



Clonal variations in cone, seed and nut traits in a *Pinus koraiensis* seed orchard in Northeast China

David Kombi Kaviriri^{1,3,4} · Yuxi Li¹ · Dawei Zhang¹ · Hongtao Li² · Zuoyi Fan² · Jingyuan Wang² · Lianfu Wang² · Qi Wang² · Deqiu Wang² · Vincent L. Chiang¹ · Xiyang Zhao¹

Received: 6 September 2019 / Accepted: 14 November 2019 / Published online: 11 January 2020
© Northeast Forestry University 2020

Abstract Korean pine (*Pinus koraiensis* Siebold & Zucc.) in Northeastern China has been genetically improved to increase seed yields in addition to timber. To assess seed yield variability and select highly productive clones, 14 cone, seed and nut traits were measured and analyzed. Variance analysis showed that all clones were significantly different in various traits ($P < 0.01$). Phenotypic coefficients of variation and repeatability of traits ranged from 9.1 to 34.4% and from 27.5 to 97.7%, respectively. Except for the cone layer and cone seed numbers, the other traits were positively or negatively correlated. Three principal components were identified. Seed and nut traits were the most important traits in the first principal component, and cone traits more important in the second. Using correlation and

principal component analyses, cone number and other traits were selected to evaluate materials. Twenty-two clones were selected using a selection rate of 10% based on cone number independently or other combined traits. The genetic gain for different traits ranged from 6.2 to 24.3%. The selected elite clones can supply seedlings for reforestation and the selection method can provide a theoretical basis for selection in other conifer species.

Keywords *Pinus koraiensis* · Clones · Cone · Seed · Nut traits · Principal component analysis

Introduction

Korean pine, *Pinus koraiensis* Siebold & Zucc., is one of the most economically important tree species in Northeastern China (Li and Löfgren 2000). Apart from its excellent wood properties, *P. koraiensis* nuts have significant nutritional and therapeutic value (Zhang et al. 2015). The use of Korean pine nuts as a nutritional source can be traced back to the Paleolithic era in eastern Asia (Rosengarten 2004). China is a major producer of pine nuts in the world, with most exported to Western countries (Destaillets et al. 2010), generating more than US \$ 250 million dollars each year (Man et al. 2012).

Because of its important economic value, *P. koraiensis* is much in demand for reforestation in Northeastern China (Ren et al. 2018). However, planting stock with desirable genetic properties is needed to meet the planting demand, but due to the difficulty of producing large numbers by cottage, tissue culture and other asexual methods, seed orchards have been an important step to improve the species (Wang and Hong 2004; Feng et al. 2010). In China, several *P. koraiensis* seed orchards were established in the

Project funding: The work was supported by the Forestry Science Technology and Development Project (KJZXSA2019046) and Fundamental Research Funds for the Central Universities (2572017DA02).

The online version is available at <http://www.springerlink.com>

Corresponding editor: Tao Xu.

✉ Xiyang Zhao
zhaoxyphd@163.com

¹ State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, People's Republic of China

² Forest Cultivation Center, Linjiang Forestry Bureau of Jilin, Linjiang 134600, People's Republic of China

³ Department of Natural Resources Management, Faculty of Agricultural Sciences, University of Kinshasa, Kinshasa, Democratic Republic of Congo

⁴ Department of Forestry, Faculty of Agricultural Sciences, Semuliki Official University, Beni, Democratic Republic of Congo

beginning of the 1960s. To date, some orchards have produced improved varieties based on progeny tests or clonal evaluations (Liang et al. 2019). Because of the demand for timber over the past several years, improved varieties of *P. koraiensis* spp. were selected based only on their growth traits or wood properties (Wang et al. 2018; Liang et al. 2019). But with the official decree in 2016 against the cutting of natural forests, we pay more attention to seeds and cone of *P. koraiensis*.

