Marker-assisted selection in *C. oleifera* hybrid population

Jin-Ling Feng¹, Ying Jiang¹, Zhi-Jian Yang¹, Shi-Pin Chen¹, Yousry A. El-Kassaby^{2*}, Hui Chen^{1*}

¹ College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China

² Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

*Corresponding authors: Yousry A. El-Kassaby: y.el-kassaby@ubc.ca; Hui Chen: zjchchenh@163.com

Abstract

Marker-assisted selection (MAS) is implemented to improve *Camellia oleifera* yield and fruit attributes for meeting China's increased demand for new varieties development. We conducted three-generational (G_0 , G_1 , and G_2) hybridization (diallel mating) and selection experiment and used Sequence-Related Amplified Polymorphism (SRAP). SRAP markers to investigate their utility in a MAS framework. The utilized SRAP markers were instrumental in hybrid authenticity and the identification of matroclinal or patroclinal inheritance presence, thus guiding mating pair selection and direction (their role as male or females). Across the studied 3 generations, estimates of genetic diversity parameters showed steady increase with percentage increase of (G_0 to G_1 and G_1 to G_2) 9.25 and 9.05: observed number of alleles; 3.12 and 7.80: means effective number of alleles; 12.35 and 22.34: Nei's gene diversity; and 14.21 and 21.77: Shannon's index), indicating lack of diversity reduction associated with selection. Estimates of genetic distance and their correlation with heterosis were useful in guiding selection of mating pairs for achieving the desired yield and fruit attributes (fruit diameter, height, weight, and index, peel thickness, number of seeds per fruit, seed weight per fruit, and seed rate). Most yield and fruit attributes exhibited high broad-sense heritability with increasing trend over generation intervals, indicating the increased potential of hybrid breeding for this species.

Keywords: : *C. oleifera*; hybridization; marker-assisted selection; selection; heritability; heterosis

Introduction

Camellia oleifera Abel. has more than 2,300 years of cultivation history and belongs to the woody edible oil plants group including olive, oil palm, and coconut (Wu et al. 2015). The species has been under cultivation in China, Vietnam, Myanmar, Thailand, Malaysia, and Japan (Ma et al. 2011). *C. oleifera* is a shrub with an entomophilous flower with a blooming season extends from autumn to early winter, followed by fruit ripening around October of the second year (Liao et al. 2014). *C. oleifera* oil is of high quality with numerous nutritional and health-care values (Zeb 2012). The FAO (Food and Agriculture Organization of the United Nations) has listed *C. oleifera* oil as the premium health-grade edible oil (Yang et al. 2016).

Since the 1950s, China has initiated an active *C. oleifera* germplasm collection and development for the production of stable high-yield varieties; however, most of the developed varieties suffered from low incremental yield improvement, uneven production level, and poor adaptability (Liang et al. 2010; Peng S et al. 2010; Qin et al. 2018). These attributes were considered as *C. oleifera* germplasm development bottleneck (Zhang et al. 2008; Peng S et al. 2010; Liang et al. 2010). The sound and rapid development of improved *C. oleifera* varieties has become an urgent task and a hybrid breeding approach was adopted as an effective avenue to broadening the genetic base permitting the selection and development of new superior clones. This became possible owing to the fact that *C. oleifera* is widely distributed and exhibits large regional differences throughout China (Huang et al. 2006).

Conventional *C. oleifera* breeding techniques are simple and cost effective through the strategy of "sexual selection and asexual use"; with the selection of families and clones

improved growth, fruit yield, stable production, and disease and insect resistance (Liang et al. 2010; Chen et al. 2011; Shi et al. 2017). However, conventional *C. oleifera* breeding methods suffer from: 1) relatively long breeding cycle (up to 10 years), 2) plants are highly heterogeneous with complex and unstable traits, and 3) unreliable early-stage quality assessment and low breeding efficiency (Huang et al. 2018). With the present and anticipated rapid development of *C. oleifera* oil market, conventional breeding is no longer meeting the industry needs/requirements.

Marker-assisted selection (MAS) is a selection method that capitalizes on the existing linkage between specific molecular markers and target attributes (i.e., morphological or functional genes), thus indirectly assist in faster and more accurate selection of desired individuals rather than selecting for the traits themselves (El-Kassaby 1982; Li et al. 2018; Bankole et al. 2017). Theoretically, MAS is expected to increase breeding efficiency, specifically for late expression traits where long-periods are required for these traits to manifest themselves. Additionally, it is anticipated that MAS is expected to be free from environmental influences as well complex genetic effects (Gamal El-Dien et al. 2015; Thistlethwaite et al. 2017 and 2019; Simko et al. 2018). MAS is suited for woody plants with long growth cycle and unstable early-stage performance (Ru et al. 2015; Ratcliffe et al. 2015; Igarashi et al. 2016; Kadam et al. 2016). Therefore, it is expected that *C. oleifera* would benefit from early-stage MAS selection through shorting and speeding the breeding cycles.

The availability of molecular markers made it possible to quantify the genetic relationship and diversity of *C. oleifera* germplasm; however, they have rarely been applied to the species MAS-framework (Wang et al. 2014; Cheng et al. 2018; Zhao X et al. 2017; Chen et al. 2016). Sequence-Related Amplified Polymorphism (SRAP) with their abundant polymorphic fragments have been frequently used along with morphological variability in screening specific DNA fragments associated with major *C. oleifera* economic traits (Peng S et al. 2011; Han et al. 2012; Wang et al. 2014). Here, utilizing MAS in a conventional *C. oleifera* hybridization breeding program we demonstrate the utility of MAS in improving fruit traits and the identification of new high-yield varieties.

Materials and Methods

Plant materials

In 1973, superior *C. oleifera* individuals with fruit yield of >2.5 kg/m² of live-crown area were selected from the species main producing areas in Fujian Province, China (Bei village (Ninghua County), ZhaoXing village (Datian County), and Fuan City (Duntou County) representing Northwestern, Central, and Northeastern Fujian Province, respectively). A total of 7 superior individuals (Min43, Min49, Min50, Min53, Min54, Ming56, and Min62) were selected from nearly 40 hectares of *C. oleifera* forest based on their 3-year average yield performance. These superior individuals were propagated through vegetative

cuttings to form 7 clonal seedlings. In 1981, using these superior individuals (G_0), artificial pollination was conducted resulting in a complete 7×7 diallel-mating crossing. In 1982, 531 trees representing the diallel crossing offspring (G_1) were planted in a complete randomized block design experiment using a single tree row-plot with 3×3 m spacing at the Fujian Minhou Tongkou Forest Farm (26° 05'N, 119° 17'E). At present, there are 392 progeny trees in representing the G_1 .

