Journal of Animal & Plant Sciences, 33(2): 2023, Page: 357-366
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2023.2.0626

GENETIC RELATIONSHIPS AMONG 15 *XANTHOCERAS SORBIFOLIUM* CULTIVARS USING SSR MARKERS AND PHENOTYPIC TRAITS

Y. L. Wang*,[1,2], L. Z. Zhu[2] and Y. Li[1,*]

[1]College of Forestry, Gansu Agriculture University, Lanzhou 753200, People's Republic of China
[2]State key laboratory of seedling bioengineering, Yinchuan 750001, People's Republic of China
Corresponding author's E-mail: liligsau@163.com; Wangyali523@126.com

ABSTRACT

Xanthoceras sorbifolium is an important plants used for biofuel production and woody edible oil, which is cultivated in northern China. A number of cultivars have been selected by breeders, but little information has been available about their genetic relationships. The genetic diversity among 15 *X. sorbifolium* cultivars were evaluated combined using the eight fruit and ten seed phenotypic traits and twenty-six SSR loci. Significant variations in all phenotypic traits studied were observed. The mean coefficient of variation for phenotypic traits was 11.86%, and seed traits (12.09%)> fruit traits (11.58%). Based on the phenotypic traits 14 cultivars (one cultivar was fruitless) were clustered in three main groups and five subgroups, and reasonable utilization of each subgroup was suggested. According the twenty-six SSR loci 97 genotypes and 89 alleles were detected. The mean value of number of alleles, polymorphism information content, gene diversity and Shannon's information index (I) were 3.4231, 0.4375,0.3104 and 0.4760, respectively, indicating that the genetic diversity was medium among 15 cultivars of *X. sorbifolium*. Based on genetic distance, the 15 cultivars were classified into three main groups with a coefficient of variation of 0.64. These results can provide a groundwork for future cultivar identification and breeding programs in *X. sorbifolium*.

Key words: *X. sorbifolium*, Genetic relationship, SSR markers, Morphological traits

Published first online December 18, 2022 Published final March 24, 2023

INTRODUCTION

X. sorbifolium belongs to Sapindaceae family and the only genus *Xanthoceras* that is a deciduous shrub or small tree and widely distributed in north China (Yu *et al*. 2017). *X. sorbifolium* offers ecological, economical, ornamental, and pharmaceutical benefits (Zhou *et al*. 2019). It is extremely cold-hardy and drought resistant; in addition, it could be used for new landscaping and as an ornamental tree due to its beautiful and colorful flowers (Wang *et al*. 2017; Zhou *et al*. 2018). Moreover, its kernels contain 55%–58% oil rich in unsaturated fatty acids (85%–93%), and this oil is considered to be a high-end woody edible oil (Zhang *et al*. 2010; Li *et al*. 2012; Ruan *et al*. 2017). Furthermore, a variety of bioactive compounds are found in it, that can be used as traditional Chinese medicines (Rong *et al*. 2016). Due to its important value, the cultivation of *X. sorbifolium* developed rapidly in recent years. Up to the year 2020, more than 5 × 10^5 ha of *X. sorbifolium* were cultivated (Shen *et al*. 2017 a; Yu *et al*. 2017).

There is tremendous phenotypic variation in *X. sorbifolium*, due to its wide distribution, ability to grow in a wide range of environments, and long-term extensive cross-pollination. Phenotypic diversity is the most convenient and traditional method to assess genetic diversity. (Schaal *et al*. 1991). Researchers have used relatively stable phenotypic traits such as leaf, flower, and fruit traits to measure genetic variation and assess genetic diversity of *X. sorbifolium*. (Shen *et al*. 2017 a; Yu *et al*. 2017). However, because the phenotypic traits are strongly affected by environmental factors they are not sufficient to reveal the genetic diversity. Whereby, complementarily, DNA polymorphism analysis might be used for reliable characterization of genetic variability and diversity among cultivars. Currently, the random amplified polymorphic DNA (RAPD), simple sequence repeats (SSR) markers, and other molecular techniques based on DNA markers have been used to analyze genetic diversity, identify, and characterize populations of *X. sorbifolium* and other plant species (Masanori, 2011; Sheng *et al*. 2010; Xie *et al*. 2011; Ao *et al*. 2016; Sun *et al*. 2019; Tam *et al*. 2020).

The number of cultivars has greatly increased over the last five years, up until 2021, about 40 cultivars of *X. sorbifolium* had been registered. To protect the rights and interests of plant breeders, the study of genetic relationships and genetic diversity of cultivars is essential and urgently needed. Therefore, the purpose of this work was to (1) evaluate the genetic relationships and genetic diversity among 15 cultivars of *X. sorbifolium*; (2) conduct a comparative analysis of SSR markers and fruit and seed phenotypic traits; (3) finally, to provide a groundwork for future cultivar identification and breeding programs in *X. sorbifolium*.

MATERIALS AND METHODS

Plant materials: The experiment was carried out at the Yinchuan Botanic Garden in Yinchuan, Ningxia, China (38°28'N; 106°22'E) from 2018 to 2019. The climate in this area is temperate continental, the average temperature is 10.1°C. The average annual precipitation and the average annual evapotranspiration is 181.2 mm and 1,882.5 mm, respectively. The soil is fine sand, and the pH is 9.07. In all, 15 cultivars/accessions were examined (Table 1), all of which were cultivated at the Yinchuan Botanic Garden. The scions of cultivars were introduced and grafted in 2015-2017; at least 30 trees of each cultivar were grafted. The rootstock was grown from seedlings of *X. sorbifolium* planted in 2010 and the seeds were harvested from a 12-year-old tree at the Yinchuan Botanic Garden as well.

Table 1. Details of plant sources and main characteristic of 15 *X. sorbifolium* cultivars.

No.	Cultivars	Origin	Main characteristics
1	Zs4	Liaoning	puberulent, white flower, leaflets lanceolate and leathery, high rate of fruit-setting
2	Zs9	Liaoning	growing vigorously and strong branch ability, white flower, high rate of fruit-setting, cylindrical fruit with prismatic
3	Smcbg	Shaanxi	puberulent, double flower, fruitless
4	Jhg	Inner Mongolia	red flower, broad-ovate leaflets, low rate of fruit-setting, big seeds
5	Smthg	Gansu	pink flower, high rate of fruit-setting, obovate fruit
6	Smjzg	Gansu	red flower, broad-ovate leaflets, cylindrical fruit, big seeds
7	Smjfg	Ningxia	pink flower, petals fringed, oval leaflets, peach shaped fruit
8	Smwgg	Ningxia	white flower, high rate of bisexual flower and fruit-setting, spheroid fruit, late phenological
9	N923	Ningxia	white flower, high rate of fruit-setting, spheroid fruit
10	N920	Ningxia	white flower, high rate of fruit-setting, spheroid fruit
11	N917	Ningxia	white flower, with purple patches at base of petals, spheroid fruit, big seeds
12	N916	Ningxia	white flower, media rate of fruit-setting, cylindrical fruit
13	Gs	Shanxi	white flower, high rate of fruit-setting, long cylindrical fruit
14	Gl	Shanxi	white flower, high rate of fruit-setting, spheroid fruit
15	Yx02	Shanxi	white flower, high rate of fruit-setting, spheroid fruit

