

RESEARCH ARTICLE

Simple Genetic Distance-Optimized Field Deployments for Clonal Seed Orchards Based on Microsatellite Markers: As a Case of Chinese Pine Seed Orchard

Huwei Yuan¹, Shihui Niu¹, Yousry A. El-Kassaby², Yue Li^{1*}, Wei Li^{1*}

1 National Engineering Laboratory for Forest Tree Breeding, Key Laboratory for Genetics and Breeding of Forest Trees and Ornamental Plants of Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China, **2** Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, Canada

* bjfuliwei@bjfu.edu.cn (WL); liyue@bjfu.edu.cn (YL)



OPEN ACCESS

Citation: Yuan H, Niu S, El-Kassaby YA, Li Y, Li W (2016) Simple Genetic Distance-Optimized Field Deployments for Clonal Seed Orchards Based on Microsatellite Markers: As a Case of Chinese Pine Seed Orchard. PLoS ONE 11(6): e0157646. doi:10.1371/journal.pone.0157646

Editor: Yuepeng Han, Wuhan Botanical Garden of Chinese Academy of Sciences, CHINA

Received: February 19, 2016

Accepted: June 2, 2016

Published: June 16, 2016

Copyright: © 2016 Yuan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This work was supported by “the Fundamental Research Funds for the Central Universities (No. 2015ZCQ-SW-02)” and “Special Fund for Forestry Scientific Research in the Public Interest (No.201104022)”. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Chinese pine seed orchards are in a period of transition from first-generation to advanced-generations. How to effectively select populations for second-generation seed orchards and significantly increase genetic gain through rational deployment have become major issues. In this study, we examined open- and control-pollinated progeny of the first-generation Chinese pine seed orchards in Zhengning (Gansu Province, China) and Xixian (Shanxi Province, China) to address issues related to phenotypic selection for high volume growth, genetic diversity analysis and genetic distance-based phylogenetic analysis of the selections by simple sequence repeats (SSRs), and phylogenetic relationship-based field deployment for advanced-generation orchards. In total, 40, 28, 20, and 13 superior individuals were selected from the large-scale no-pedigree open-pollinated progeny of Zhengning (ZN-NP), open-pollinated families of Zhengning (ZN-OP), open-pollinated families of Xixian (XX-OP), and control-pollinated families of Xixian, with mean volume dominance ratios of 0.83, 0.15, 0.25, and 0.20, respectively. Phylogenetic relationship analysis of the ZN-NP and XX-OP populations showed that the 40 superior individuals in the ZN-NP selected population belonged to 23 families and could be further divided into five phylogenetic groups, and that families in the same group were closely related. Similarly, 20 families in the XX-OP population were related to varying degrees. Based on these results, we found that second-generation Chinese pine seed orchards in Zhengning and Xixian should adopt a grouped, unbalanced, complete, fixed block design and an unbalanced, incomplete, fixed block design, respectively. This study will provide practical references for applying molecular markers to establishing advanced-generation seed orchards.

Introduction

Selection and deployment of improved materials are important in establishing an advanced-generation seed orchard. To increase genetic gain, advanced-generation seed orchards usually contain only a few elite clones. To decrease the risks of inbreeding depression [1, 2], propagation populations in advanced-generation seed orchards are often composed of elite individuals from elite, pedigreed families selected through controlled pollination. However, controlled pollination is a time, resource- and cost-intensive activity, and the number of tested families is always limited, which greatly restricts the genetic quality of the selected germplasm. Open-pollinated progenies contain all possible crosses and greater numbers of superior individuals; however, such individuals are rarely used when establishing an advanced-generation seed orchard because they are not pedigreed.

Molecular markers developed in the 1980s have brought phylogenetic analysis to a new level [3]. Microsatellite markers and relevant analytical softwares have made marker-based phylogenetic analysis a reality [4, 5], and are widely used in phylogenetic analysis of natural and hybrid populations [5–8]. El-Kassaby introduced a breeding strategy called “breeding without breeding” (BWB), which has been proven highly convenient for tree breeding [9]. The efficiency of this strategy has been evaluated by progeny testing, parental selection, and construction of pedigrees [9–11]. Later, the BWB strategy was demonstrated in a number of tree species [12–14] and it has been extensively used in different areas of forest tree breeding, including phylogenetic analysis [15, 16], mating systems [12, 17–19], estimation of genetic parameters and breeding value [13, 20], and spatial variation [17]. Selecting superior individuals directly from the open-pollinated progeny of a seed orchard or from plantations established using seed orchard seeds, coupled with identifying the phylogenetic relationship of the selected materials based on molecular markers, could decrease the reliance on controlled pollination and shorten the breeding cycle by 10–15 years [9]. Genetic distance reflects the genetic relationships among materials. Simple sequence repeats (SSRs) are high-resolution markers that can identify different individuals within the same species. The combination of phenotypic selection, genetic distance-based phylogenetic analysis of selected individuals using SSR markers and phylogenetic relationship-based field deployment would simplify breeding activities, decrease inbreeding, and expand genetic diversity among seed orchard progenies.

Field deployment is one of the most important activities in seed orchard establishment. The most important criteria of field deployment is to maximize genetic gain of target traits while maintaining an acceptable level of genetic diversity [21]. Advanced-generation seed orchards of conifers often contain a small number of clones with varying origin (including backward and forward selections) [22], which has increased the probability of inbreeding and enhanced the complexity of deployment. A number of designs, including permuted neighborhood design [23], systematic layout [22], randomized, replicated, staggered clonal-row (R^2 SCR) design [24], have been employed in the deployment of advanced-generation clonal seed orchards. It was proved that unequal clonal deployment could improve genetic gain at a certain level of genetic diversity [21, 25, 26], and the mating system of advanced-generation clonal seed orchards could be controlled by allocating clones based on their phylogenetic relationships [27].

Chinese pine (*Pinus tabulaeformis*) is a major afforestation tree species that is naturally distributed in the 3,000,000 km² of mountainous areas between 32°–43°N and 102°–122°E in northern China [28–30]. Genetic improvement of Chinese pine was initiated in the 1970s [31]. Sixteen first-generation seed orchards containing more than 4,000 plus-tree clones were established, and about 1,000 open-pollinated and 200 control-pollinated families were tested [32]. Variations in flowering and fruiting characteristics among clones in the seed orchards were

studied extensively [33–38] and improved seeds from the orchards were used for afforestation. Chinese pine seed orchards are currently in a period of transition from first-generation to advanced-generation orchards; thus, how to select superior individuals from the available material and form rational deployment designs of clones in advanced-generation seed orchards have become major concerns for breeders.