Over four consecutive years (2007 - 2011), fruit yield of the G_1 392 trees were assessed following the national selection criteria of *C. oleifera* (China's State Forestry Administration 2008). In 2011, 5 G_1 individuals (27-5, 41-4, 46-2, 51-2, and 56-2; where the first and the second numbers representing the block and plot, respectively), each representing one of the original 7×7 diallel crossing, were selected and used in an additional 5×5 complete diallel crossing to produce the second generation (G_2). A total of 20 cross combinations were carried out and all but one cross (56-2×41-4) successfully produced fruits that were harvested and phenotyped and the phenotypic data were used as the main indicators of performance's data analyses.

Artificial pollination method

Pollen was collected from mature flowers using clean brush and if not immediately used (i.e., the same day); it was placed in silica-gel dryer at room temperature for 3-5 days followed by storage at -4°C.

Female flowers at the full-bloom stage were selected at the top of branches and were emasculate with tweezers then isolated with paper bags (15×15 cm). Within 2-3 days after emasculation, a special mucus was secreted from the stigma indicating pollen receptivity time. Before pollen application, petals were gently removed to maximize stigma access. Using a brush, pollen was applied to receptive stigmas and immediately followed by replacing the pollination bags. Eight days after pollination, the petal colour faded and the stigma dried, demarking the time for pollination bags removal.

DNA extraction

Healthy leaves were washed with alcohol and their midribs were removed. Using liquid nitrogen, 2-3 g of clean leaves tissue were quickly ground into powder and were immediately transferred to 50 ml pre-cooled centrifuge tube, with 625 µl of β mercaptoethanol and 15 ml preheated to 65°C 2 % CTAB solution, shaken, bathed in 65°C water for 2 hrs. This was added to centrifuge tube with 10 ml chloroform-isoamyl alcohol (24:1), gently shaken for 10 minutes, and centrifuged at 4°C at 6,000 rpm for 10 minutes. The supernatant was transferred to a new centrifuge tube with 10 % of the supernatant volume of 10 % CTAB solution, gently shaken, then 10 ml chloroform-isoamyl alcohol (24:1) was added, gently shaken for 10 minutes and centrifuged for 10 minutes at 4°C at 6,000 rpm.

The supernatant was transferred to a new 50 ml centrifuge tube with 10 % of the supernatant volume of 3 mol / L NaAC solution (pH 5.2) and a pre-cooled at -24°C 110 % volume of isopropyl alcohol, gently shaken until flocculation appeared, then placed at -20°C for 30 minutes. Then the tube was

centrifuged at 4°C at 6,000 rpm for 10 minutes. After 3-5 minutes, 7 ml of high-salt TE and 5 µl of RNase (10 mg/ml) were added to completely dissolving the DNA in 55°C water. Then 7 ml chloroform-isoamyl alcohol (24:1) was added. The centrifuge tube was gently shaken for 10 minutes, centrifuged at 4°C and 6,000 rpm for 10 minutes. The supernatant was removed to a new centrifugal tube with 2 times the supernatant volume of pre-cooled ethanol, gently shaken until the white flocc (DNA) appeared, which was moved to 1.5 ml centrifuge tube. DNA was washed twice with 75 % ethanol solution, then with 100 % ethanol for 3-5 times, dried until the DNA became transparent. DNA was added to 50-100 µl TE buffer in 1.5 ml centrifuge tube and refrigerated at 4°C for immediate use while the remaining was stored in -20°C for further use.

DNA quality and concentration detection

DNA quality and concentration of all samples were evaluated using 1 µl of the fully dissolved DNA solution that was mixed with 1 µl of 6×Loading buffer, then loaded on 1.5 % TAE agarose gel with 15,000 bp DNA maker as a control and electrophoresed in 1×TAE buffer (voltage 150 V, current 120 mA) for 30 minutes. Gel was placed in ethidium bromide (EB) for 10 minutes and imaged under UV spectroscopy to determine its concentration, fragment size, degradation, and the presence of RNA and protein pollution levels.

DNA Dilution

The DNA was diluted using different levels. The different dilutions were amplified by PCR, detected by 6 % polyacrylamide gel for SRAP amplification and according to bands brightness and clarity, the appropriate dilution of DNA was selected.

SRAP reaction system

PCR reactions were performed in a 20 µl final volume which contained 0.2 µl Taq (5 U/µl, TaKaRa), 3 µl 10×buffer (Mg²⁺ Plus), 1.6 µl dNTPs mixture (2.5 mM), 1 µl forward primer (10 µM, Sangon Biotech Shanghai), 1 µl reverse primer (10 µM, Sangon Biotech Shanghai), 2 µl DNA, and ddH₂O 11.2 µl. Ten pairs of forward and reverse primers designed by Li and Quiros (2001), which combined 100 SRAP primers were used in the present study (Table 1). The combined primers were screened for polymorphism and stability. The PCR program was a pre-denaturation at 94°C for 5 minutes, denaturation at 94°C for 1 minutes, 35°C renaturation for 1 minute, extension at 72°C for 1 minute, with a total of 5 cycles; followed by 35 cycles at 94°C denaturation for 1 minute, 35°C renaturation for 1 minute, 72°C extension 1 minute and a final extension at 72°C for 5 minutes, stored at 4°C (Eppendorf Master cycler gradient PCR 5333, Germany). This program was used for all primer pairs (Xue et al. 2009; Ji et al. 2009).

Using 6 % acrylamide gel, the SRAP-PCR amplification product mixed with 1/6 volume of 6×Loading Buffer and 4 µl mixture was loaded onto 6 % acrylamide gel. Using 1×TBE electrophoresis buffer solution, at a constant voltage of 350 V and a current of 180 mA, electrophoresis was carried out in a vertical electrophoresis tank (DYCZ-28D, Beijing Liuyi Instrument Factory) for 2 hrs, then the acrylamide gel was fixed,

silver stained and wrapped in plastic wrap. After the acrylamide gel got dry, the bands were counted and scanned.

Table 1

SRAP random primers and their forward and backward sequences

Primer name	Forward primers sequence (5' TO 3')	Primer name	Reverse primers sequence (5' TO 3')
Me1	TGA gTC CAA ACC ggA TA	Em1	gAC TgC gTA CgA ATT AAT
Me2	TGA gTC CAA ACC ggA GC	Em2	gAC TgC gTA CgA ATT TGC
Me3	TGA gTC CAA ACC ggA AT	Em3	gAC TgC gTA CgA ATT GAC
Me4	TGA gTC CAA ACC ggA CC	Em4	gAC TgC gTA CgA ATT TgA
Me5	TgA gTC CAA ACC ggA Ag	Em5	gAC TgC gTA CgA ATT AAC
Me6	TgA gTC CAA ACC ggT AA	Em6	gAC TgC gTA CgA ATT gCA
Me7	TgA gTC CAA ACC ggT AC	Em7	gAC TgC gTA CgA AAT CAA
Me8	TgA gTc CAA ACC ggT Gc	Em8	gAC TgC gTA CgA ATT CTg
Me9	TTC Agg gTg GCC ggA Tg	Em9	gAC Tgc gTA CgA ATT CgA
Me10	TTC Agg gTg GCC ggA AT	Em10	gAC Tgc gTA CgA ATT CAg

Yield and fruit morphological index

Individuals' yields were measured for four consecutive years (2009 - 2012) and the morphological indexes were: fruit diameter, height, weight, and index (= fruit height / fruit diameter), peel thickness, number of seeds per fruit, seed weight per fruit and seed rate (= (seed weight per fruit / fruit weight) × 100 %). Fruit diameter and height were measured using Vernier caliper with 0.01mm accuracy.