Fruit and seed phenotypic traits: Using the sample quartiles, we randomly selected 30 fruits and 30 seeds from each cultivar. Fruit morphological traits, including fruit transverse diameter (FTD/mm), fruit longitudinal diameter (FLD/mm), fruit weight (FW/g), seed weight of single fruit (SWSF/g), fruit shell thickness (FST/mm) and seed number of single fruit (SNSF/No.); and seed morphological traits, including seed transverse diameter (STD/mm), seed longitudinal diameter (SLD1/mm), seed lateral diameter (SLD2/mm), seed weight (SW/g), kernel weight (KW/g), seed shell weight (SSW/g), seed shell thickness (SST/mm), and seed 100-grain weight (SHGW/g) were measured using an electronic balance and a vernier caliper. Fruit shape indexes (FSI) and seed shape indexes (SSI) were calculated by dividing the transverse by the longitudinal diameter in each case. To calculate seed percentage (SP/%) and kernel percentage (KP/%), divide seed weight by fruit weight, and kernel weight by seed weight, respectively.

DNA extraction: Young, fresh leaves from each cultivar were used for DNA extraction. A Takara MiniBEST plant genomic DNA extraction kit (Dalian, China) was used to extract the genomic DNA from the samples that were frozen in liquid nitrogen and stored at -80 °C. The concentration of DNA was adjusted to ≥ 20 ng/µL as per spectrophotometric estimations conducted at $1.70 \leq 260/280 \leq 2.30$; $260/230 \geq 1.50$.

SSR marker amplification by polymerase chain reaction (PCR): A total of 38 highly polymorphic genomic SSR markers developed by (Bi *et al*. 2014) were used (Table S1) and synthesized by Beijing Capital Bio Technology Co., LTD. The PCR marker amplification was performed in 10 µL containing 4 µL of 50 ngµL^{-1} DNA template, 2 µL of 1.25 µmolL^{-1} primer, and 4 µL of 2× PCR mix. The following PCR protocol was used: preliminary denaturing step at 94 °C for 5 min,12 cycles of denaturing at 94 °C for 30 s; annealing at 60 °C for 45 s, and an extension at 72 °C for 45 s (each cycle decrease 0.8 °C); 12 cycles of denaturing at 94 °C for 30 s; annealing at 50 °C for 45 s; and extension at 72 °C for 45 s; and the final extension step at 72 °C for 10 min. PCR products were separated on 6% polyacrylamide gels

(PAGE) at 200 V. Silver nitrate was used to stain the gels to check the DNA banding patterns (Ji *et al*. 2007).

Data analysis: Variance analysis, multiple comparison, and cluster analysis were performed using SPSS 20.0. SSR alleles were determined manually by comparison with the standard marker size, fragment lengths were rounded to the nearest whole number for analysis; and genotypes were scored with harbored different allele combinations. The amplified products were scored as present “1” or absent “0”. The coefficient of variation (CV/%) was calculated by dividing the standard deviations by the average. The PowerMarker V3.25 software (Liu *et al*. 2005) was used to analyze major allele frequency, observed allele number, gene diversity, Shannon information index, and polymorphic information content (PIC). The cluster analysis was performed using the unweighted pair group method with arithmetic mean (UPGMA) based on Nei’s (1983) genetic distance.

RESULTS AND DISCUSSION

Fruit and seed phenotypic traits: Variations in phenotypic traits due to genotype and environment interactions (Santos *et al*. 2012). Although the 15 *X. sorbifolium* cultivars studied herein were planted in the same area and grown under the same conditions, the cultivars still showed significant differences with respect to the fruit and seed traits under study. Eight fruit (Table 2) and ten seed (Table 3) phenotypic traits were measured for 14 of the 15 cultivars under study, as cultivar Smcbg was fruitless and characterized by double flowers. Cultivar Gl showed the lowest values of FTD (47.66 mm), FW (30.26 g), FST (4.54 mm), SLD1 (12.51 mm), SLD2 (9.36 mm), SW (0.67 g), KW (0.37 g), SSW (0.31 g), and SHGW (63.62 g); additionally, Gl showed the greatest values for SP (60.36%) and KP (54.74%). In turn, cultivar Smwgg showed the lowest values for FLD (48.63 mm), SWSF (14.02 g), and STD (11.65 mm); it also showed the largest values of FSI (1.15). Cultivar Jhg showed the lowest values for FSI (0.76) and SNSF (7.27); Jhg showed the greatest values for FLD (79.96 mm), SW (1.38 g), and SHGW (131.16 g). Cultivar Smthg showed the largest values for SLD1 (16.49 mm), SLD2 (11.81 mm), and KW (0.71 g), while cultivar Smjzg registered the largest values for FST (7.98 mm), SSI (1.07), SSW (0.73 g), and SST (1.56 mm), whereas it showed the lowest values for KP (44.03%). Cultivar N923 showed the largest values for SNSF (23.80), while cultivar N917 showed the largest values for FTD (64.52 mm), FW (76.95 g), and STD (15.24 mm). Finally, cultivar Yx02 showed the greatest values for SWSF (30.74 g).

Table 2. Fruit phenotypic traits for 14 of the 15 cultivars of *X. sorbifolium*.