In this study, we examined large-scale no-pedigree open-pollinated progeny, and open- and control-pollinated families of the first-generation Chinese pine seed orchards in Zhengning and Xixian as a base population. Our aims were to: screen polymorphic SSR primers adapted to Chinese pine, pre-select superior individuals in the base population based on volume growth, identify the relative phylogenetic relationships among the selected individuals, and develop a simplified clone deployment strategy for the corresponding advanced-generation seed orchards. Our results will provide practical references for applying molecular markers in Chinese pine improvement programs.

Materials and Methods

Screening of polymorphic SSR primers for Chinese pine

We extracted genomic DNA from the needles of Chinese pine using the CTAB method [39]. Ours is the leading cooperative group researching Chinese pine, and comprises the Zhengning Forestry Institute in Gansu Province and the Lüliang Mountain State-owned Forest Administration in Shanxi Province. Both authorities gave full permission to conduct all studies at the experimental sites. This study did not involve endangered or protected species. Twenty-one SSR primers from Chinese pine and related species [40] were synthesized by Sangon Biotech Co. Ltd. (Beijing, China) and used for polymerase chain reaction (PCR) amplification. PCR amplification was performed on a Veriti Dx 96-well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) using a method described previously [41]. The amplified PCR fragments were separated on 8% polyacrylamide gels using an HT-SCZ04 Vertical Electrophoresis Tank (Hongtaojiye Science & Technology Co. Ltd., Beijing, China) according to the manufacturer's instructions. Primers that could present polymorphic amplification products among individuals were screened and used for the following experiment. Detailed information regarding the 11 polymorphic SSR primers is shown in Table 1.

Base population and selection of superior individuals

The first-generation Chinese pine seed orchards in Zhengning (Gansu Province) and Xixian (Shanxi Province) contained about 300 clones each. We regarded no-pedigree open-pollinated progeny of clones in the Zhengning seed orchard (ZN-NP), open-pollinated families of clones in the Zhengning seed orchard (ZN-OP), open-pollinated families of clones in the Xixian seed orchard (XX-OP), and control-pollinated families of clones in the Xixian seed orchard (XX-CP) as the base population, with ages of 21, 21, 22, and 22 years, respectively. We measured individual's heights and diameters at 1.3 m, and calculated individual's volume as:

$$V = 2\pi \cdot (D_{1.3}/2) \times (H + 3) \times f$$

where V is the individual's volume, f is 0.3578 for Chinese pine, $D_{1.3}$ is diameter at 1.3 m height, and H is height. Individuals or families in the base population were ranked according to their volume dominance ratio (R_d), which was calculated as:

$$R_d = (V - \bar{V})/\bar{V}$$

where V is an individual tree's volume or a family's mean volume; and \bar{V} is the mean volume of

Table 1. Characteristics of the *P. tabuliformis* polymorphic simple sequence repeat (SSR) primers.

Primer No.	Locus	Primer sequences (5'-3')	Expected size (bp)	T _a (°C)
1	LOP1	F: GGCTAATGGCCGGCCAGTGCT	158	58
1	LOP1	R: GCGATTACAGGTTGCAGCCT	158	58
2	PiTX2146	F: CCTGGGGATTGGATTGGGTATTTG	200	51
2	PiTX2146	R: ATATTTTCCTTGCCCTTCCAGACA	200	51
3	PiTX3107	F: AAACAAGCCACATCGTCAATC	160	54
3	PiTX3107	R: TCCCCTGGATCTGAGGA	160	54
4	PiTX3116	F: CCTCCCAAAGCCTAAAGAAT	160	54
4	PiTX3116	R: CATAACAAGGCCTTATCTTACAGAA	160	54
5	PiTX4001	F: CTATTTGAGTTAAGAAGGGAGTC	210	52
5	PiTX4001	R: CTGTGGGTAGCATCATC	210	52
6	SPAC12.5	F: CTTCTTCACTAGTTTCCTTTGG	130	56
6	SPAC12.5	R: TTGGTTATAGGCATAGATTGC	130	56
7	ctg1367	F: CGATATTATGGATTTTGCTTGTGA	120	53
7	ctg1367	R: AAATGCATGCCAAACTTAAATAC	120	53
8	pdms221	F: GAGAGTTGTATGACGGAAATAC	165	54
8	pdms221	R: CCCACACAAAAGTGTACTTC	165	54
9	CR354677	F: ATGGTGGTTGATCTGCAGGCTAAT	230	51
9	CR354677	R: GTTTAACCAACCGCACTCATTTTTCACT	230	51
10	PiTX2123	F: GAAGAACCACAAACACAAG	210	57
10	PiTX2123	R: GGGCAAGAATCAATGATAA	210	57
11	RPS160	F: ACTAAG AACTCTCCCTCTCACC	240	55
11	RPS160	R: TCATTGTTCCCAATCAT	240	55

Note: T_a, annealing temperature.

doi:10.1371/journal.pone.0157646.t001

the corresponding population. The selected population in ZN-NP was composed of individuals with high-ranked R_d values in the base population; the ZN-OP, XX-OP and XX-CP selected populations were composed of high-ranked individuals (individuals with high R_d values) within the high-ranked families (families with high R_d values) in the corresponding base populations. We conducted analyses of variance (ANOVAs) and multiple comparisons using R 2.15.3 software to describe differences in R_d values among the selected populations.

Genetic distance-based phylogenetic analysis and genetic diversity analysis of superior individuals in the ZN-NP and XX-OP selected populations

We used superior individuals in the ZN-NP and XX-OP selected populations as samples for the phylogenetic analysis. We performed DNA extraction, PCR amplification, and polyacrylamide gel electrophoresis using the methods described above, and analyzed the genotyping results using PowerMarker 3.25 software. We also calculated frequency and frequency-based distance (Nei 1983), and reconstructed genetic distance-based phylogenetic trees using the unweighted pair-group method with the arithmetic average method. Phylogenetic trees were viewed using Fig.Tree 1.4.2 software.

We analyzed the genetic diversity of superior individuals in the ZN-NP and XX-OP selection populations using PowerMarker 3.25 software. The parameters analyzed included observed heterozygosity (H_o), expected heterozygosity (H_e), and polymorphism information content (PIC).