Data analysis

The electrophoretic gel pattern was read manually, and each band was regarded as a one locus, bands with clear and reproducible pattern were used in the analysis. Bands with mobility similar to that of their primers, were recorded as "1" (no "0" bands were recorded). These data were converted to a binary data matrix and the UPGMA method of NTSYSpc software (2018) was used to calculate the genetic distance among samples (Salgotra et al. 2015; Zheng et al. 2017).

Broad sense heritability (H^2) was measured using analysis of variance component method (Fehr 1987). $H^2 = \sigma_g^2 / \sigma_p^2$; where, σ_g^2 = genetic variance (additive, dominance, and epistatic) and σ_p^2 = phenotypic variance (Parikh et al. 2016). Generation gain was calculated according to the formula Generation gain = (Indicator_{curr} - Indicator_{base}) / Indicator_{base} × 100 %, where Indicator_{curr} and Indicator_{base} represent the mean indicator of the current and base generations, respectively. Mid-parent heterosis (MPH), over high-parent heterosis (OPH) and below lower-parent heterosis (BPH) were used to measure the yield of *C. oleifera* heterosis. MPH (%) = (F₁ - MP) / MP × 100 %, OPH (%) = (F₁ - HP) / HP × 100 %, BPH (%) = (F₁ - LP) / LP × 100 %. In which F₁ was the average of the hybrid progeny traits, MP was the average of the two parental traits, HP was the character value of the higher of the parents, LP was the character value of the lower of the parents (Sharma et al. 2016; Ismaeel et al. 2018; Yang H et al. 2018; Yang J et al. 2018). Regression and cluster analyses were conducted using SPSS version 16.0.

Results

C. oleifera hybrid identification

To identify hybrid offspring authenticity, we used 17 pairs of SRAP primers which were selected from 100 pair of primers combinations using the 10 pairs of upstream and downstream primers (Table 1). These 17 pairs of SRAP primers produced clear, stable bands with high polymorphism (Table 2).

Table 2

SRAP polymorphism (# of total bands (TNB); # of polymorphic bands (NPB); and % of polymorphic bands (PPB) across primer combination for *C. oleifera* hybrid population

Primer	TNB	NPB	PPB
Me3-Em3	11	10	90.91
Me3-Em6	17	13	76.47
Me3-Em7	14	14	100.00
Me3-Em9	11	11	100.00
Me4-Em8	13	12	92.31
Me5-Em6	9	8	88.89
Me5-Em8	13	12	92.31
Me6-Em7	9	9	100.00
Me8-Em3	16	14	87.50
Me8-Em4	10	10	100.00
Me8-Em7	13	11	84.62
Me8-Em10	15	14	93.33
Me9-Em1	13	12	92.31
Me9-Em3	9	7	77.78
Me9-Em5	12	11	91.67
Me9-Em7	4	4	100.00
Me9-Em9	18	16	88.89
Total	207	188	90.82
Average	12.2	11.1	91.60

Single primer or primer combinations diagnosed parental and hybrid plants. According to the method of Huang et al. (2014) and Yu et al. (2016), hybrids were successfully determined if they share bands exclusively amplified in male parents. However, when single primers were used, hybrid identification success rate varied across primers and ranged between 37.50 and 87.50 % for Me3-Em8 and Me3-Em7, respectively, with an overall average of 60.29 % (Supplementary Table 1). Thus, the use of a single prime is deemed unreliable and therefore we used primer combinations which produced better results. For example, the primer combination of "Me3-Em7 + Me4-Em8" and "Me5-Em8 + Me8-Em7", identified all G_1 (5 individuals) and G_2 (19 individuals) offspring as hybrids (Supplementary Table 1).

SRAP polymorphism and genetic relationship among *C. oleifera* hybrids

SRAP polymorphism for the 17 primers was assessed using 31 *C. oleifera* accessions in a hybrid population (Table 2). The total number of bands (TNB), number of polymorphic bands (NPB), and percentage of polymorphic bands (PPB) were 207, 188, and 90.82 % for all primers, respectively (Table 2). TNB varied across primers ranged between 4 and 18 (average = 12.2

bands) while band size varied between 100 and 1,500 bp, NPB ranged between 4 and 16 (average = 11.1 bands), and PPB ranged between 76.47 and 100 % (average = 91.60 %) (Table 2).

Across the three generations (G_0 - G_2), Nei's gene diversities (H) varied between 0.259 and 0.491 (average = 0.404), Shannon's index (I) ranged from 0.412 to 0.684 (average = 0.585), and mean effective number of alleles (Ne) varied from 1.403 to 1.966 (average = 1.733) (Table 3). It is noteworthy to mention that Na , Ne , H , and I showed steady increase across the selection experiment generations (Table 4). For example, Na increased from 1.613 (G_0) to 1.746 (G_1) to 1.904 (G_2) with gains of 8.25 and 9.05 % between G_0 - G_1 and G_1 - G_2 , respectively (Table 4). Similarly, Ne increase from 1.444 (G_0) to 1.489 (G_1) to 1.605 (G_2) with gains of 3.12 and 7.8 % between G_0 - G_1 and G_1 - G_2 , respectively (Table 3). H increased from 0.251 (G_0) to 0.282 (G_1) to 0.345 (G_2) with gains of 12.35 and 22.34 % between G_0 - G_1 and G_1 - G_2 , respectively (Table 4). Finally, and following the same trend, I increased from 0.366 (G_0) to 0.418 (G_1) to 0.509 (G_2) with gains of 14.21 and 21.77 % between G_0 - G_1 and G_1 - G_2 , respectively (Table 4).