Cultivars	FTD	FLD	FSI	FW	SWSF	SP	FST	SNSF
Zs4	55.04 ± 3.39^{d}	56.66 ± 2.63^{e}	0.97 ± 0.05^{cd}	47.21 ± 5.38^{e}	25.51 ± 3.11^{b}	54.51 ± 7.71^{b}	6.33 ± 0.42^{d}	18.53 ± 2.92^{cd}
Zs9	61.37 ± 2.59^{bc}	66.11 ± 5.05^{cd}	0.93 ± 0.07^{de}	70.22 ± 8.16^{abc}	29.04 ± 3.81^{a}	41.33 ± 2.06^{cde}	6.80 ± 0.56^{c}	17.40 ± 2.35^{d}
Jhg	60.85 ± 4.60^{c}	79.96 ± 4.98^{a}	0.76 ± 0.06^{g}	57.49 ± 14.07^{d}	15.30 ± 4.29^{d}	26.54 ± 3.24^{gh}	7.91 ± 0.48^{a}	7.27 ± 2.28^{g}
Smthg	58.87 ± 1.72^{c}	59.22 ± 2.65^{e}	1.00 ± 0.03^{bc}	65.16 ± 6.18^{cd}	25.11 ± 2.48^{b}	38.58 ± 2.32^{e}	7.47 ± 0.43^{b}	19.53 ± 1.60^{bcd}
Smjzg	63.84 ± 4.20^{ab}	73.84 ± 6.05^{b}	0.87 ± 0.03^{f}	68.27 ± 12.21^{bc}	19.96 ± 3.42^{c}	29.57 ± 4.06^{g}	7.98 ± 0.93^{a}	11.67 ± 2.53^{f}
Smjfg	55.69 ± 4.39^{d}	63.88 ± 4.38^{d}	0.87 ± 0.03^{f}	42.31 ± 11.15^{ef}	18.65 ± 5.38^{c}	44.03 ± 6.22^{c}	5.68 ± 0.41^{f}	10.07 ± 3.13^{f}
Smwgg	55.65 ± 3.67^{d}	48.63 ± 2.11^{f}	1.15 ± 0.10^{a}	36.86 ± 7.71^{fg}	14.02 ± 1.89^{d}	38.87 ± 5.55^{e}	5.64 ± 0.53^{f}	18.60 ± 2.38^{cd}
N923	53.32 ± 3.74^{d}	56.06 ± 4.17^{e}	0.95 ± 0.05^{cde}	45.94 ± 11.09^{cde}	19.25 ± 3.68^{c}	42.57 ± 4.37^{cd}	5.82 ± 0.37^{f}	23.80 ± 3.69^{a}
N920	54.59 ± 2.31^{d}	57.74 ± 3.68^{e}	0.95 ± 0.06^{cde}	41.03 ± 5.71^{cdef}	14.06 ± 2.02^{d}	34.47 ± 3.99^{f}	5.90 ± 0.42^{ef}	17.80 ± 2.34^{d}
N917	64.52 ± 5.77^{a}	68.61 ± 7.90^{c}	0.95 ± 0.09^{cde}	76.95 ± 18.71^{a}	15.18 ± 4.91^{d}	19.41 ± 3.10^{i}	7.82 ± 0.52^{ab}	10.80 ± 3.90^{f}
N916	59.91 ± 5.20^{c}	66.23 ± 5.67^{cd}	0.91 ± 0.05^{ef}	71.56 ± 16.30^{abc}	17.95 ± 5.81^{c}	24.70 ± 5.27^{h}	6.20 ± 0.30^{de}	14.53 ± 4.42^{e}
Gs	59.20 ± 2.44^{c}	57.80 ± 3.36^{e}	1.03 ± 0.07^{b}	58.77 ± 7.34^{d}	23.13 ± 2.47^{b}	39.58 ± 3.56^{de}	6.76 ± 0.60^{c}	19.40 ± 2.26^{bcd}
Gl	47.66 ± 3.24^{e}	50.17 ± 3.17^{f}	0.95 ± 0.06^{cde}	30.26 ± 4.43^{f}	18.20 ± 2.49^{c}	60.36 ± 3.66^{a}	4.54 ± 0.46^{g}	21.33 ± 2.82^{b}
Yx02	61.59 ± 2.91^{bc}	63.57 ± 3.32^{d}	0.97 ± 0.04^{cd}	76.19 ± 7.88^{ab}	30.74 ± 3.62^{a}	40.39 ± 3.10^{cde}	7.00 ± 0.25^{c}	20.07 ± 2.02^{bc}

means	54.14± 15.53	57.90± 17.98	0.88± 0.26	52.55± 22.53	19.07± 8.01	38.21± 11.50	6.12± 1.95	15.39± 6.67

Table 3. Seed phenotypic traits for 14 of the 15 cultivars of *X. sorbifolium*.

Cultivars	STD	SLD1	SLD2	SSI	SW	KW	KP	SSW	SST	SHGW
Zs4	13.43± 1.27^{b}	14.73± 1.00^{abc}	10.76± 0.78^{b}	0.92± 0.11^{de}	1.04± 0.23^{bcd}	0.47± 0.12^{bc}	45.24± 5.26^{ef}	0.57± 0.11^{bc}	1.17± 0.14^{bc}	93.93± 1.18^{de}
Zs9	13.43± 0.91^{b}	14.80± 1.24^{ab}	10.65± 0.93^{bc}	0.91± 0.08^{de}	1.09± 0.20^{b}	0.51± 0.11^{bc}	46.19± 4.82^{def}	0.59± 0.09^{b}	1.28± 0.11^{b}	98.95± 1.94^{d}
Jhg	15.22± 1.09^{a}	15.33± 0.96^{a}	11.76± 1.16^{a}	0.99± 0.07^{bc}	1.38± 0.27^{a}	0.68± 0.17^{a}	49.39± 6.21^{bcd}	0.70± 0.14^{a}	1.09± 0.15^{bcd}	131.16± 3.34^{a}
Smthg	14.95± 1.56^{a}	16.49± 2.12^{a}	11.81± 1.08^{a}	0.92± 0.12^{cde}	1.34± 0.35^{a}	0.71± 0.19^{a}	53.26± 1.65^{ab}	0.62± 0.16^{b}	1.21± 0.15^{b}	126.96± 5.08^{ab}
Smjzg	15.10± 0.89^{a}	14.13± 0.56^{bcd}	12.12± 1.00^{a}	1.07± 0.09^{a}	1.26± 0.13^{a}	0.55± 0.10^{b}	44.03± 5.57^{f}	0.73± 0.07^{a}	1.56± 0.22^{a}	120.25± 6.20^{b}
Smjfg	12.82± 1.13^{bcd}	13.52± 0.72^{de}	10.08± 0.79^{bcd}	0.95± 0.09^{cde}	0.87± 0.15^{e}	0.43± 0.06^{cd}	50.25± 5.40^{bc}	0.43± 0.11^{d}	0.95± 0.14^{ef}	82.34± 4.02^{g}
Smwgg	11.65± 0.99^{e}	13.23± 1.13^{def}	9.87± 1.08^{bcd}	0.88± 0.04^{ef}	0.72± 0.21^{f}	0.38± 0.11^{d}	52.94± 4.95^{ab}	0.35± 0.11^{e}	0.91± 0.15^{fg}	72.29± 6.23^{h}
N923	11.66± 1.05^{e}	12.89± 1.06^{ef}	9.79± 0.84^{cd}	0.91± 0.07^{def}	0.71± 0.18^{f}	0.37± 0.09^{d}	52.81± 4.81^{ab}	0.32± 0.09^{e}	0.78± 0.13^{h}	68.92± 6.42^{hi}
N920	11.94± 2.08^{de}	14.19± 2.36^{bcd}	10.21± 2.74^{bcd}	0.84± 0.07^{f}	0.98± 0.29^{bcde}	0.49± 0.16^{bc}	49.67± 2.32^{bcd}	0.48± 0.14^{d}	0.95± 0.17^{ef}	80.64± 2.24^{g}
N917	15.24± 1.33^{a}	14.99± 0.97^{b}	11.75± 0.64^{a}	1.02± 0.10^{ab}	1.27± 0.18^{a}	0.65± 0.11^{a}	50.95± 3.41^{abc}	0.62± 0.09^{b}	1.05± 0.14^{de}	109.46± 2.92^{c}
N916	12.58± 0.67^{bcd}	13.78± 0.71^{cde}	10.56± 0.50^{bc}	0.92± 0.05^{de}	0.90± 0.11^{de}	0.44± 0.06^{cd}	48.68± 2.55^{cde}	0.45± 0.05^{d}	0.97± 0.14^{def}	90.80± 7.66^{ef}
Gs	12.43± 0.95^{cde}	14.76± 0.88^{abc}	10.02± 0.96^{bcd}	0.84± 0.05^{f}	0.93± 0.14^{cde}	0.47± 0.09^{bc}	51.05± 9.34^{abc}	0.45± 0.06^{d}	1.17± 0.15^{bc}	84.55± 2.16^{fg}
Gl	11.96± 0.52^{de}	12.51± 1.10^{f}	9.36± 0.63^{d}	0.97± 0.11^{bcd}	0.67± 0.06^{f}	0.37± 0.05^{d}	54.74± 2.72^{a}	0.31± 0.02^{e}	0.83± 0.12^{gh}	63.62± 0.66^{i}
Yx02	12.91± 0.97^{bc}	14.58± 1.26^{bc}	10.15± 0.82^{bcd}	0.89± 0.10^{ef}	1.07± 0.16^{bc}	0.54± 0.10^{b}	50.06± 2.61^{bc}	0.51± 0.07^{cd}	1.22± 0.23^{b}	106.45± 0.82^{c}
means	13.24± 1.73	13.33± 3.88	9.93± 2.98	0.87± 0.25	0.95± 0.38	0.47± 0.20	46.62± 13.59	0.51± 0.16	1.01± 0.36	95.02± 21.24