Results

Selected populations in the second-generation Chinese pine seed orchards

The second-generation selected populations consisted of superior individuals with high-ranking R_d values. The ZN-NP, ZN-OP, XX-OP, and XX-CP selected populations had R_d values of 0.20–2.15, 0.09–0.40, 0.11–0.60, and 0.09–0.29, with mean values of 0.83, 0.15, 0.25, and 0.20, respectively (Fig 1a–1d). ANOVA showed that R_d values were significantly higher in the ZN-NP selected population than in the other selected populations, whereas there were no statistically significant differences in R_d values among the ZN-OP, XX-OP, and XX-CP selected populations (Fig 1e).

Genetic distance-based phylogenetic analysis of superior individuals in the ZN-NP and XX-OP selected populations

The fingerprints revealed that the amplified bands were polymorphic among individuals in the ZN-NP and XX-OP selected populations (Fig 2). The phylogenetic analysis revealed that the 40 individuals in the ZN-NP selected population belonged to 23 families with various numbers of individuals. The largest family consisted of eight individuals (nos. 8, 9, 10, 11, 12, 14, 19, and 20). The numbers of families containing one, two, three, and four individuals were 15, 5, 1, and 1, respectively (Fig 3a). We found a close relationship between the family consisting of individuals 28 and 29 and the family consisting of individuals 15, 16, and 18, whereas the relationships between these two families and the other families were distant. Similarly, families consisting of individuals 13, 17, 22, 27, and 38 were closely related but they were distantly related to the other families. The details of the phylogenetic relationships among the families in the ZN-NP selected population are shown in Fig 3a.

The 20 individuals in the XX-OP selected population belonged to 20 different families (Fig 3b). The families were numbered the same as the individuals. Family 3 was distantly related to the other families. Families 5, 12, and 14 were closely related but were distantly related to the other families. Families 2, 7, and 20 were closely related but were distantly related to the other families. The details of the phylogenetic relationships among families in the XX-OP selected population are shown in Fig 3b.

Genetic diversity of superior individuals in the ZN-NP and XX-OP selected populations

Table 2 shows that the major allele frequency (*MAF*), the number of genotypes (No. genotypes), and the number of alleles (No. alleles) of the 11 polymorphic SSR loci in the ZN-NP selected population had ranges of 0.3250–0.6875, 4.0000–10.0000, and 3.0000–11.0000, with means of 0.4568, 6.0909, and 5.5455, respectively. Expected heterozygosity (H_e) and observed heterozygosity (H_o) values varied from 0.4609 to 0.7931 and from 0.1000 to 1.000, respectively, with means of 0.6690 and 0.5614. Polymorphism information content (*PIC*) changed from 0.3977 to 0.7671, with a mean of 0.6202. The fixation index (*F*) had a range of –0.2676–0.8532, with a mean of 0.1732.

Table 3 shows that the *MAF*, numbers of genotypes, and numbers of alleles of the 11 polymorphic SSR loci in the selected XX-OP population had ranges of 0.2750–0.9500, 2.0000–13.0000, and 2.0000–10.0000, with means of 0.5846, 7.1818, and 6.0000, respectively. H_e and H_o varied from 0.0950 to 0.8438 and from 0.1000 to 0.8500, respectively, with means of 0.5703 and 0.5955. The *PIC* changed from 0.0905 to 0.8266, with a mean of 0.5324. *F* ranged between –0.2925 and 0.3956, with a mean of –0.0184.

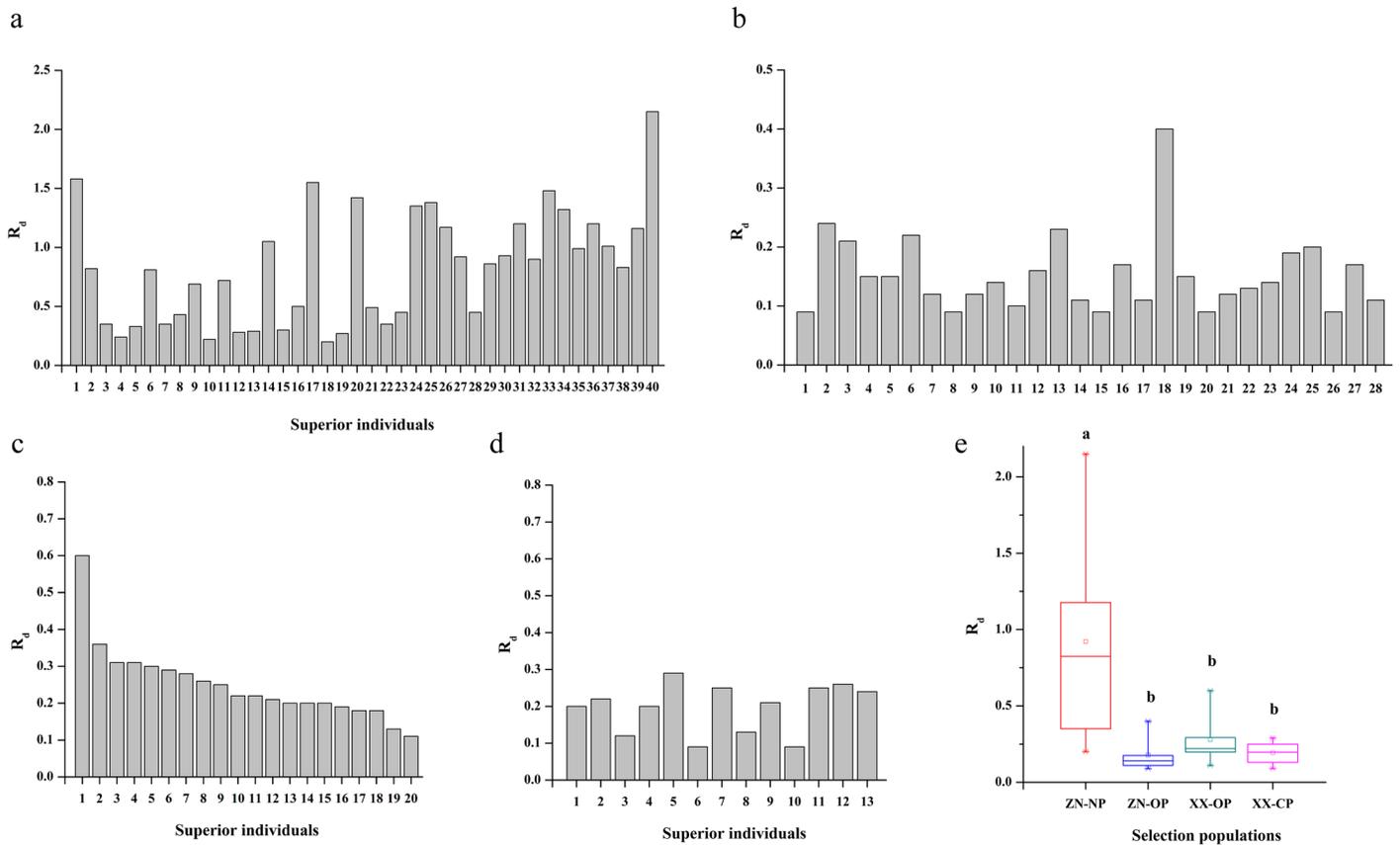


Fig 1. Volume dominance ratios (R_d) of the second-generation selected populations of Chinese pine. Comments: (a) Volume dominance ratios of superior individuals in the ZN-NP selected population; (b) volume dominance ratios of superior individuals in the ZN-OP selected population; (c) volume dominance ratios of superior individuals in the XX-OP selected population; (d) volume dominance ratios of superior individuals in the XX-CP selected population; (e) Variations in volume dominance ratios among the four selected populations.