Table 3

Genetic diversity parameters for *C. oleifera* accessions

Locus	N	Na	Ne	H	I
Me3-Em9	31	2	1.763	0.420	0.608
Me9-Em3	31	2	1.966	0.491	0.684
Me3-Em3	31	2	1.935	0.483	0.676
Me4-Em8	31	2	1.730	0.418	0.608
Me9-Em7	31	2	1.920	0.477	0.669
Me6-Em7	31	2	1.874	0.466	0.659
Me5-Em6	31	2	1.743	0.419	0.607
Me8-Em10	31	2	1.801	0.441	0.632
Me3-Em7	31	2	1.563	0.347	0.526
Me9-Em9	31	2	1.403	0.259	0.412
Me5-Em8	31	2	1.465	0.279	0.428
Me8-Em7	31	2	1.632	0.350	0.514
Me9-Em1	31	2	1.635	0.378	0.562
Me9-Em5	31	2	1.797	0.439	0.630
Me3-Em6	31	2	1.751	0.429	0.620
Me3-Em8	31	2	1.674	0.380	0.556
Me8-Em4	31	2	1.762	0.413	0.596
Average	31	2	1.733	0.404	0.585

Note: N, sample size; Na , means observed number of alleles; Ne , means effective number of alleles; H , means Nei's gene diversity; I , means Shannon's Information index.

In order to understand the genetic relationships among the *C. oleifera* 3 generational 31 accessions, we estimated Nei's genetic distances which ranged between 0.145 (cross 30 (56-2x46-2) x 29 (56-2x51-2)) and 0.602 (crosses 22 (46-2x27-5) x 7 (Min49), and 22 (46-2x27-5) x 11 (51-2)) with overall average of 0.378 (Table S2). In depth evaluation of the genetic distance within each of the G_1 and G_2 indicated that either matroclinal or patroclinal inheritance played a role depending on the mating direction (i.e., different genetic distance value from the mating pair depending on which individuals acted as male or female) (Supplementary Table 2).

Table 4
Genetic diversity parameters across *C. oleifera* crossing generations with their respective gain

Generation	Sample size	<i>N_a</i>	<i>N_e</i>	<i>H</i>	<i>I</i>
G ₀	7	1.613	1.444	0.251	0.366
G ₁	5	1.746	1.489	0.282	0.418
G ₂	19	1.904	1.605	0.345	0.509
Across generations	31	2.000	1.732	0.404	0.585
G ₀ -G ₁ generation gain (%)		8.25	3.12	12.35	14.21
G ₁ -G ₂ generation gain (%)		9.05	7.80	22.34	21.77

Note: *N_a*, means observed number of alleles; *N_e*, means effective number of alleles; *H*, means Nei's gene diversity; *I*, means Shannon's Information index. Generation gain = [(Indicator_{curr} - Indicator_{base}) / Indicator_{base}] × 100 %, where Indicator_{curr} and Indicator_{base} represent the mean indicator of the current and base generations, respectively.

Yield and fruit indicators H² and generation gain
Broad-sense heritability (H²) of all yield and fruit indicators across the three generations were generally high and ranged between 0.457 and 0.986 for fruit diameter and seed rate (G₂), and fruit index (G₀), respectively (Table 5). H² estimates varied with generational advance showing both inconsistent increasing and decreasing trends among the studied yield and fruit indicators (Table 5).

Table 5
C. oleifera yield and fruit indicators broad-sense heritability (H²) estimates across generations

Indicator	H ²			
	G ₀	G ₁	G ₂	Average
Yield	0.508	0.672	--	0.590
Fruit diameter	0.571	0.737	0.457	0.588
Fruit height	0.862	0.882	0.859	0.868
Fruit weight	0.700	0.844	0.850	0.798
Fruit index	0.986	0.660	0.603	0.750
Peel thickness	0.929	0.967	0.968	0.955
Number of seeds per fruit	0.844	0.739	0.862	0.815
Seed weight per fruit	0.875	0.868	0.802	0.848
Seed rate	0.900	0.775	0.457	0.711

Heterosis of the yield and fruit indicators in *C. oleifera* hybrids

On general, all heterosis estimates (mid-parents heterosis (MPH), over high-parent heterosis (OPH), and below lower-parent heterosis (BPH)), yield and fruit indicators average values of G₂ were higher than that of G₁ (Supplementary Tables 3 - 5). However, few exceptions were detected for MPH (number of seeds per fruit), OPH (fruit diameter), and BPH (fruit index) (Supplementary Tables 3 - 5). It is noteworthy to indicate some negative heterosis values were observed in both G₁ and G₂ (Supplementary Tables 3 - 5).

Regression analysis of yield and fruit indexes

In order to reveal the intrinsic relationship between *C. oleifera* fruit indicators, a stepwise regression model was developed with yield as the dependent variable and the fruit indicators as the independent variables. This analysis produced optimal regression equation of $Y = 10468.848 + 436.015X_1 + 2345.035X_2$, where *Y*, *X*₁, and *X*₂ represent yield, fruit weight, and number of seeds per fruit, respectively. This relationship was significant (*R* = 0.643, *P* < 0.05) and G₂ yield can be predicted by the resulting regression equation. According to the value of predicted yield, the yield productivity of G₂ were predicted and the top (#29 = 50.2 kg, #30 = 47.4 kg, and #31 = 44.7 kg) and bottom (#25 = 21.9 kg, #21 = 23.1 kg, and #27 = 25.1 kg) yield producer accessions were identified (Table 6).

Table 6
Predicted *C. oleifera* G₂ accession yield with the regression equation

Accession	Fruit weight (g)	Number of seeds per fruit	Predicted yield (g)
13	26.90	4.78	33,406.92
14	27.09	4.47	32,762.80
15	26.44	3.85	31,025.47
16	22.32	4.00	29,579.53
17	28.44	5.78	36,423.42
18	20.23	3.60	27,730.25
19	30.22	7.27	40,693.19
20	28.27	5.00	34,520.17
21	12.89	3.00	23,124.19
22	22.18	4.00	29,519.80
23	20.18	5.00	30,992.81
24	27.45	7.25	39,439.84
25	15.41	2.00	21,877.91
26	15.96	4.00	26,805.61
27	14.64	3.50	25,059.73
28	20.15	4.00	28,636.00
29	41.98	9.15	50,229.83
30	38.57	8.58	47,406.35
31	34.54	8.16	44,664.29
Average	24.94	5.13	33,363.06

Yield and fruit indicators relationship with genetic distance and heterosis

Relationships between heterosis types (MPH, OPH and BPH) and the genetic distance between hybrids' parents (male and female (Supplementary Table 6)) in G₁ hybrids indicated that yield, single fruit weight, and number of seeds per fruit of G₁ generally were negative and in some cases significant (Table 7). For G₂, the relationships were also negative and were significant only for fruit weight (MPH and BPH) (Table 7).