The Data are reported as mean±standard deviation; Each column with the same letter indicates that the differences are not statistically significant (P<0.05).

The CV values of the morphological traits was shown in (Fig.1), among the 8 fruit phenotypic traits studied, SNSF (18.89%) showed the largest CV, while the CV for SWSF (18.67%), FW (17.48%) were higher than the mean of CV for fruit (11.58%). And among10 seed phenotypic traits the KW (21.44%) showed the largest CV, while the CV for SW (18.80%), SST (14.32%) were higher than the mean (12.09%) of CV for seed. The mean CV of seed characteristics (12.09%) was higher than that for the fruit characteristics (11.58%). The mean CV of all 14 cultivars was 11.86%. The CV values for all 14 cultivar ranked as follows: N920 (14.61%)> Jhg (13.75%)> Smwgg (13.72%)> Smjfg (13.58%)> N917 (13.56%)> N923 (13.42%)> N916 (12.23%)> Zs4 (11.30%)> Smjzg (10.93%)> Smthg (10.74%)> Gs (10.23%)> Zs9 (9.98%)> Yx02 (9.25%)> Gl (8.77%). This broad phenotypic variability in fruit and seed characteristics was previously reported (Zhang *et al.* 2019; Yu *et al.* 2017; Chai *et al.* 2013; Shen *et al.* 2017 b). The morphological variability observed in fruits and seeds among the cultivars included in this study was induced by the long term use of extensive cross-pollination, which enables a more feasible and convenient selection of breeds.

Clustering based on phenotypic traits: The phenotypic traits of 14 cultivars (1 cultivar out of 15 Smcbg was fruitless) of *X. sorbifolium* were clustered using UPGMA cluster analysis (Fig. 2). The14 cultivars were clustered in three main groups and five subgroups mainly according to fruit and seed size and weight, and no significant correlation between the classification of groups and geographical origin was found. The first group included seven cultivars and was divided into subgroups Ⅰ and Ⅱ ; the subgroup Ⅰ included four cultivars (Zs9, Yx02, N917, and N916) with large fruits and showed higher

FTD, FLD, and FW. The subgroup Ⅱ included three cultivars, namely, Jhg, Smjzg, and Smthg with large seeds and higher FST, STD, SLD1, SLD2, SW, KW, SSW, SST and SHGW. The second group included six cultivars and was divided into subgroups Ⅲ and Ⅳ, subgroup Ⅲ comprises three cultivars (i.e., Smjfg, Gs, and Zs4), which showed the average level of most of these traits, whereas subgroup Ⅳ comprised three cultivars, Smwgg, N923 and Gl, all with small fruits and seeds and showing the lowest level for most of the traits described. The third group included only one cultivar, N920, which showed lower values for all these traits. Moreover, fruit weight (FW) and seed percentage (SP) are reportedly important seed yield components (Chai *et al*. 2013); thus the cultivars in subgroup Ⅰ (Zs9, Yx02, N917, and N916) with large fruits and higher FTD, FLD, and FW were expected in high-yield. It facilitates the rational use of cultivars in the future.