Characteristics of the cone and seeds, nut numbers and shape, are important traits to evaluate for improving seed yield (Davis 1967; Matziris 1998; Roy et al. 2004; Salvatore et al. 2010; Rawat and Bakshi 2011). In this study, 110 clones of 38-year-old *P. koraiensis* were selected and 14 cone, seed and nut characteristics were investigated and analyzed (Table 1). The primary objectives were to: (1) compare the performance values of different *P. koraiensis* clones for each trait; (2) estimate the genotypic and variation parameters of different traits; (3) identify the relationship among different traits; and, (4) perform a comprehensive assessment of each clone and predict selection gains.

Materials and methods

Site description

Data were collected from the Naozhi forestry seed orchard located on the western hillside of Changbai Mountain, Linjiang City, Jilin Province, Northeast China (41°05' N, 126°06' E). Natural *P. koraiensis* forest and broad-leaved mixed forest are typical of this region 700 to 1100 m. The region has

a moderate temperate and a continental monsoon climate with average annual temperatures, rainfall, and frost-free period is 5 °C, 800 mm, and 135 days, respectively. The soil is Albi-boric Argosols according to the U.S. soil taxonomy (Xu et al. 2018), dominated by dark brown soil and about 40 cm thick, with a textural proportion of sand (15.1%), silt (63.3%), and clay (21.6%) (Zhu et al. 2010).

Materials

In 1979, 110 clones (Table 2) were selected, based on growth traits from the natural distribution of *P. koraiensis* in Linjiang City. Cuttings were grafted the following year and plantations were established with 4-year-old plants in the spring of 1984. The experimental design consisted of 10 large blocks, and each containing five small blocks. Each clone was planted using a randomized complete block design containing one tree at 7.0 m × 7.0 m in each small block.

Trait measurements

The cones were collected at maturity. Cones from each ramet were counted (cone number, CN) and directly weighed (cone weight, CWei). In order to have a stable cone number per ramet, as cone numbers are different each year, previous years data on cone number were taken into account and cone number recorded for 5 years (from 2014 to 2018). Averages were calculated and used in the analysis. Measurements of the other traits were only for the study year (2018). Cone lengths (CL) and widths (CWd), (diagonally at the widest portion), were measured with a digital caliper. Scales

Table 1 ANOVA, genetic and variation parameters on 110 *Pinus koraiensis* clones in the Naozhi seed orchard

Traits	Unit	Mean	Min	Max	SD	df	MS	F	PCV (%)	R
Cones per clone	–	18.1	11.0	26.0	6.2	109	44.7**	1.4	34.3	27.5
Cone length	(cm)	132.9	11.6	201.1	17.6	109	2502.8***	13.7	13.2	92.7
Cone width	(cm)	70.4	24.5	125.4	10.8	109	1536.5***	43.8	15.3	97.7
Cone weight	(g)	216.7	80.0	813.0	55.6	109	21,799.7***	10.8	25.7	90.7
Cone layer number	–	13.5	5.0	24.0	2.3	109	38.3***	10.5	17.4	90.5
Seed number per cone	–	137.7	42.0	268.0	26.6	109	3230.7***	5.7	19.3	82.5
Seed length	(mm)	16.4	10.2	21.2	1.5	109	14.2***	10.3	9.1	90.2
Seed width	(mm)	10.8	9.3	12.7	1.3	109	7.3***	6.0	11.8	83.2
Single seed weight	(g)	0.7	0.5	1.0	0.2	109	0.2***	15.1	21.8	93.4
Coat thickness	(mm)	0.9	0.8	1.1	0.1	109	11.4***	11.4	13.5	91.3
Thousand seed weight	(g)	724.9	614.8	1098.8	132.0	109	2.6***	5.5	18.2	81.8
Nut length	(mm)	13.4	5.6	19.7	1.3	109	0.0***	7.0	9.7	85.6
Nut width	(mm)	7.1	4.1	11.2	0.8	109	0.1***	5.0	11.0	80.1
Nut weight	(g)	0.3	0.2	0.4	0.1	109	40,456.3***	3.9	30.0	74.5

SD, df, MS, PCV, *F* and *R*, are standard deviation, degrees of freedom, mean square, phenotypic coefficient of variance, the *F* Fisher-Yates coefficients and Repeatability respectively

***Represents very significant differences and **represents strongly significant differences

Table 2 Comprehensive evaluation of traits investigated in 110 *P. koraiensis* clones