Table 7

C. oleifera yield (G_1) and single fruit weight and number of seed per fruit (G_1 and G_2) relationship between genetic distance and heterosis type

Generation and heterosis index	Fruit weight	Number of seeds per fruit	Yield
G1	-0.768*	-0.737*	-0.502
MPH	-0.560	-0.871*	-0.763*
OPH	-0.738*	-0.947**	-0.795*
BPH	-0.287	-0.517	-0.713
G2	-0.026	-0.127	na
MPh	-0.447*	-0.318	na
OPH	-0.216	-0.291	na
BPH	-0.606**	-0.243	na

* $p < 0.05$; ** $p < 0.01$

Cluster analysis of the predicted *C. oleifera* yield

Cluster analysis of G_2 accessions predicted yields divided the 19 accessions into two clusters. The first has three accessions (29, 30 and 31, all hybrids with 56-2 as the female parent) and all the remaining accessions were classified in a second cluster. The second cluster can be further divided into several sub-clusters, each with a specific female parent (e.g., the first sub-cluster contained four accessions (28, 27, 26, 25, all hybrids with 51-2 as the female parent) (Fig. 1). The most common feature of all sub-cluster is the role of a specific female parent on yield.

Fig. 1

Cluster analysis of *C. oleifera* predicated yield (numbers represent G_2 accessions).

Discussion

SRAP identified *C. oleifera* hybrid authenticity

Traditional hybrid authenticity identification methods using morphological traits often are unreliable as they are affected by environment. Conversely, DNA molecular markers have proven to be effective in hybrid identification with unprecedented accuracy and relative ease (Nandakumar et al. 2004; Hu et al. 2017; Villano et al. 2017). SRAP markers amplify open reading frames and display a co-dominant inheritance mode, which makes it easier for hybrid identification. This hybrid identification method could further be augmented with additional morphological variability as has been previously demonstrated for *C. oleifera* (Fan et al. 2011; Peng S et al. 2011; Peng F 2012; Huang et al. 2014; Abbas et al. 2015; Yu et al. 2016; Zhao Y et al. 2017). In the present study, we used 17 SRAP primes which resulted in stable amplification and produced unambiguous highly polymorphic DNA fragments (average % polymorphic bands (PPB) across all accessions = 91.60 % (Table 2)). When a single SRAP primer was used for *C. oleifera* hybrid identification, we identified progenies with 60.29 % success rate; however, the success rate has increased to 100 % with the use of primer combination and all G_1 (5 individuals) and G_2 (19 individuals) offspring were identified as hybrids (Supplementary Table 1). These results increased our confidence in further using the SRAP markers information in subsequent genetic distance, cluster, and yield and fruits indicators analyses.

Marker-assisted selection of hybrids in *C. oleifera*

Heritability estimates are indicative of the magnitude of genetic control as well as the environment role over a specific attribute (Falconer and Mackay 1996; Kahrizi et al. 2010). Most yield traits (e.g., economic attributes) are complex in nature and follow Fisher's (1918) infinitesimal model in their inheritance mode that is controlled by large number of genes, each with a small effect, and are influenced by environmental conditions. In the present study, broad-sense heritability estimates of *C. oleifera* yield and fruit indicators across the three generations were generally high (range: 0.457 - 0.986) and clearly indicate the presence of high genetic component (Table 5). This observed high heritability estimates indicate that selection for *C. oleifera* yield and fruit indicators is expected to be effective owing to the minor part of environmental influences (Maniee et al. 2009; Kahrizi et al. 2010). Attributes with high broad-sense heritability ($H^2 > 0.800$) across the three generations such as peel thickness, fruit height, number of seeds per fruit, and seed weight per fruit could be used as choice indicator for *C. oleifera* parents breeding and selection, specifically in G_1 .

The use of genetic markers through estimates of genetic distances among accessions were instrumental in the identification of matroclinal and/or patroclinal inheritance which are driven by parental mating directions (Supplementary Table 2). We further offered opportunities for their exploitation during *C. oleifera* breeding through the proper selection of parents and their role in the mating as male or female (see Table S2 for details), a situation could not have reached without genetic markers.

Heterosis is a common biological phenomenon (Lush 1950), thus understanding its modes is helpful for yield prediction and eventually increasing the potential of breeding outcome (Fang et al. 2016; Yadav et al. 2018). Early heterosis prediction can shorten the breeding cycle, improve selection efficiency, and accelerate breeding process (Jan et al. 2016). Our findings indicated that heterosis was associated with genetic distance (Table 7 and Supplementary Tables 3-5), confirming previous observation for many organisms (e.g., sesame, maize, coral reef butterflyfishes) where correlations between the genetic distance of molecular marker and heterosis were reported (Pandey et al. 2018; Seifert et al. 2018; Montanari et al. 2014). *C. oleifera* yield and fruit indicators have shown negative significant relationship with genetic distance and the observed three-heterosis modes (MPH, OPH, and BPH) (Table 7). These results are supported by the presence of the similar relationship in other organisms (Yu et al. 2005; Yao et al. 2016) and provide impetus for its exploitation to further increasing *C. oleifera* yield and fruit indicators in hybrid breeding programs, and highlight the role of DNA markers in unraveling these relationships.

The correct selection of parents is key to the success of *C. oleifera* hybrid breeding. Through the use of molecular marker, in the present study and in particular, genetic distance estimates were instrumental in the identification and selection of parents that maximized yield and fruit indicators and capitalized on heterosis as basis/reference to selection of hybrid parents as clearly illustrated in the present study. Finally, it is

also noteworthy to state that the unsuccessful 56-2x41-4 mating (i.e., lack of progeny production) has been previously reported (Uggla and Carlson-Nilsson 2005; Vander Mijnsbrugge 2015; El'chinova et al. 2014) and maybe hints to the presence of incompatibility mechanisms between these two parents.

Additionally, we assessed the phenotypic correlations among all fruits' attributes to investigate the presence of consistent traits associations across generations (G_0 , G_1 , and G_2) (Supplementary Tables 7-9). While the correlations among fruits attributes varied substantially, we identified 18, 11, and 15 significant out of the calculated 28 correlations across the three generations, respectively. The overall across traits correlations averages were 0.53, 0.52, and 0.53 for G_0 , G_1 , and G_2 , respectively, values supporting consistent associations, hinting to the presence of pleiotropic effect; however, this cannot be verified with the present data. We expected that these consistent significant and positive correlations will ease the selection process. Furthermore, and more interestingly, the presence of negative correlations was rare with 1 for G_1 and G_2 and 3 for G_3 and all were low and nonsignificant.

Marker-assisted early selection of high yield *C. oleifera* hybrids

All genetic diversity parameters (observed number of alleles, means effective number of alleles, Nei's gene diversity, and Shannon's index) over generations showed increasing trend, indicating that directional breeding did not affect the breadth and depth of genetic diversity, an observation reported for several tree species under directional selection (Chaisurisri and El-Kassaby 1994; El-Kassaby and Ritland 1996; Stoehr and El-Kassaby 1997) (Table 4). Additionally, all heterosis modes (MPH, OPH, and BPH) indicated that, on average, G_2 was higher than G_1 in all fruit indicators (Supplementary Tables 3-5). Broad sense heritability (H^2) of yield and fruit weight increased with generations advance, and number of seed per fruit generally increased from G_0 to G_2 (Table 5). We predicted G_2 yield, using the developed yield-fruit indicators stepwise regression model, and results indicated that the average of yield was higher than G_0 and G_1 (Tables 5-6, and Supplementary Tables 3-5). The clustering analysis of G_2 accessions (Fig. 1), indicated that clusters were female parent-specific and in agreement with the yield prediction results of the developed regression equation (see above) and highlighting the validity of the DNA marker-based cluster analysis.