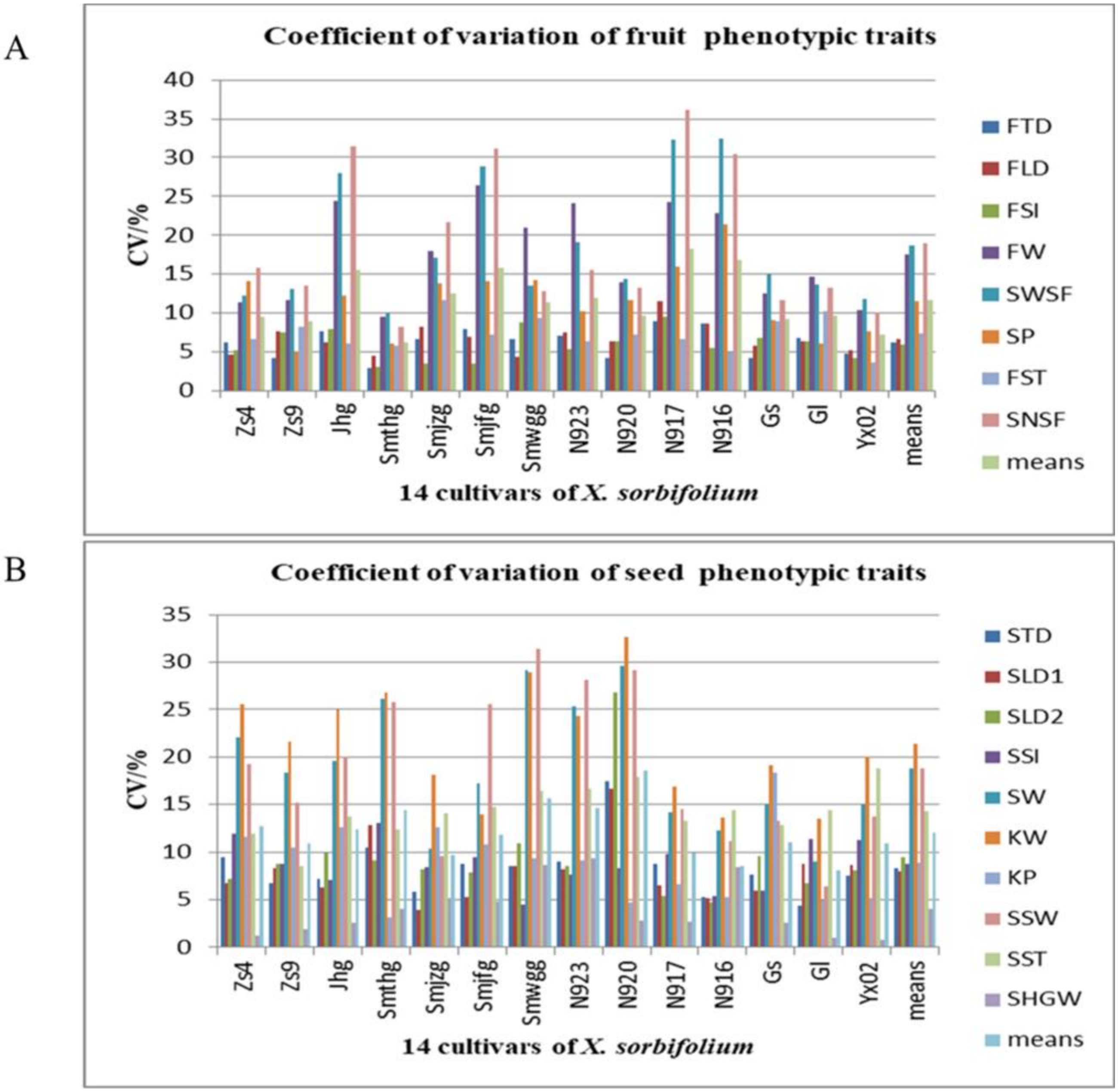

Figure 1. A. Variation coefficient bar chart of eight fruit phenotypic traits among 14 cultivars of *X. sorbifolium*(%); B. Variation coefficient bar chart of ten seed phenotypic traits among 14 cultivars of *X. sorbifolium*(%).

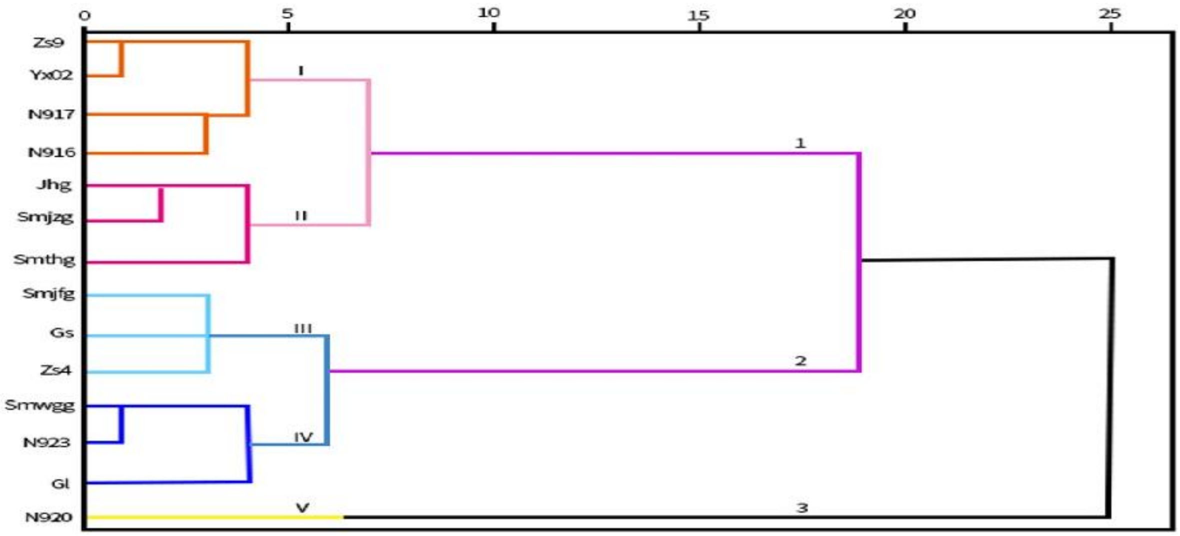

Figure 2. Clustering diagram for14 cultivars of *X. sorbifolium* based on eight fruit and ten seed phenotypic traits

Polymorphism detected by SSRs: There was considerable variation in amplified fragment patterns using 26 out of 38 primer pairs, 12 yielded no products, so they are not included in the result. The banding patterns were markedly different in amplifications using 26 primer pairs (Table 4). There were 97 genotypes and 89 alleles detected and the calculated proportion of polymorphic loci was 91.75%. Genotypes per locus ranged from 2 (QXR639, QXHB116, QBLB1, QXH002, QXH282, QXH555, B57) to 10 (QXH177), with a mean of 3.7308. The number of alleles (Na) ranged from 1 (QXH459, QBLB1, QXR231, QXH555, B57) to 10 (QXH274) with a mean of 3.4231, which was smaller to what was used by the researchers who first reported these microsatellite primer sequences, they used 38 primers in 36 plants representing two populations of *X. sorbifolium*, and the average number of alleles per locus was 6 (Bi *et al*. 2014). This is mainly related to the samples chosen for this study were the cultivars selected from the primary distribution areas of *X. sorbifolium*. This sample represents, therefore, a mini subset of *X. sorbifolium*, which could explain the relatively lower alleles and heterozygosity. Further, this value was much smaller than the value found in ISSR and RAPD studies (Wang *et al*. 2012; Chanhoon *et al*. 2016). But the average of alleles was similar to those previously reported (Shen *et al* 2017 *a*; 2017 *b*) (3.35 and 3.48). Genetic diversity is a component of all biological diversity, as it is an indicator of population genetic variation, mainly reflecting the evolution of genetic differences between different loci (Xu *et al*. 2020). As genetic diversity rises, organisms are more likely to adapt to diverse environments, and DNA sequences variations are the primary mechanisms behind this diversity (Qiao *et al*. 2020). Nei's gene diversity and Shannon's information index are common indicators represent genetic diversity of populations. In this study the gene diversity ranged from 0.1286 (QXH274) to 0.4995 (QXHS028), with an average of 0.3104; the Shannon's information index (I) ranged from 0.2512 to 0.6926, with an average of 0.4760. These results were smaller than the previous research (Shen *et al*. 2017 *b*), the gene diversity and Shannon's information index among 11 population of *X. sorbifolium* were 0.50 and 0.79. Allele frequencies are indicator of genetic diversity in a population which indicate the frequency of gene variants (zou *et al*. 2020). The major allele frequency ranged from 0.0691 (QBLB1) to 0.6884 (QBLB58) with an average of 0.2575. PIC values, which reflect allelic diversity and frequency across genotypes, usually classified as low (PIC < 0.25), medium (0.5 > PIC > 0.25) or high (PIC > 0.5) (Botstein *et al*.1980). In this study, PIC values ranged from 0.0691 (QBLB1) to 0.8647 (QXHS028), with an average of 0.4375, which was medium informative. There are ten marker showed a high PIC, and sixteen showed low or medium PIC. In previous study, the mean PIC values of *X. sorbifolium* were 0.47 (Bi *et al*. 2014), 0.49 (Shen *et al* 2017 *a*), and 0.51 (Shen *et al* 2017 *b*), respectively, which were also medium or high informative, supporting that the PIC values being moderate informative is acceptable. The most polymorphic locus was QXHS028 which exhibited high levels of genetic variation. The proportion of polymorphic loci, the allele number, the gene diversity, Shannon's information index and the PIC value were all similar or smaller than those reported in previous studies indicating a medium informative degree of genetic diversity among 15 cultivars of *X. sorbifolium*.