Clones	Qi	Clones	Qi	Clones	Qi	Clones	Qi	Clones	Qi
PK52	2.86	PK23	2.70	PK80	2.63	PK24	2.60	PK47	2.54
PK60	2.81	PK4	2.70	PK75	2.63	PK78	2.60	PK88	2.53
PK110	2.80	PK13	2.70	PK95	2.63	PK97	2.59	PK82	2.53
PK49	2.79	PK22	2.70	PK93	2.63	PK30	2.59	PK1	2.53
PK84	2.78	PK35	2.69	PK67	2.62	PK14	2.59	PK108	2.52
PK101	2.78	PK7	2.69	PK11	2.62	PK79	2.59	PK74	2.52
PK42	2.77	PK58	2.69	PK5	2.62	PK73	2.59	PK45	2.52
PK106	2.76	PK105	2.68	PK64	2.62	PK103	2.58	PK99	2.51
PK3	2.76	PK18	2.68	PK41	2.62	PK89	2.58	PK100	2.51
PK56	2.75	PK8	2.68	PK9	2.62	PK37	2.58	PK29	2.50
PK39	2.75	PK51	2.68	PK68	2.62	PK50	2.56	PK46	2.50
PK17	2.75	PK12	2.68	PK61	2.61	PK48	2.55	PK81	2.49
PK109	2.75	PK6	2.67	PK59	2.61	PK102	2.55	PK76	2.49
PK55	2.74	PK57	2.67	PK2	2.61	PK96	2.55	PK54	2.48
PK43	2.74	PK33	2.67	PK16	2.61	PK77	2.55	PK21	2.48
PK32	2.72	PK94	2.66	PK65	2.61	PK28	2.54	PK90	2.45
PK34	2.72	PK38	2.66	PK83	2.60	PK107	2.54	PK70	2.45
PK25	2.72	PK36	2.65	PK10	2.60	PK71	2.54	PK87	2.44
PK19	2.72	PK63	2.65	PK91	2.60	PK98	2.54	PK66	2.42
PK20	2.72	PK104	2.65	PK40	2.60	PK85	2.54	PK72	2.42
PK15	2.72	PK27	2.65	PK31	2.60	PK62	2.54	PK44	2.41
PK69	2.71	PK26	2.64	PK92	2.60	PK53	2.54	PK86	2.38

were counted per layer, and the number of layers/cone were recorded (CLN). After measurements were recorded, seeds were removed and the number of seeds per cone per ramet were counted (CSN). To calculate clone thousand seed weight (TSWei), seeds from different ramets were mixed, 400 seeds were randomly chosen and divided into four equal portions for replication, and every 100 seeds were weighed. A thousand seed weight was obtained by extrapolation. To determine seed characteristics, 50 seeds from each clone were randomly chosen and measured for length (SL), width (SWd), and weight (SWei). Finally, nuts were extracted from the 50 seeds and length (NL), width (NWd), and weight (NWei) were determined. Coat thickness (CTH) was measured 10 times for each seed using a Vernier caliper, and an average value were taken.

Data analyses

One-way analyses of variance (F test) were performed for the 14 traits using R Software (Version 2.5-3). Clonal and environment/random effects were estimated using IBM SPSS, version 20 (Field 2013) and the linear model Eq. (1) (Li et al. 2017)

$$Y_{ij} = \mu + a_i + \epsilon_{ij}, \tag{1}$$

where Y_{ij} is the performance of individual tree j of clone i , μ is the overall mean, a_i is the clone effect, and ϵ_{ij} is random error. The phenotypic coefficient of variance (PCV) in Eq. (2) expresses the range of individual tree performances compared to the clonal mean values for a given trait. PCV was calculated using the ratio of the standard deviation (SD) to the average values (\bar{X}) (Brancher et al. 2019).

$$PCV = \frac{SD}{\bar{X}}, \tag{2}$$

Repeatability was calculated from the clone averages to estimate the stability of traits in different ramets (Xiao et al. 2019). And it calculated by subtracting the ratio of the F (Fisher-Yates) coefficients determined by ANOVA ($1/F$) from the unit value (1) (Zheng et al. 2015).