doi:10.1371/journal.pone.0157646.g001

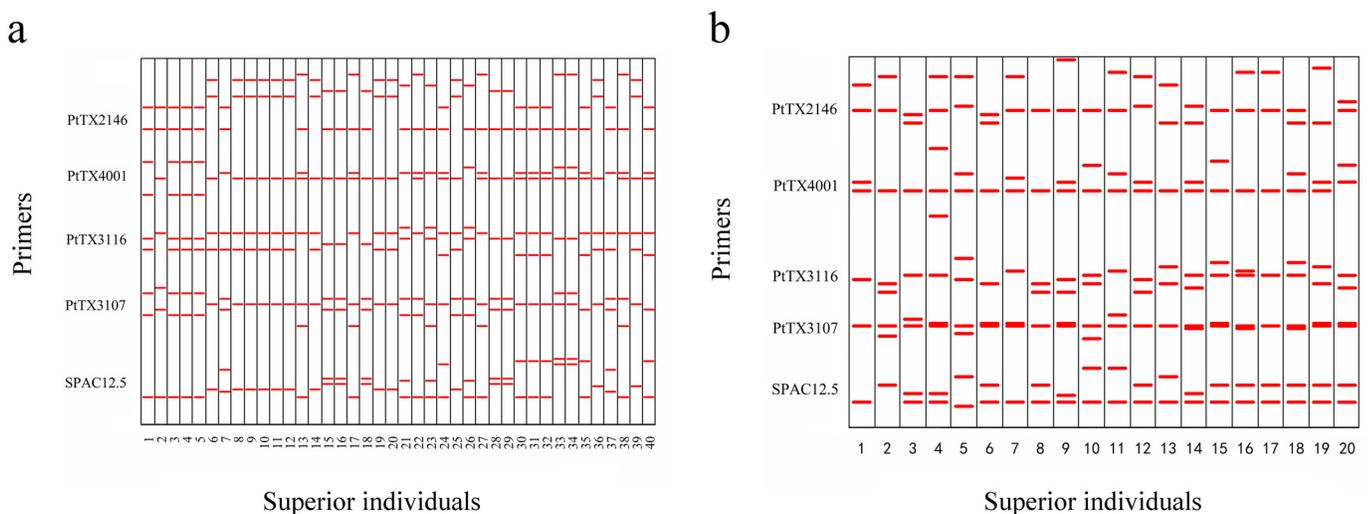


Fig 2. Fingerprints of five simple sequence repeat (SSR) primers in the second-generation ZN-NP (a) and XX-OP (b) selected populations of Chinese pine.

doi:10.1371/journal.pone.0157646.g002

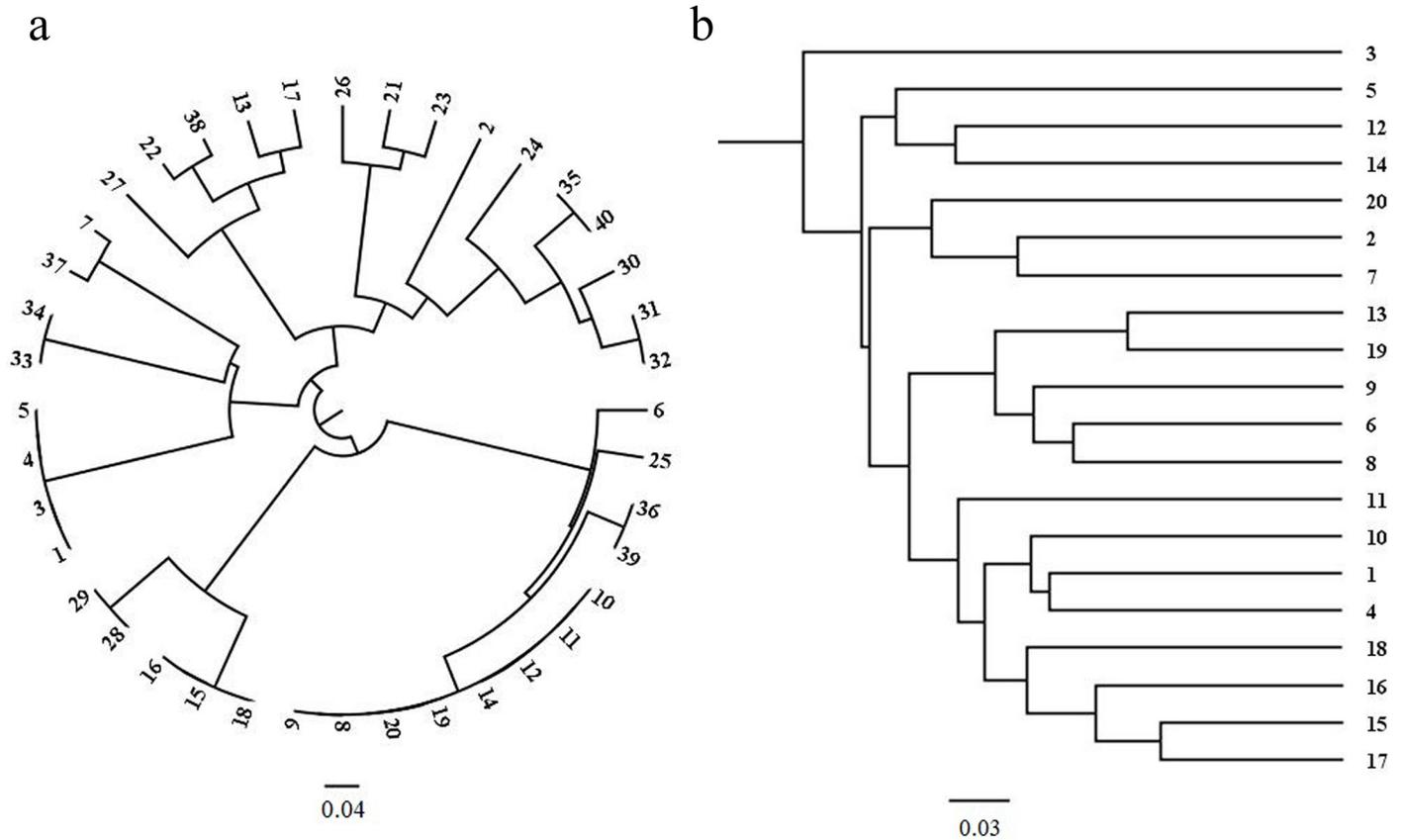


Fig 3. Phylogenetic trees of superior individuals in the second-generation ZN-NP (a) and XX-OP (b) selected populations of Chinese pine.