Clustering based on SSR loci: The genetic distance among the 15 cultivars (Table 5) ranged from 0.4494 (between Smjfg and N923) to 0.8621 (between Zs4 and N923), with a mean value of 0.6603. Using Nei's genetic distance coefficient (Nei's 1983), a dendrogram (Fig. 3) was generated by cluster analysis. With a coefficient of

Table 4. Genetic diversity among 15 cultivars of *X. sorbifolium* using 26 primer pairs of SSR.

Primer	Major allele frequency	Genotype No.	Allele No.	Gene diversity	Shannon's information index	PIC
QXHS028	0.1945	4.0000	4.0000	0.3110	0.4889	0.8647
QXHS043	0.2139	10.0000	10.0000	0.2790	0.4378	0.2564
QXH049	0.3334	4.0000	4.0000	0.3667	0.5462	0.3999
QXH177	0.2139	10.0000	10.0000	0.2790	0.4378	0.2831
QXH262	0.1985	4.0000	4.0000	0.1982	0.3284	0.6024
QXH274	0.2320	4.0000	4.0000	0.2360	0.3716	0.5438
QXH365	0.3322	4.0000	4.0000	0.4422	0.6346	0.5565
QXHS371	0.3671	3.0000	3.0000	0.3950	0.5731	0.1986
QXH643	0.3181	3.0000	3.0000	0.3416	0.4978	0.5582
QXR639	0.1548	2.0000	1.0000	0.2617	0.4310	0.2617
QXHB116	0.2441	2.0000	1.0000	0.3690	0.5557	0.369
QBRB203	0.4655	2.0000	1.0000	0.4976	0.6908	0.4976
QBRS192	0.0691	3.0000	3.0000	0.1286	0.2512	0.2857
QBLB1	0.0691	2.0000	1.0000	0.1286	0.2512	0.1287
QBLB4	0.2139	7.0000	7.0000	0.2790	0.4378	0.3363
QBLB58	0.6884	2.0000	2.0000	0.4234	0.6138	0.429
QBLB65	0.3340	3.0000	3.0000	0.3543	0.5366	0.5296
QBLB72	0.2203	4.0000	4.0000	0.2594	0.4052	0.6377
QBLB51	0.2286	5.0000	5.0000	0.3162	0.4801	0.6476
QXH002	0.4836	2.0000	2.0000	0.4995	0.6926	0.5323
QXH083	0.3184	4.0000	4.0000	0.3966	0.5780	0.5195
QXH323	0.1451	4.0000	4.0000	0.2433	0.4049	0.3062
QXR231	0.1056	3.0000	1.0000	0.1889	0.3372	0.1889
QXH282	0.1834	2.0000	2.0000	0.2761	0.4408	0.841
QXH555	0.1835	2.0000	1.0000	0.2997	0.4767	0.2997
QBLB57	0.1835	2.0000	1.0000	0.2997	0.4767	0.2997
Mean	0.2575	3.7308	3.4231	0.3104	0.4760	0.4375

Table 5. Nei genetic distance among 15 cultivars of *X. sorbifolium*.

Cultivars	Zs4	Zs9	Smcbg	Jhg	Smthg	Smjzg	Smjfg	Smwgg	N923	N920	N917	N916	Gs	Gl	Yx02
Zs4	***														
Zs9	0.62	***													
Smcbg	0.83	0.62	***												
Jhg	0.73	0.69	0.75	***											
Smthg	0.79	0.70	0.80	0.65	***										
Smjzg	0.77	0.61	0.61	0.65	0.49	***									
Smjfg	0.69	0.58	0.61	0.56	0.61	0.57	***								
Smwgg	0.70	0.63	0.67	0.61	0.67	0.66	0.62	***							
N923	0.86	0.66	0.66	0.57	0.62	0.65	0.45	0.58	***						
N920	0.72	0.63	0.63	0.61	0.65	0.62	0.53	0.66	0.57	***					
N917	0.86	0.66	0.64	0.69	0.66	0.67	0.56	0.58	0.72	0.64	***				
N916	0.72	0.67	0.61	0.52	0.67	0.69	0.55	0.55	0.69	0.72	0.71	***			
Gs	0.83	0.67	0.72	0.63	0.70	0.78	0.50	0.55	0.55	0.65	0.66	0.63	***		
Gl	0.83	0.65	0.70	0.65	0.70	0.82	0.53	0.60	0.57	0.65	0.64	0.63	0.64	***	
Yx02	0.83	0.71	0.69	0.60	0.78	0.79	0.58	0.63	0.70	0.66	0.70	0.66	0.76	0.74	***