Finally, it should be stated that the use of DNA markers in the present study was effective in hybrid authenticity verification, highlighting the presence of heterosis from the genetic distance analysis and validating the presence of matroclinal and/or patroclinal inheritance and guiding the selection of parents and their respective roles in mating as either males or females. While we considered the present study as a pilot experiment, the promise of MSA should not be overlooked. Naturally, we recommend the development of an overall breeding strategy that accommodate the use of markers. Effective, recurrent selection breeding program is recommended with the inclusion of substantial number of individuals, so the

genetic diversity is maintained over generations to secure the species ability to cope with unforeseen environmental contingencies such as climate change, changing breeding targets objectives, and providing ample opportunities for genetic gain capture.

Conclusions

The utility of marker-assisted selection (MAS) scheme in selecting improved *C. oleifera* varieties yield and fruit attributes using Sequence-Related Amplified Polymorphism markers was implemented through three generation (G_0 , G_1 , and G_2) of mating and selection experiment. The results indicated that the markers used were effective in: 1) clearly identifying hybrid offspring authenticity, 2) demonstrating lack of genetic diversity reduction across the implemented multi-generational scheme, 3) identification of matroclinal or patroclinal inheritance presence which were mainly driven by mating direction, 4) existence of correlated between yield and fruit indicators with genetic distance and heterosis, and 5) clustering of the G_2 accessions was driven by the accessions' maternal parent. The multi-generational mating and selection experiment permitted the development of stepwise regression model for predicting G_2 yield utilizing only two fruit attributes (fruit weight, and number of seeds per fruit). Finally, the utility of MSA in the development of new *C. oleifera* varieties is illustrated.

Acknowledgements

This work was supported by Central Finance Project of China Grants (K4317202A) and Fujian Province Major Science and Technology Project (2013NZ0207). We thank Minhou Tongkou State Forest Farm of Fujian Province for serving as our experimental base.

References

- Abbas G, Hameed A, Rizwan M, Ahsan M, Asghar MJ, Iqbal N (2015) Genetic confirmation of mungbean (*Vigna radiata*) and mashbean (*Vigna mungo*) interspecific recombinants using molecular markers. *Frontiers in Plant Science* 6:1107-1107. <https://doi.org/10.3389/fpls.2015.01107>
- Bankole F, Menkir A, Olaoye G, Crossa J, Hearne S, Unachukwu N, Gedil M (2017) Genetic gains in yield and yield related traits under drought stress and favorable environments in a maize population improved using marker assisted recurrent selection. *Frontiers in Plant Science* 8:808. <https://doi.org/10.3389/fpls.2017.00808>
- Chaisurisri K, El-Kassaby YA (1994) Genetic diversity in a seed production population vs. natural populations of Sitka spruce. *Biodiversity and Conservation*, 3:512-523. <https://doi.org/10.1007/bf00115157>
- Chen Y, Dai X, Hou J, Guan H, Wang Y, Li Y, Yin T (2016) DNA fingerprinting of oil camellia cultivars with SSR markers. *Tree Genetics & Genomes* 12:1. <https://doi.org/10.1007/s11295-015-0966-7>
- Chen Y, Wang D, Liu Y, Peng S, Su L, Chen L (2011) Research on selection of superior male-sterile clone of *Camellia oleifera*. *Journal of Central South University of Forestry & Technology* 31:1-6.
- Cheng J, Jiang D, Cheng H, Zhou X, Fang Y, Zhang X, Xiao X, Deng X, Li L (2018) Determination of *Camellia oleifera* Abel. germplasm resources of genetic diversity in China using ISSR markers. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 46:501-508. <https://doi.org/10.15835/nbha46211092>
- China's State Forestry Administration (2008) *Camellia oleifera*-Part 1: Regulations of selection and breeding for plus tree and superior clone. LY/T 1730.1-2008. Beijing: Standards Press of China 12-01.
- El'chinova GI, Ivanov AV, El'kanova LA, Revazova YA, Zinchenko RA (2014) Acceptability of using Karachay surnames as a quasigenetic marker in population and genetic studies. *Russian Journal of Genetics* 50:769-771. <https://doi.org/10.1134/s1022795414070060>
- El-Kassaby YA (1982) Association between allozyme genotypes and quantitative traits in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco]. *Genetics* 101:103-115.
- El-Kassaby YA, Ritland K (1996) Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodiversity and Conservation* 5:795-813. <https://doi.org/10.1007/bf00051787>
- Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. New York: Longman. [https://doi.org/10.1016/0168-9525\(96\)81458-2](https://doi.org/10.1016/0168-9525(96)81458-2)
- Fan X, Lin P, Zhang S (2011) Analysis of genetic diversity in camellia *oleifera* germplasm family. *Acta Agriculturae Boreali-Sinica* 26:11-17.
- Fang J, Zhu X, Wang C, Shangquan L (2016) Applications of DNA technologies in agriculture. *Current Genomics* 17:379-386. <https://doi.org/10.2174/1389202917666160331203224>
- Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb.* 52:399-433. <https://doi.org/10.1017/s0080456800012163>
- Fehr WR (1987) *Heterosis In: Principles of cultivar development: Theory and Techniques* (Vol.1). Macmillan Publishing Company. New York pp:115.
- Gamal El-Dien O, Ratcliffe B, Klápště J, Chen C, Porth I, El-Kassaby YA (2015) Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics* 16:370. <https://doi.org/10.1186/s12864-015-1597-y>
- Han X, Zhang D, Wang Z, Song Z, Gu Z, Wang J, Kuang X, Tian Y (2012) SRAP-based molecular identification of elite cultivars of camellia *oleifera* from south Jiangxi province. *Journal of Central South University of Forestry & Technology* 32:147-151.
- Hu F, Liu G, Hu Y, Guo R, Zhu L, Luo, Wang F (2017) Authenticity identification and leaf blight resistance evaluation of the F1 hybrids from two *Lilium* cultivars 'Sorbonne' and 'Francia'. *Physiological and Molecular Plant Pathology* 100:194-200. <https://doi.org/10.1016/j.pmpp.2017.10.005>
- Huang CQ, Liu GD, Bai CJ, Wang WQ, Tang WJ (2014) Application of SRAP markers in the identification of *Stylosanthes guianensis* hybrids. *Molecular Biology Reports* 41:5923-5929. <https://doi.org/10.1007/s11033-014-3467-0>
- Huang Y, Chen X, Zhuang X, Lei Z, Chen Y, Peng S (2006) Analysis of genetic diversity in camellia *oleifera* germplasm. *Scientia Sylvae Sinicae* 42:38-43.
- Huang X, Chen J, Yang X, Duan S, Long C, Ge G, Rong J (2018) Low genetic differentiation among altitudes in wild *Camellia oleifera*, a subtropical evergreen hexaploid plant. *Tree Genetics & Genomes* 14:21. <https://doi.org/10.1007/s11295-018-1234-4>
- Igarashi M, Hatsuyama Y, Harada T, Fukasawa-Akada T (2016) Biotechnology and apple breeding in Japan. *Breeding Science* 66:18-33. <https://doi.org/10.1270/jsbbs.66.18>
- Ismaeel M, Shah SMA, Suliman S, Raza A, Anwar M (2018) Assessment of genetic variability, heterosis and heritability for morphological parameters in rice. *Pure and Applied Biology* 700174. <https://doi.org/10.19045/bspab.2018.700174>
- Jan HU, Abbadi A, Lücke S, Nichols RA, Snowdon RJ (2016) Genomic prediction of testcross performance in canola (*Brassica napus*). *PLoS One* 11:e0147769. <https://doi.org/10.1371/journal.pone.0147769>
- Ji Y, Zhang X, Ma X, Chu X, Li F, Meng Y (2009) Genetic analysis of *Lolium* multi-florum hybrids and their progenies detected using SRAP markers. *Acta Prataculturae Sinica* 18:260-265.
- Kadam DC, Potts SM, Bohn MO, Lipka AE, Lorenz AJ (2016) Genomic prediction of single crosses in the early stages of a maize hybrid breeding pipeline. *G3: Genes, Genomes, Genetics* 6:3443-3453. <https://doi.org/10.1534/g3.116.031286>