doi:10.1371/journal.pone.0157646.g003

Table 2. Genetic diversity in the selected second-generation Chinese pine ZN-NP population.

Primer No.	MAF	No. genotype	No. allele	H_e	H_o	PIC	F
1	0.3250	10.0000	7.0000	0.7663	0.9000	0.7314	-0.1623
2	0.3500	5.0000	7.0000	0.7931	1.0000	0.7671	-0.2490
3	0.4500	9.0000	7.0000	0.7328	0.5750	0.7036	0.2274
4	0.4000	7.0000	6.0000	0.7416	0.6750	0.7042	0.1023
5	0.6125	6.0000	5.0000	0.5609	0.4250	0.5121	0.2542
6	0.3625	9.0000	11.0000	0.7872	0.4500	0.7631	0.4386
7	0.5750	4.0000	5.0000	0.5863	0.7500	0.5301	-0.2676
8	0.6875	4.0000	3.0000	0.4609	0.3250	0.3977	0.3064
9	0.3625	5.0000	3.0000	0.6653	0.1000	0.5913	0.8532
10	0.4000	4.0000	4.0000	0.6550	0.3000	0.5877	0.5509
11	0.5000	4.0000	3.0000	0.6097	0.6750	0.5335	-0.0946
Mean	0.4568	6.0909	5.5455	0.6690	0.5614	0.6202	0.1732

Comments: MAF, H_e , H_o , PIC, and F are major allele frequency, expected heterozygosity, observed heterozygosity, polymorphism information content, and the fixation index, respectively.

doi:10.1371/journal.pone.0157646.t002

Table 3. Genetic diversity in the second-generation XX-OP selected population of Chinese pine.

Primer No.	MAF	No. genotype	No. allele	H_e	H_o	PIC	F
1	0.6000	10.0000	7.0000	0.6013	0.6000	0.5725	0.0277
2	0.4000	12.0000	10.0000	0.7838	0.8500	0.7634	-0.0590
3	0.6250	8.0000	8.0000	0.5700	0.7500	0.5387	-0.2925
4	0.2750	13.0000	10.0000	0.8438	0.7000	0.8266	0.1952
5	0.6500	8.0000	7.0000	0.5450	0.6500	0.5170	-0.1678
6	0.5000	8.0000	7.0000	0.6625	0.8500	0.6169	-0.2593
7	0.9500	2.0000	2.0000	0.0950	0.1000	0.0905	-0.0270
8	0.7000	3.0000	3.0000	0.4338	0.5500	0.3589	-0.2440
9	0.5500	8.0000	6.0000	0.6388	0.4000	0.6015	0.3956
10	0.5500	3.0000	3.0000	0.5950	0.5000	0.5280	0.1845
11	0.6500	4.0000	3.0000	0.5050	0.6000	0.4425	-0.1633
Mean	0.5864	7.1818	6.0000	0.5703	0.5955	0.5324	-0.0184

Comments: MAF, H_e , H_o , PIC and F are major allele frequency, expected heterozygosity, observed heterozygosity, polymorphism information content and the fixation index, respectively.

doi:10.1371/journal.pone.0157646.t003

Deployment designs for the second-generation Chinese pine clonal seed orchards in Zhengning and Xixian

We conducted secondary selection among the elementarily selected individuals in the ZN-NP and XX-OP selected populations before preparing the deployment designs. Families with ≥ 2 individuals underwent within-family selection and only individuals with the highest R_d values were selected in the corresponding families. Designs were prepared based on the criteria of maintaining maximum genetic distance and genetic diversity, increasing the frequency of the best selections, and fixed deployment.

Seventeen individuals were removed by secondary selection, and 23 families with only one individual each were included in the final ZN-NP selected population. The 23 families were divided into six phylogenetic groups (G_p) according to their phylogenetic relationships (Fig 3a). Families within the same group were closely related, and families in different groups were distantly related. For example, families 6, 25, 36, and 20 were included in the same G_p , and families 15 and 29 were in the same G_p (Fig 3a). The details of the grouping referred to Fig 3a. From two to five families were in different G_p . We adopted a grouped, unbalanced, complete, fixed block (GUCFB) design for the second-generation Chinese pine seed orchards in Zhengning (Fig 4). We chose one family from each of the six G_p , and the six selected families with a distant genetic relationship formed a deployment group (G_d ; families in the same rectangle of Fig 4 were in the same G_d). Five G_d were needed when all 23 families were included. The families with the highest R_d values in each G_p (families 20, 29, 1, 17, 26, and 40 in blue in Fig 4) were used to complement if there were fewer than five families in the corresponding G_p . Next, we added the G_d consisting of the families with the highest R_d values in each G_p (the rectangle with family no. in blue in Fig 4). Finally, six G_d were included in the design and formed a block (areas with grey shading belonged to a complete block). All blocks were the same in the design (fixed design). Within each block, 19 (52.78%, families with yellow shading in Fig 4), 6 (16.67%, families with the blue shading in Fig 4), and 11 (30.56%, families without shading in Fig 4) best families (families with the highest R_d of volume in each G_p), worst families (families with the lowest R_d of volume in each G_p), and moderate families (the remaining families) were used. Consequently, the families were deployed unequally. Four blocks are shown in Fig 4.