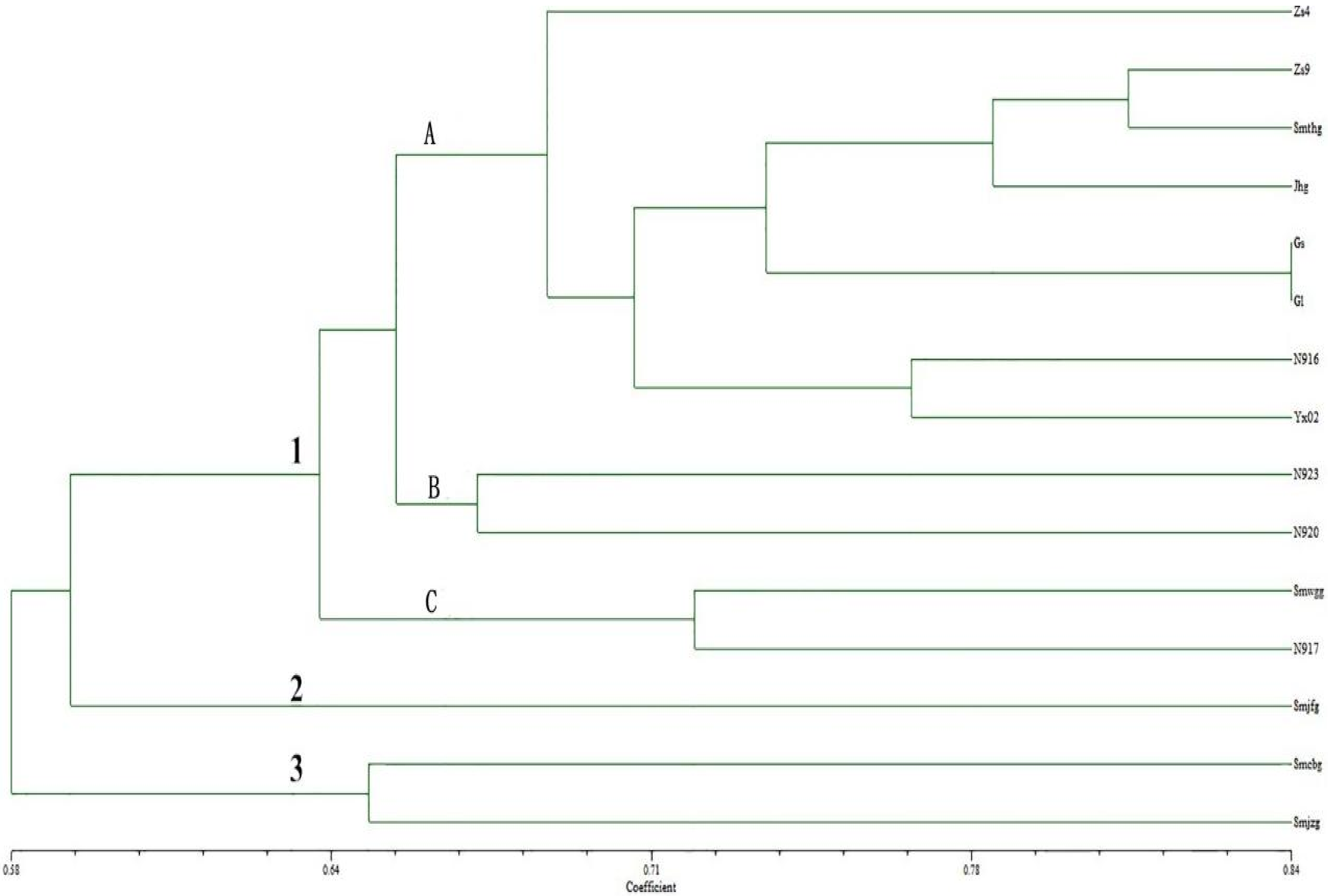

Figure 3. UPGMA cluster analysis among 15 cultivars of *X. sorbifolium* based on Nei's genetic distances coefficient of twenty six SSR loci.

variation of 0.64, the 15 cultivars were classified into three main groups. The first group included twelve cultivars was divided into subgroup A,B,C; the subgroup A contain eight cultivars (Zs4, Zs9, Smthg, Jhg, Gs, Gl, N916 and Yx02); the subgroup B includ two cultivars (N923 and N920); the subgroup C included Smwgg and N917 two cultivars. The second group included one cultivar Smjfg. Finally, the third group included cultivars Smcbg and Smjzg.

Although the results from cluster analysis based on SSR markers and phenotypic traits were not fully concordant, a part of the group of the genotypes was consistently identified by both methods; for example, the cultivars Smthg and Jhg were both ornamental cultivars, and they were clustered in the same subgroup based on phenotypic traits as well as on SSR makers. This result might be partly explained by the fact that phenotypic traits were determined by the environment and by multiple genes (Santos *et al*. 2012). Furthermore, phenotypic traits cannot completely represent the genetic relationships at the gene level. This indicates that the two analyses are complementary and that both are necessary for a reliable characterization of a gene bank (Fu *et al*. 2018).

Conclusions: Overall, we determined the genetic relationships among 15 *X. sorbifolium* germplasm accessions using 26 SSR markers, and combined these findings with data on fruit and seed morphological traits. Our results suggest that combined analyses of morphological and molecular data could be used to analyze genetic diversity and related of cultivars, thereby providing information for genetic improvement of *X. sorbifolium*, as well as preservation and evaluation. However, only 15 cultivars were used in this study, therefore, the need to identify the increasing number of commercial *X. sorbifolium* cultivars certainly further research.

Funding: This research was funded by the Natural Science Foundation of Ningxia Province, grant number 2022AAC03754; and the Conventional Science and Technology Assistance Program to Developing Countries, grant number KY202002011.

Author Contributions: Author Y. Li conceived and designed the project. Author Y. L. Wang and L. Z. Zhu executed the experiment and analyzed the data. All authors interpreted the data, critically revised the manuscript for important intellectual contents and approved the final version.

REFERENCES

Ao, Y., Ma, L. Y. (2016). Research progress in molecular biology of *Xanthoceras sorbifolia* Bunge. Chinese J. Molecular Plant Breeding. 14(12): 3570-3576.

Bi, Q. X., Guan, W. B. (2014). Isolation and characterisation of polymorphic genomic SSRs markers for the endangered tree *Xanthoceras sorbifolium* Bunge. Conservation Genet Resour. 6: 895-898.

Botstein, D., White, R. L., Skolnick, M., Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 32(1): 314-331.

Chai, C. S., Lu, J., Cai, G. J., Wang, S. Y., Qi, J. L., Wang, Z. T., Xue, R. (2013). Fruit phenotypic diversity and variation of *Xanthoceras sorbifolia* artificial population. Chinese J. Forest Research. 26(2): 181-191.

Chanhoon, A., Hyunseok, L., Jeonghoon, L., Eun, J. C., Li, Y. H., Jaeseon, Y. (2016). Analysis of genetic diversity and differentiation of artificial populations of yellowhorn (*Xanthoceras sorbifolium*) in China using ISSR markers. J. For. Res. 27(5): 1099-1104.

Fu, H. F., Liu, X. H., Chen, J. Y., Li, G. J. (2018). Genetic diversity analysis of *Capsicum* germplasm based on phenotypic traits With SSR markers. Chinese J. Nuclear Agricultural Sciences. 32(7): 1309 -1319.