$$R = 1 - \frac{1}{F} \tag{3}$$

The relationships between characteristics were estimated by Pearson’s correlation coefficient ($r_{A(x,y)}$), which was the partition of the covariance of related traits $COV_{P(x,y)}$ divided by the result of multiplying their respective variances $\sigma_p(x) \times \sigma_p(y)$ (Zhao et al. 2016).

$$r_A(xy) = \frac{COV_{P(x,y)}}{\sigma_p(x)\sigma_p(y)}, \tag{4}$$

Considering the correlation matrix, factor analysis was performed by principal components using PAST software (Hammer and Harper 2001). It determined the major traits that explained most of the observed variance. These traits were the most suitable to be used to generate the clone selection index (Ait-Sahalia and Xiu 2019). The highlighted traits were then used to perform a comprehensive evaluation using the Q_i value. Q_i was obtained by squaring the sums of the ratios of the clonal averages to the maximum values of overall means of the clone (Zhao et al. 2016). A selection rate of 10% was applied to retain elite clones for future seed production.

$$Q_i = \sqrt{\sum_{j=1}^n a_i}, \quad (5)$$

Selection success was verified by the genetic gain obtained for the selected clones. The genetic gain was calculated by the product of the repeatability R , and the ratio of the selection difference W/\bar{X} , for a given trait (Gonçalves et al. 2019).

$$\Delta G = R \times W/\bar{X}, \quad (6)$$

Finally, the expected economic gain of the Naozhi seed orchard was considered as a function of seed yield. It was estimated by extrapolating the unitary economic value per kilogram of seeds multiplied by the recorded production and realized gain in seed weight.

$$\Delta Eco. = \text{Price} \times \text{unitary seed weight} \times \text{realized gain in seed weight}, \quad (7)$$

Results

Average values of different traits

Variations in the traits were distinct among clones, with moderate to high variation between clone averages. Overall mean values, minimum and maximum values, standard deviations (SD) and phenotypic coefficients of variance (PCV) are shown in Table 1. With regards to cone characteristics, overall means of cone numbers, lengths and widths were 18.1 cm, 132.9 cm, and 70.4 cm, values ranged from 11.0 to 26.0 cm, 11.6 to 201.1 cm, and 24.5 to 125.4 cm, respectively. Cone weight varied from 80.0 to 813.0 g, with a mean of 216.7 g. The number of layers per cone and seeds per cone ranged from 5 to 24 layers and from 42 to 268, with overall averages of 13.5 layers and 137.5 seeds, respectively. Traits related to single seed dimensions showed large average of 16.4 mm, 10.8 mm, 0.9 mm, and 0.7 g in length, width, coat thickness, and single seed weight, respectively. A value of 724.9 g was calculated for a thousand seed weight.

For nut traits, average values of length width and weight were 13.4 mm, 7.1 mm and 0.3 g respectively.

ANOVA, PCV and R

ANOVA results show that clone performances were significantly different for the traits at probability levels of 0.001 and 0.01 (Table 1). The coefficient of variance (PCV) for all cone, seed and nut traits varied from 9.1 to 34.3% (Table 1). The highest PCV value (34.3%) was recorded for cone number, followed by 29.8% for nut weight and 25.7% for cone weight. The lowest PCVs, below 15%, were found for seed length (9.1%), nut length (9.7%), nut width (11.0%), and seed width (11.0%). The repeatability values (R) for all traits ranged from 27.5 to 93.4% (Table 1) with most having high repeatability values ($R > 0.80$). Cone dimensions (CL, CWd, CWei, CLN, and CSN) were strongly inherited ($R \geq 90\%$) as well as seed traits (SL, SWd, SWei and TSWei, higher than 80%). Low R values for cones per clone were 27.5%.