3	6	12	11	3	6	22	21	3	6	12	11
20	1	7	18	30	1	7	28	20	1	7	18
19	15	5	8	29	25	5	8	19	15	5	8
14	10	2	4	24	10	2	4	14	10	2	4
13	16	9	17	23	26	9	27	13	16	9	17
3	6	22	21	3	6	12	11	3	6	22	21
30	1	7	28	20	1	7	18	30	1	7	28
29	25	5	8	19	15	5	8	29	25	5	8
24	10	2	4	14	10	2	4	24	10	2	4
23	26	9	27	13	16	9	17	23	26	9	27

Fig 5. Field deployment of the second-generation Chinese pine seed orchard in Xixian. Note: Number represents family no. Families in the same rectangle belong to the same block. Families in black, blue, and red fonts are the best, moderate, and control-pollinated families. The yellow and blue shaded areas represent the two kinds of complete blocks, and six blocks are shown.

doi:10.1371/journal.pone.0157646.g005

unbalanced, incomplete, fixed block (UIFB) design for the second-generation Chinese pine seed orchards in Xixian. We included twenty families in each block with single-tree plot, and considered the phylogenetic relationships among families. We divided the 20 families in the XX-OP selected population into two groups. Families 1–10 were the best families (with high R_d values; families in black font in Fig 5) and families 11–20 were the moderate families (with lower R_d values; families in blue font in Fig 5). The best families were deployed in every block, and the moderate families and control-pollinated families (families in red font in Fig 5) were deployed in every other block. The frequency of the best families was twice that of the moderate families and the control-pollinated families (unequal deployment) across the entire design.

Blocks containing the same families were of the same (fixed design). The two kinds of blocks (one containing families 1–20 with yellow shading in Fig 5; the other containing families 1–10 and 21–30 with blue shading in Fig 5) were alternately deployed in the adjacent blocks. Six blocks are shown in Fig 5.

Discussion

Strategies for selecting the second-generation selected populations of Chinese pine

Four second-generation base populations were used in this study. The ZN-NP base population consisted of bulked no-pedigree open-pollinated progeny of the first-generation Chinese pine seed orchard in Zhengning with a large planting scale. Neither of the paternal or maternal information in the ZN-NP base population were known. The ZN-OP and XX-OP base populations were composed of open-pollinated families of the first-generation Chinese pine seed orchards in Zhengning and Xixian. The maternal sources of the individuals were known but the paternal information was not. The XX-CP base population contained the control-pollinated families of the first-generation Chinese pine seed orchards in Xixian, with known parental information.

The second-generation selected populations were composed of individuals with relatively high R_d values in the corresponding base population. The mean R_d values in the selected ZN-NP, ZN-OP, XX-OP, and XX-CP populations were 0.83, 0.15, 0.25, and 0.20, respectively (Fig 1a–1d). The mean R_d value was significantly higher in the ZN-NP selected population than in the other populations (Fig 1e). This may have occurred because the ZN-NP base population comprised randomly mated progeny of all clones in the first-generation seed orchard, contained nearly all possible crosses needed for testing, and had relatively high genetic diversity and selection potential, whereas the open- and control-pollinated families contained only limited numbers of families that needed to be tested, and had relatively low levels of genetic diversity and selection potential. The mean R_d value was slightly but not significantly higher in the XX-OP selected population than in the XX-CP selected population (Fig 1e), which might be explained that although the XX-OP base population had higher number of tested families and level of genetic diversity, the XX-CP base population had better genetic basis, as all tested families were derived from pre-selected clones in the first-generation seed orchard.

In general, R_d values were significantly higher in selected populations derived from large-scale bulked no-pedigree open-pollinated progeny than in those derived from the common open- and control-pollinated progeny, indicating that elite families could be selected effectively from the large-scale open-pollinated progeny for future genetic improvement of Chinese pine, which could significantly increase the genetic gain of the target trait. In addition, we detected no differences between the open- and control-pollinated selected populations, but the control-pollination process is much more labor- and time-consuming. Thus, selecting elite families from the open-pollinated progeny could significantly shorten the breeding cycle and increase improvement efficiency.

Effectiveness of SSR markers in the phylogenetic analysis of Chinese pine progeny

SSR molecular markers developed in the 1980s are widely used in phylogenetic analysis and population genetics because of their neutral evolution, abundant polymorphic loci, even distribution, co-dominance, stability, and repeatability [42–43]. In this study, we screened 11 polymorphic SSR primers from studies of Chinese pine and related species [40]. We used these

primers to amplify stable, clear, polymorphic, and repeatable 120- to 240-bp bands in different Chinese pine individuals at annealing temperatures of 51–58°C (Table 1).

Inbreeding is a serious problem when attempting to improve tree traits because it can cause inbreeding depression and decrease the improvement effect [1, 2]. Advanced-generation seed orchards usually contain fewer clones derived from backward selection, forward selection, or both, compared with those of first-generation seed orchards [22], leading to a great increase in inbreeding rates. Consequently, the phylogenetic relationships among clones must be considered to minimize inbreeding and ensure improvement. Of the four selected populations, both maternal and paternal information was unknown for the ZN-NP population, and the paternal information was unknown for the ZN-OP and XX-OP populations. The representative ZN-NP and XX-OP selected populations were chosen for phylogenetic analysis. We used the 11 polymorphic primers screened in the previous section to identify the phylogenetic relationships among individuals in the ZN-NP and XX-OP selected populations. The amplified bands were polymorphic among individuals from the ZN-NP and XX-OP selected populations (Figs 2 and 3), with mean *PIC* values of 0.6202 (Table 2) and 0.5324 (Table 3), respectively. Families in the ZN-NP selected population had varying levels of phylogenetic relationship. Some families in the ZN-NP selected population had more than one individual, and the largest family had eight individuals (individuals 8, 9, 10, 11, 12, 14, 19, and 20) (Fig 3a), demonstrating that this family had a significant advantage in volume growth. It also reveals that there was a high probability that the same families would be selected multiple times if they were selected from the no-pedigree open-pollinated progeny, so more candidate individuals must be used to ensure a particular number of families. All individuals in the ZN-NP selected population were divided into five phylogenetic groups (G_p), and families within the same G_p were closely related, whereas families in different G_p were distantly related. For example, the family composed of individuals 28 and 29 and the family composed of individuals 15, 16, and 18 were in the same G_p . Similarly, individuals 13, 17, 22, 27, and 38 were in the same G_p (Fig 3a). The phylogenetic analysis of individuals in the XX-OP selected population showed that family 3 was distantly related to the other families. Families 5, 12, and 14, as well as families 2, 7, and 20 were closely related (Fig 3b). Although the exact parents of the individuals were not identified using the SSR molecular markers, the relative phylogenetic relationships among individuals in the selected population were revealed. Inbreeding can be avoided by deployment designs that consider the relative phylogenetic relationships among individuals and families. Consequently, SSR molecular markers could be used efficiently during marker-assisted selection and to determine the phylogenetic relationships among individuals within the Chinese pine population.