Ji, W., Wang, L. X., Sun, H., Wang, M. Y., Zhao, C. P. (2007). Predigestion of wheat SSR analysis protocol. Chinese J. Agricultural Biotechnology. 15(5): 907-908.

Li, J., Fu, Y. J., Qu, X. J., Wang, W., Luo, M., Zhao, C. J., Zu, Y. G. (2012). Biodiesel production from yellow horn (*Xanthoceras sorbifolia* Bunge) seed oil using ion exchange resin as heterogeneous catalyst. Bioresour Technol. 108(3): 112-118.

Liu, K., Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker data. Bioinformatics. 21(9): 2128-2129.

Masanori, H., Tsukasa, N., Sono, K., Takayoshi, Y., Hiromichi, Y., Megumi, H., Susumu, Y., Masami, M. (2011). Strawberry cultivar identification based on hypervariable SSR markers. Breed Sci. 61(4): 420-425.

Qiao, D. Y., Wang, P., Wang, S. A., Li, L. F., Gao, L. L., Yang, R. T., Wang, Q., Li, Y. (2020). Genetic diversity analysis of *Lagerstroemia* germplasm resources based on SNP markers. Chinese J. Nanjing Forestry University (Natural Sciences Edition). 32(5): 1161-1172.

Rong, W., Sun, Z., Li. Q., Liu, R., Zhang, T., Wang, T., Yang, W., Li, Z., Bi, K. (2016). Characterization and simultaneous quantification of seven triterpenoid saponins in different parts of

Xanthoceras sorbifolia Bunge by HPLC-ESI-TOF. Anal Methods. 8: 2176-2184.

Ruan, C. J., Yan, R., Wang, B. X., Mopper. S., Guan, W. K., Zhang, J. (2017). The importance of yellow horn (*Xanthoceras sorbifolia*) for restoration of arid habitats and production of bioactive seed oils. Ecol Eng. 99(5): 504-512.

Sheng, D. C., Feng, Q. R., Li, F. L., Liu, Z. H. (2010). Pollination characteristics and fruit developing law of *Xanthoceras sorbifolia*. Chinese J. Liaoning Forestry Science Technology.4:1-4.

Santos, R. C., Pires, J. L., Correa, R. X. (2012). Morphological characterization of leaf, flower, fruit and seed traits among Brazilian *Theobroma* L. species. Genet Resour Crop Ev. 59(3): 327-345.

Schaal, B. A., Leverich, W. J., Rogstad, S. H. (1991). A comparison of methods for assessing genetic variation in plant conservation biology. In: Falk DA, Holsinger KE (eds) Genetics and conservation of rare plants. Oxford University Press, New York. 123-124.

Shen, Z., Duan, J., Ma, L. Y. (2017 a). Genetic diversity of *Xanthoceras sorbifolium* Bunge germplasm using morphological traits and microsatellite molecular markers. PLoS One. 12(6): e0177577.

Shen, Z., Zhang, K. Q., Ma, L. Y., Duan, J., Ao, Y. (2017 b). Analysis of the genetic relationships and diversity among 11 populations of *Xanthoceras sorbifolia* using phenotypic and microsatellite marker data. Electron J Biotechn. 26(3): 33-39.

Sun, C. W., Jia, L. M., Xi, B. Y., Liu, J. M., Wang, L. C., Weng, B. X. (2019). Genetic diversity and association analyses of fruit traits with microsatellite ISSRs in *Sapindus*. J. For. Res. 30(1): 193-203.

Tam, M. N., Duy, D. V., Duc, M. N., Hien, P. D., Long, K. P., Phuong, X. B. (2020). Microsatellite analysis reveals genetic diversity of the endangered species *Dipterocarpus dyeri.* J. For. Res. 25(3): 198-201.

Wang, Q., Yang, L., Ranjitkar, S., Wang, J. J., Wang, X. R., Zhang, D. X., Wang, Z. Y., Huang, Y. Z., Zhou, Y. M., Deng, Z. X., Yi, L. B., Luan, X. F., Wen, B. G. (2017). Distribution and in situ conservation of a relic Chinese oil woody species *Xanthoceras sorbifolium* (yellowhorn). Can. J. For. Res. 47(8): 1450-1456.

Wang, L. W., Wu, S. Q., Kou, Y. Y., Luo, G. J. (2012). RAPD analysis of germplasm resources of *Xanthoceras sorbifolia*. Chinese J. Liaoning Forestry Science Technology. 20(1) : 11-13.

Xie, R. J, Zhou, J., Wang, G. Y., Zhang, S. M., Chen, L., Gao, Z. S. (2011). Cultivar identification and genetic diversity of Chinese bayberry (*Myrica rubra*) accessions based on fluorescent SSR markers. Plant Mol Biol Rep. 29(4): 554-562.

Xu, Q. Y., Zheng, J. H., Nie, H. T., Wang, Q. Z., Yan, X. W. (2020). Genetic diversity and population structure of *Meretrix petechialis* in China revealed by sequence-related amplified polymorphism markers. Peer J. 8: e8723.

Yu, H. Y., Fan, S. Q., Bi, Q. X., Wang, S. X., Hu, X. Y., Chen, M. Y., Wang, L. B. (2017). Seed morphology, oil content and fatty acid composition variability assessment in yellow horn (*Xanthoceras sorbifolium* Bunge) germplasm for optimum biodiesel production. Ind Crop Prod. 97(6): 425-430.

Zhang, Y., Ao, Y., Liu, J. F., Zhao, L. L. (2019). Variation pattern of fruit and seed characters and superior individual selection of *X. sorbifolium*. Chinese J. Northeast Forestry University. 47(9): 1-5.

Zhou, Q. Y., Cai, Q. (2018). The superoxide dismutase genes might be required for appropriate development of the ovule after fertilization in *Xanthoceras sorbifolium.* Plant Cell Rep. 37(2): 727-739.

Zhou, Q. Y., Cai, Q., Zheng, Y. R., Wu, Z. Y., Mao, J. F. (2019). Floral development and the formation of functionally unisexual flowers in *Xanthoceras sorbifolium* (Sapindaceae), a morphologically and romonoecious tree endemic to northern china. Trees Struct Funct. 09: 1-11

Zhang, S., Zu, Y. G., Fu, Y. J., Luo, M., Liu, W., Li, J., Efferth, T. (2010). Supercritical carbon dioxide extraction of seed oil from yellow horn (*Xanthoceras sorbifolia* Bunge) and its antioxidant activity. Bioresource Technol. 101(5): 2537-2544.

Zou, X., He, G., Wang, M. (2020). Genetic diversity and phylogenetic structure of four Tibeto-Burman-speaking populations in Tibetan-Yi corridor revealed by insertion deletion polymorphisms. Mol Genet Genom Med. 8(4): e1140.