Correlation analysis

The correlation coefficients between the investigated traits are shown in Table 3. Cone number (CN) was negatively correlated to the other traits, ranging from -0.382 (CN with nut length, NL) to -0.005 (CN with seeds/cone/ramet, CSN). Cone length (CL), width (CWd) and weight (CWei)

were positively correlated with each other and with the other traits (r ranged from 0.196 to 0.799). In addition, CN, seed length (SL), seed width (SWd), seed weight (SWei), nut length (NL), nut width (NWd), nut weight (NWei), seed coat thickness (CTH) and thousand seed weight (TSWei) were positively correlated, with coefficients ranging from -0.382 (CN with NL) to 0.937 (SL with NL). There were no significant correlations between layers/cone (CLN) and the other traits except for CL, CWd, CWei and seeds/cone/ramet (CSN) which showed significant positive correlations (r ranging from 0.311 to 0.647). Excluding SWd, SWei, NWei, CTH and TSWei, CLN was not significantly correlated with the other traits (r ranged from -0.369 to 0.565).

Principal component analysis

Three principal components with high eigenvalues were obtained from the principal components analysis (PCA) with a cumulative contribution of 75.56% (Table 4). From the first principal component (PC I), seed and nut traits (SL, SWd, SWei, NL, NWd, NWei and TSWei) were the most important with

Table 3 Pearson correlation coefficients of each investigated traits in 110 *P. koraiensis* clones

Traits	CN	CL	CWd	CWei	CLN	CSN	SL	SWd	SWei	NL	NWd	NWei	CTH	TSWei
CN	1													
CL	-0.278**	1												
CWd	-0.250**	0.799**	1											
CWei	-0.188*	0.448**	0.382**	1										
CLN	-0.132	0.568**	0.647**	0.311**	1									
CSN	-0.005	0.487**	0.475**	0.505**	0.633**	1								
SL	-0.369**	0.509**	0.428**	0.565**	0.182	0.246**	1							
SWd	-0.352**	0.502**	0.421**	0.466**	0.186	0.165	0.821**	1						
SWei	-0.376**	0.486**	0.415**	0.491**	0.141	0.175	0.874**	0.871**	1					
NL	-0.382**	0.506**	0.412**	0.492**	0.136	0.196*	0.937**	0.771**	0.879**	1				
NWd	-0.236*	0.422**	0.414**	0.523**	0.104	0.237*	0.743**	0.829**	0.836**	0.745**	1			
NWei	-0.308**	0.481**	0.446**	0.383**	0.093	0.127	0.780**	0.792**	0.891**	0.828**	0.853**	1		
CTH	-0.311**	0.304**	0.196*	0.334**	0.139	0.150	0.495**	0.533**	0.646**	0.477**	0.383**	0.447**	1	
TSWei	-0.242*	0.378**	0.257**	0.227*	-0.034	0.008	0.599**	0.654**	0.753**	0.634**	0.611**	0.811**	0.496**	1

**Correlation is significant at the 0.01 level (2-tailed) *Correlation is significant at the 0.05 level (2-tailed) CN, CL, CWd, CWei, CLN, CSN, SL, SWd, SWei, NL, NWd, SWei, NL, NWd, NWei, CTH, and TSWei are cone number, cone length, cone width, cone width, cone weight, layers/cone, seed/cone, seed length, seed width, seed weight, nut length, nut width, nut weight, seed coat thickness, and thousand seed weight, respectively. The same abbreviations are used in the table below

Table 4 Variance and cumulative variance of different components

Component	Total explained variance	Variance (%)	Cumulative variance (%)
1	7.32	52.25	52.25
2	2.30	16.45	68.70
3	0.96	6.86	75.56
4	0.84	5.99	81.55
5	0.70	4.99	86.53
6	0.40	2.87	89.40
7	0.39	2.78	92.18
8	0.34	2.46	94.64
9	0.26	1.82	96.46
10	0.21	1.51	97.97
11	0.13	0.94	98.91
12	0.06	0.42	99.33
13	0.05	0.36	99.69
14	0.04	0.31	100.00