- Kahrizi D, Cheghamirza K, Kakaei M, Mohammadi R, Ebadi A (2010) Heritability and genetic gain of some morphophysiological variables of durum wheat (*Triticum turgidum* var. durum). *African Journal of Biotechnology* 9:4687-4691.
- Li C, Xu W, Guo R, Zhang J, Qi X, Hu L, Zhao M (2018) Molecular marker assisted breeding and genome composition analysis of Zhengmai 7698, an elite winter wheat cultivar. *Scientific Reports* 8:322. <https://doi.org/10.1038/s41598-017-18726-8>
- Li G, Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. *Theoretical and Applied Genetics* 103:455-461. <https://doi.org/10.1007/s001220100570>
- Liao T, Yuan DY, Zou F, Gao C, Yang Y, Zhang L, Tan XF (2014) Self-sterility in *Camellia oleifera* may be due to the prezygotic late-acting self-incompatibility. *PLoS One* 9:e99639. <https://doi.org/10.1371/journal.pone.0099639>
- Liang Y, Yang Z, Huang H (2010) Status and development trend of *Camellia oleifera* breeding in China. *Jiangxi Forestry Science and Technology* 2:24-28.
- Lush JL (1950) The theory of inbreeding. *American journal of human genetics* 2:97-100.
- Ma J, Ye H, Rui Y, Chen G, Zhang N (2011) Fatty acid composition of *Camellia oleifera* oil. *Journal für Verbraucherschutz und Lebensmittelsicherheit* 6:9-12. <https://doi.org/10.1007/s00003-010-0581-3>
- Maniee M, Kahrizi D, Mohammadi R (2009) Genetic variability of some morpho-physiological traits in durum wheat (*Triticum turgidum* var. Durum). *Journal of Applied Sciences* 9:1383-1387. <https://doi.org/10.3923/jas.2009.1383.1387>
- Montanari SR, Hobbs JPA, Pratchett MS, Bay LK, Herwerden LV (2014) Does genetic distance between parental species influence outcomes of hybridization among coral reef butterflyfishes? *Molecular Ecology* 23:2757-2770. <https://doi.org/10.1111/mec.12762>
- Nandakumar N, Singh AK, Sharma RK, Mohapatra T, Prabhu KV, Zaman FU (2004) Molecular fingerprinting of hybrids and assessment of genetic purity of hybrid seeds in rice using microsatellite markers. *Euphytica* 136:257-264. <https://doi.org/10.1023/b:euph.0000032706.92360.c6>
- NTSYSpc (2018) Numerical Taxonomy System. Version 2.2. <http://www.exetersoftware.com/cat/ntsyspc/ntsyspc.html>.
- Pandey SK, Dasgupta T, Rathore A, Vemula A (2018) Relationship of parental genetic distance with heterosis and specific combining ability in sesame (*Sesamum indicum* L.) based on phenotypic and molecular marker analysis. *Biochemical Genetics* 56:188-209. <https://doi.org/10.1007/s10528-017-9837-2>
- Parikh L, Mmbaga MT, Kodati S, Blair M, Hui D, Meru G (2016) Broad-sense heritability and genetic gain for powdery mildew resistance in multiple pseudo-F2 populations of flowering dogwoods (*Cornus florida* L.). *Scientia Horticulturae* 213:216-221. <https://doi.org/10.1016/j.scienta.2016.09.038>
- Peng F, Wu Y, Hao M, Chen L, Chen Y (2012) Genetic diversity of *camellia oleifera* using ISSR and SRAP markers. *Journal of Nanjing Forestry University (Natural Science Edition)* 36:19-25.
- Peng S, Chen Y, Wang X, Wang R, Ma L, Chen L (2010) Research progress and development strategies of breeding of oil-tea *Camellia*. *Hunan Forestry Science & Technology* 37:32-38.
- Peng S, Zhang D, Chen Y, Song Z, Nan X, Zhang H (2011) Genetic diversity analysis of elite cultivars of *camellia oleifera* by SRAP. *Journal of Central South University of Forestry & Technology* 31:80-85.
- Qin S, Rong J, Zhang W, Chen J (2018) Cultivation history of *Camellia oleifera* and genetic resources in the Yangtze River Basin. *Biodiversity Science* 26:384-395. <https://doi.org/10.17520/biods.2017254>
- Ratcliffe B, Gamal El-Dien O, Klápště J, Porth I, Chen C, Jaquish B, El-Kassaby YA (2015) A comparison of genomic selection models across time in Interior Spruce (*Picea engelmannii* × *glauca*) using high density SNP imputation methods. *Heredity* 115:547-555. <https://doi.org/10.1038/hdy.2015.57>
- Ru S, Main D, Evans K, Peace C (2015) Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. *Tree Genetics & Genomes* 11:8. <https://doi.org/10.1007/s11295-015-0834-5>
- Salgotra RK, Gupta BB, Bhat JA, Sharma S (2015) Genetic diversity and population structure of Basmati rice (*Oryza sativa* L.) germplasm collected from north western Himalayas using trait linked SSR markers. *PLoS One* 10:e0131858. <https://doi.org/10.1371/journal.pone.0131858>
- Seifert F, Thiemann A, Grant-Downton R, Edelmann S, Rybka D, Schrag TA, Frisch M, Dickinson HG, Melchinger AE, Scholten S (2018) Parental expression variation of small RNAs is negatively correlated with grain yield heterosis in a maize breeding population. *Frontiers in Plant Science* 9:13. <https://doi.org/10.3389/fpls.2018.00013>
- Sharma HP, Dhakal KH, Kharel R, Shrestha J (2016) Estimation of heterosis in yield and yield attributing traits in single cross hybrids of maize. *Journal of Maize Research and Development* 2:123-132. <https://doi.org/10.3126/jmrd.v2i1.16223>
- Shi S, Nan L, Smith KF (2017) The current status, problems, and prospects of alfalfa (*Medicago sativa* L.) breeding in China. *Agronomy* 7:1. <https://doi.org/10.3390/agronomy7010001>
- Simko I, Hayes RJ, Truco MJ, Michelmore RW, Antonise R, Massoudi M (2018) Molecular markers reliably predict post-harvest deterioration of fresh-cut lettuce in modified atmosphere packaging. *Horticulture Research* 5:21. <https://doi.org/10.1038/s41438-018-0022-5>
- Stoehr MU, El-Kassaby YA (1997) Levels of genetic diversity at different stages of the domestication cycle on interior spruce in British Columbia. *Theor. Appl. Genet.* 84:83-90. <https://doi.org/10.1007/s001220050385>
- Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, El-Kassaby YA (2017) Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* 18:930. <https://doi.org/10.1186/s12864-017-4258-5>
- Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, El-Kassaby YA (2019) Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Heredity* 1. <https://doi.org/10.1038/s41437-018-0172-0>
- Uggle M, Carlson-Nilsson BU (2005) Screening of fungal diseases in offspring from crosses between *Rosa* sections *Caninae* and *Cinnamomeae*. *Scientia Horticulturae* 104:493-504. <https://doi.org/10.1016/j.scienta.2004.11.001>
- Vander Mijnsbrugge K (2015) Morphological dissection of leaf, bud and inflorescence traits of the interfertile native *A. glutinosa* and non-native *A. incana* in Flanders (northern part of Belgium). *Trees* 29:1661-1672. <https://doi.org/10.1007/s00468-015-1247-7>
- Villano C, Lisanti MT, Gambuti A, Vecchio R, Moio L, Frusciantè L, Aversano R, Carputo D (2017) Wine varietal authentication based on phenolics, volatiles and DNA markers: State of the art, perspectives and drawbacks. *Food Control* 80:1-10. <https://doi.org/10.1016/j.foodcont.2017.04.020>
- Wang P, Cai L, Wei C, Huang J, Jiang X, Li M, Qin Z, Chen Y, Lu Y, Wu Y (2014) Genetic diversity analysis of *camellia oleifera* in Guangxi province using SRAP molecular markers. *Molecular Plant Breeding* 12:681-686.
- Wu W, Hu JJ, Duan WJ, Shen YF, Cao ZH, Shu QL (2015) Study on the chemical components of shoots and leaves of *Camellia oleifera* from different elevations. *Pakistan Journal of Botany* 47:1111-1114.
- Xue D, Guo H, Zheng Y, Chen X, Liu J (2009) Hybrid identification of progenies of *Zoysia* crosses by SRAP marker. *Acta Prataculturae Sinica* 18:72-79.
- Yadav V, Singh NP, Sharma R, Gupta A, Baranwal A, Ahmad SF, Raina V (2018) Crossbreeding systems of livestock. *The Pharma Innovation Journal* 7:8-13.
- Yang C, Liu X, Chen Z, Lin Y, Wang S (2016) Comparison of oil content and fatty acid profile of ten new *Camellia oleifera* cultivars. *Journal of Lipids* 3982486. <https://doi.org/10.1155/2016/3982486>
- Yang H, Wang X, Wei Y, Deng Z, Liu H, Chen J, Dai L, Xia Z, He G, Li D (2018) Transcriptomic analyses reveal molecular mechanisms underlying growth heterosis and weakness of rubber tree seedlings. *BMC Plant Biology* 18:10. <https://doi.org/10.1186/s12870-017-1203-3>
- Yang J, Luo S, Li J, Zheng Z, Du X, Deng Y (2018) Transcriptome analysis of growth heterosis in pearl oyster *Pinctada fucata martensii*. *FEBS Open Bio* 8:1794-1803. <https://doi.org/10.1002/2211-5463.12502>
- Yao J, Li H, Ye J, Shi L (2016) Relationship between parental genetic distance and offspring's heterosis for early growth traits in *Liriodendron*: implication for parent pair selection in cross breeding. *New Forests* 47: 63-177. <https://doi.org/10.1007/s11056-015-9508-2>
- Yu C, Yin Y, Creech DL, Lu Z, Xu J (2016) Morphological characters and SRAP analysis of two hybrids between *Hibiscus dasycalyx* and *Hibiscus* 'Moy Grande'. *Scientia Horticulturae* 198:118-124. <https://doi.org/10.1016/j.scienta.2015.11.037>
- Yu CY, Hu SW, Zhao HX, Guo AG, Sun GL (2005) Genetic distances revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.).

