbura, M., Chalo, R., et al., 2016. Nuclear microsatellites reveal the genetic architecture and breeding history of tea germplasm of East Africa. *Tree Genet. Genomes* 12, 11. <https://doi.org/10.1007/s11295-015-0963-x>.
- Wang, H., Pan, G., Ma, Q., et al., 2015. The genetic diversity and introgression of *Juglans regia* and *Juglans sigillata* in Tibet as revealed by SSR markers. *Tree Genet. Genomes* 11, 804. <https://doi.org/10.1007/s11295-014-0804-3>.
- Wang, H., Pei, D., 2008. Genetic diversity and structure of walnut populations in central and southwestern China revealed by microsatellite markers. *J. Am. Soc. Hortic. Sci.* 133, 197–203. <https://doi.org/10.21273/JASHS.133.2.197>.
- Wu, G.L., Meng, H.J., Hao, Y.Y., Liu, Q.L., Wang, D., Tian, J.B., 2009. Thirty years of breeding walnut in China. *ISHS Acta Hortic.* 861. <https://doi.org/10.17660/ActaHortic.2010.861.14>. VI International Walnut Symposium.
- Xu, Z.C., Jin, Y.C., Milne, R.I., et al., 2020. Development of 32 novel microsatellite loci in *Juglans sigillata* using genomic data. *Appl. Plant Sci.* 8, e11328. <https://doi.org/10.1002/aps3.11328>.
- Yuan, X.Y., Sun, Y.W., Bai, X.R., et al., 2018. Population structure, genetic diversity, and gene introgression of two closely related walnuts (*Juglans regia* and *J. sigillata*) in southwestern China revealed by EST-SSR markers. *Forests* 9, 646. <https://doi.org/10.3390/f9100646>.
- Zhang, F., Batley, J., 2019. Exploring the application of wild species for crop improvement in a changing climate. *Curr. Opin. Plant Biol.* 13, 1–5. <https://doi.org/10.1016/j.pbi.2019.12.013>.