Genetic distance-based field deployments for second-generation Chinese pine seed orchards

Seed orchards are designed and established to produce superior seeds or other products [44, 45]. The most important purpose of a seed orchard is to maximize the genetic gain of target traits while maintaining certain levels of genetic diversity. The effective population size, mating system, and flowering phenology can directly or indirectly affect genetic diversity or genetic gain and are major considerations in orchard deployment design [22]. The first-generation seed orchards in Zhengning and Xixian adopted an ordinal offset staggered design, which is feasible for first-generation seed orchards because the number of clones is relatively high and the clones are distantly related. Advanced-generation seed orchard designs are much more complex because of the limited number of clones, diverse sources of the clones (backward selection, forward selection, and so on) and higher level of inbreeding probability.

The second-generation Chinese pine seed orchards in Zhengning adopted a GUCFB design (Fig 4). The 23 families in the secondary-selected population at Zhengning were divided into six deployment groups (G_d) according to their relative phylogenetic relationships. Each G_d contained six families, with one best individual each, derived from a different G_p . Families in the same G_d were distantly related. In each block, the proportions of best, moderate, and worst families were 52.78%, 30.56%, and 16.67%, respectively. All 23 families were included in a block, and the blocks were the same across the entire design (Fig 4). The grouped fixed design increases the possibility of outbreeding and allows closely related progeny to be used effectively in an advanced-generation seed orchard. Following a method reported previously [25, 46], we used an unbalanced block design, which increased the mean R_d value from 1.00 in the balanced design to 1.76, thus maximizing genetic gain at the same level of genetic diversity.

Twenty open-pollinated families (Fig 1c) and 10 control-pollinated families (top 10 families in Fig 1d) were included in the second-generation seed orchard at Xixian. A UIFB design was adopted. Twenty families were included in each block. The 10 best families were deployed in every block, and the 10 moderate and 10 control-pollinated families were deployed in every other block. The relative positions were fixed (Fig 5); this is the recommended systematic design for advanced-generation seed orchards [22]. We considered the relative phylogenetic relationships among the families when determining the relative positions of the families to maintain a certain distance between related families and to control inbreeding. We also considered an unbalanced design, in which the number of best families was twice that of moderate and control-pollinated families across the entire design. This unbalanced design increased the mean R_d value from 0.24 in the balanced design to 0.34, which increased genetic gain at the same level of genetic diversity.

Acknowledgments

We thank the workers in the Shanxi and Zhengning first-generation Chinese pine seed orchards for sample collection.

Author Contributions

Conceived and designed the experiments: YL YAE WL. Performed the experiments: HY SN. Analyzed the data: HY SN. Wrote the paper: HY YL WL.

References

1. Namkoong G. Inbreeding effects on estimation of genetic additive variance. *Forest Sci.* 1966; 12: 8–13.
2. Griffin AR, Lindgren D. Effect of inbreeding on production of filled seed in *Pinus radiata*—experimental results and a model of gene action. *Theor Appl Genet.* 1985; 71(2): 334–343.
3. Pena SDJ, Chakraborty R. Paternity testing in the DNA era. *Trends Genet.* 1994; 10: 204–209. PMID: [8073534](#)
4. Luikart G, England PR. Statistical analysis of microsatellite DNA data. *Trends in Ecol Evol.* 1999; 14: 253–256.
5. Nielsen R, Mattila DK, Clapham PJ, Palsboll PJ. Statistical approaches to paternity analysis in natural populations and applications to the Northern Atlantic humpback whale. *Genetics.* 2001; 157(4): 1673–1682. PMID: [11290722](#)
6. Neff BD, Repka J, Gross MR. Parentage analysis with incomplete sampling of parents and offspring. *Mol Ecol.* 2000; 9: 515–528.
7. Jones AG, Ardren WR. Methods of parentage analysis in natural populations. *Mol Ecol.* 2003; 12: 2511–2523. PMID: [12969458](#)
8. Grattapaglia D, Ribeiro VJ, Rezende GDSP. Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. *Theor Appl Genet.* 2004; 109: 192–199. PMID: [15004676](#)