Theoretical and Applied Genetics 110:511-518.

<https://doi.org/10.1007/s00122-004-1858-7>

Zeb A (2012) Triacylglycerols composition, oxidation and oxidation compounds in camellia oil using liquid chromatography-mass spectrometry. *Chemistry and Physics of Lipids* 165:608-614.

<https://doi.org/10.1016/j.chemphyslip.2012.03.004>

Zhang R, Ding Z, Zhang X, Jin X, Wen L (2008) Advances in genetic improvement of tea oil camellia. *Acta Horticulturae* 769:33-42.

<https://doi.org/10.17660/actahortic.2008.769.2>

Zhao X, Zhang J, Zhang Z, Wang Y, Xie W (2017) Hybrid identification and genetic variation of *Elymus sibiricus* hybrid populations using EST-SSR markers. *Hereditas* 154:15.

<https://doi.org/10.1186/s41065-017-0053-1>

Zhao Y, Ruan CJ, Ding GJ, Mopper S (2017) Genetic relationships in a germplasm collection of *Camellia japonica* and *Camellia oleifera* using SSR analysis. *Genetics & Molecular Research* 16:16019526.

<https://doi.org/10.4238/gmr16019526>

Zheng Y, Xu S, Liu J, Zhao Y, Liu J (2017) Genetic diversity and population structure of Chinese natural bermudagrass [*Cynodon dactylon* (L.) Pers.] germplasm based on SRAP markers. *PloS One* 12: e0177508.

<https://doi.org/10.1371/journal.pone.0177508>