9. El-Kassaby YA, Lstibůrek M. Breeding without breeding. *Genet Res.* 2009; 91(02): 111–120.
10. El-Kassaby YA, Cappa EP, Liewlaksaneeyanawin C, Klápště J, Lstibůrek M. Breeding without breeding: Is a complete pedigree necessary for efficient breeding. *PLoS One.* 2011; 6(10): e25737. doi: [10.1371/journal.pone.0025737](https://doi.org/10.1371/journal.pone.0025737) PMID: [21991342](https://pubmed.ncbi.nlm.nih.gov/21991342/)
11. El-Kassaby YA, Klápště J, Guy RD. Breeding without breeding: selection using the genomic best linear unbiased predictor method (GBLUP). *New Forest.* 2012; 43: 631–637.
12. Wang X, Torimaru T, Lindgren D, Fries A. Marker-based parentage analysis facilitates low input 'breeding without breeding' strategies for forest trees. *Tree Genet Genomes.* 2010; (6): 227–235.
13. Korecký J, Klápště J, Tibůrek ML, Koblíha J, Son CDN, Assaby YAE. Comparison of genetic parameters from marker-based relationship, sibship, and combined models in Scots pine multi-site open-pollinated tests. *Tree Genet Genomes.* 2013; 9: 1227–1235.
14. Zhao P, Zhang S, Woeste K. Genotypic data changes family rank for growth and quality traits in a black walnut (*Juglans nigra* L.) progeny test. *New Forest.* 2013; 44: 357–368.
15. Massah N, Wang J, Russell JH, Van Niejenhuis A, El-Kassaby YA. Genealogical relationship among members of selection and production populations of yellow cedar (*Callitropsis nootkatensis* [D. Don] Oerst.) in the absence of parental information. *J Hered.* 2010; 101: 154–163. doi: [10.1093/jhered/esp102](https://doi.org/10.1093/jhered/esp102) PMID: [19945992](https://pubmed.ncbi.nlm.nih.gov/19945992/)
16. Moriguchi Y, Yomogida H, Iwata H, Takahashi M, Taira H, Tsumura YK. Paternity analysis in a progeny test of *Cryptomeria japonica* revealed adverse effects of pollen contamination from outside seed orchards on morphological traits. *Tree Genet Genomes.* 2011; 7: 1089–1097.
17. Shimono A, Wang X, Torimaru T, Lindgren D, Karlsson B. Spatial variation in local pollen flow and mating success in a *Picea abies* clone archive and their implications for a novel "breeding without breeding" strategy. *Tree Genet Genomes.* 2011; 7: 499–509.
18. Ng C, Lee S, Ng KK, Lee C, Muhammad N. Establishment of full-sib families in *Shorea platyclados* using paternity analysis: an alternative to controlled pollination. *J For Res.* 2012; 17: 479–487.
19. Wang J, El-Kassaby YA, Ritland K. Estimating selfing rates from reconstructed pedigrees using multilocus genotype data. *Mol Ecol.* 2012; 21(1): 100–116. doi: [10.1111/j.1365-294X.2011.05373.x](https://doi.org/10.1111/j.1365-294X.2011.05373.x) PMID: [22106925](https://pubmed.ncbi.nlm.nih.gov/22106925/)
20. Klápště J, Lstibůrek M, El-Kassaby YA. Estimates of genetic parameters and breeding values from western larch open-pollinated families using marker-based relationship. *Tree Genet Genomes.* 2014; 10: 241–249.
21. Weng Y, Park Y S, Krasowski M J. Managing genetic gain and diversity in clonal deployment of white spruce in New Brunswick, Canada. *Tree Genet Genomes.* 2010; 6(3): 367–376.
22. Hodge GR, White TL. Advanced-generation wind-pollinated seed orchard design. *New Forest.* 1993; 7: 213–236.
23. Bell GD, Fletcher AM. Computer organised orchard layouts (COOL) based on the permuted neighbourhood design concept. *Silvae Genet.* 1978; 27(6): 223–225.
24. El-Kassaby YA, Fayed M, Klápště J, Lstibůrek M. Randomized, replicated, staggered clonal-row (R2SCR) seed orchard design. *Tree Genet Genomes.* 2014; 10(3): 555–563.
25. Weng Y, Park Y S, Lindgren D. Unequal clonal deployment improves genetic gains at constant diversity levels for clonal forestry. *Tree Genet Genomes.* 2012; 8: 77–85.
26. Ahlinder J, Mullin TJ, Yamashita M. Using semidef inite programming to optimize unequal deployment of genotypes to a clonal seed orchard. *Tree Genet Genomes.* 2014; 10(1): 27–34.
27. Ndez JSF, Nez SCGL. Allocating individuals to avoid inbreeding in ex situ conservation plantations: so far, so good. *Conserv Genet.* 2009; 10: 45–57.
28. Xu HC. Chinese pine. Beijing: China Forestry Publishing. 1993.
29. Wu CY. Vegetation of China, 2nd ed. Beijing: Science Press. 1995.
30. Chen KM, Abbott RJ, Milne RI, Tian XM, Liu JQ. Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant species of coniferous forest in northern China. *Mol Ecol.* 2008; 17: 4276–4288. PMID: [19378405](https://pubmed.ncbi.nlm.nih.gov/19378405/)
31. Wang SS, Shen XH, Yao LH, Li Y, Li JG, Li AD. A comparative study on growth patterns of *Pinus tabulaeformis* clones in Xingcheng seed orchard, Liaoning province. *J Beijing For Univ.* 1985; 7(4): 72–83.
32. Chen BW, Zheng Y, Shen XH. Studies on selection methods and criteria of plus trees in *Pinus tabulaeformis* Carr. *J Beijing For Univ.* 1992; 14(1): 7–13.
33. Wang XR, Shen XH. Studies on increasing seed production in seed orchards of *Pinus tabulaeformis* Carr.—An analysis of seed losses caused by aborted ovules and emptied seeds. *J Beijing For Univ.* 1989; 11(3): 60–65.

34. Zhang HX, Shen XH. Progress on reproductive system of forest seed orchards. *Scientia Silvae Sinicae*. 2002; 38(2): 129–134.
35. Li W, Wang XR, Li Y. Stability in and correlation between factors influencing genetic quality of seed lots in seed orchard of *Pinus tabulaeformis* Carr. over a 12-year span. *PLoS One*. 2011; 6(8): e23544. doi: [10.1371/journal.pone.0023544](https://doi.org/10.1371/journal.pone.0023544) PMID: [21887269](https://pubmed.ncbi.nlm.nih.gov/21887269/)
36. Li W, Wang XR, Li Y. Variation in floral phenological synchronization in a clonal seed orchard of *Pinus tabuliformis* in Northeast of China. *Silvae Genet*. 2012; 61(4–5): 133–142.
37. Yuan HW, Li ZX, Fang P, Li W, Li Y. Variation and stability in female strobili production of a first-generation clonal seed orchard of Chinese Pine (*Pinus tabuliformis*). *Silvae Genet*. 2014; 63(1–2): 41–47.
38. Yuann HW, Niu SH, Zhou XQ, Du QP, Li Y, Li W. Evaluation of seed production in a first-generation seed orchard of Chinese pine (*Pinus tabuliformis*). *J For Res*. 2016; doi: [10.1007/s11676-016-0238-x](https://doi.org/10.1007/s11676-016-0238-x)
39. Porebski S, Bailey L, Baum B. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep*. 1997; 15: 8–15.
40. Zhang DM, Yang Y, Shen XH, Ru GX. Selection of primers and establishment of SSR-PCR reaction system on *Pinus tabulaeformis* Carr. *J Beijing For Univ*. 2007; 29(2): 13–17.
41. Fang P, Niu S, Yuan H, Li Z, Zhang Y, Yuan L, et al. Development and characterization of 25 EST-SSR markers in *Pinus sylvestris* var. *mongolica* (Pinaceae). *Appl Plant Sci*. 2014; 2(1): 1300057.
42. Gomes F, Costa R, Ribeiro MM, Figueiredo E, Canhoto JM. Analysis of genetic relationship among *Arbutus unedo* L. genotypes using RAPD and SSR markers. *J For Res*. 2016; 24(2): 227–236.
43. Lu YQ, Zhang JL, Tong ZK. Exploit SSR in *Cryptomeria fortunei* and application in genetic diversity analysis. *J For Res*. 2015; 26(4): 963–969.
44. Burczyk J, Chalupka W. Flowering and cone production variability and its effect on parental balance in a Scots pine clonal seed orchard. *Ann Forest Sci*. 1997; 54: 129–144.
45. Kang KS, Lindgren D. Fertility variation among clones of Korean pine (*Pinus koraiensis* S. et Z.) and its implications on seed orchard management. *Forest Genet*. 1999; 6(3): 191–200.
46. Lindgren D, Danusevicius D, Rosvall O. Unequal deployment of clones to seed orchards by considering genetic gain, relatedness and gene diversity. *Forestry*. 2009; 82(1): 17–28.