Table 5 Component values and genetic gain of different traits

Trait	PC I	PC II	PC III	Genetic gain (%)
CN	-0.42	0.05	0.77	7.2
CL	0.68	0.51	-0.16	9.6
CWd	0.61	0.59	-0.17	14.1
CW	0.61	0.30	0.32	17.9
LN	0.33	0.81	-0.13	8.4
CSN	0.36	0.74	0.28	8.0
SL	0.90	-0.13	0.07	7.3
SWd	0.89	-0.18	0.03	7.1
SWei	0.94	-0.24	0.02	24.3
NL	0.89	-0.18	0.03	8.4
NW	0.85	-0.17	0.26	6.2
Nwei	0.89	-0.26	0.07	22.7
CTH	0.60	-0.15	-0.20	6.9
TSWei	0.73	-0.38	-0.02	11.4

PC I, PC II and PC III are the principal components I, II and III

high eigenvalue values (0.73 to 0.94). From the second component (PC II), number of layers, seeds per cone, cone length and cone width had high positive values (\hat{c} 0.51). In PC III, only cone number per clone had a high positive absolute value.

Comprehensive evaluation and genetic gain

The number of cones is one of the most important traits of *P. koraiensis*, but the correlation analysis indicated that cone number was negatively correlated with the other traits; consequently, cone number was evaluated separately. Eleven clones (PK14, PK75, PK24, PK88, PK44, PK4, PK5, PK103, PK82, PK72, PK2) showed the highest performance at a selection rate of 10% based on cone number, and the genetic gain was 7.2% (Table 5). PCA results show that cone seed and nut traits, including SL, SWd, SWei, NL, NWd, NWei, TSWei, CLN and CSN, could be used together to perform a comprehensive evaluation of *P. koraiensis* clones. The Qi values of the clones based on the selected traits are shown in Table 2. The eleven clones (PK52, PK60, PK110, PK49, PK84, PK101, PK42, PK106, PK3, PK56 and K39) were selected as elite clones with a selection rate of 10%. The corresponding genetic gains ranged from 6.2 to 24.3%. The greatest genetic gains were found in seed weight, following by nut weight (22.7%), cone weight (17.9%), and cone width (14.1%) (Table 5).

Discussion

ANOVA and variation parameters of different traits

Average values and their comparison through variance analysis enable the selection of genetic resources in tree

improvement research. In fact, an analysis of variance helps to characterize phenotypic variations and variance components among individual trees (Guerra et al. 2016), while average performance values are used to establish a succession order for selection of improved material (clone or families) in different traits according to their ranks (Hietz et al. 2017). Cones, seeds and nuts are the most important traits considered by scientists and foresters in the breeding of economically important coniferous tree species, but their use and output are different because of the sizes and nutritional requirements of different species. The average values for the cone traits (number, length, width, weight) in this study established a succession order to carry out a valuable selection. However, these values are higher than those in studies on *Pinus halepensis* Mill. (Matziris 1998), *Pinus roxburghii* Sarg. (Roy et al. 2004) and *Pinus sylvestris* L. (Sivacioglu 2010). Results from previous studies indicate that *P. koraiensis* seeds are more suitable for consumption, supporting the basis of this current study for a more in depth seed trait characterization for a breeding program in North-east China.

Variation parameters such as the phenotypic coefficient of variation and repeatability are basic elements for selecting material for improvement programs, as these elements explain the significance of degrees of difference in the materials (Pan et al. 2019), and the range of variability and stability of traits in tree populations (Bai et al. 2019). The traits investigated in this study showed significant differences among clones, indicating that the methods of this study were effective. Our results are in accordance with previous studies that estimated the variation parameters in cone and seed traits for *Pinus wallichiana* A. B. Jacks. (Rawat and Bakshi

2011), *P. sylvestris* L. (Sivacioglu and Ayan 2008) and *P. koraiensis* (Miyaki 1987). The repeatability of most of the traits was higher than 80%, indicating that the production and quality of seeds from *P. koraiensis* are stable and effective. Our repeatability values are similar to those of Yuan et al. (2016) for *Pinus tabuliformis* Carr. However, they were markedly higher than those reported by Matziris (1998) and Sivacioglu (2010) for *P. sylvestris* and *P. halepensis* with 70% and 80% heritability, respectively.

Correlation analysis

Correlation coefficients reveal the degree of correlation between different traits and are thus useful for the selection of multiple characters (Xia et al. 2016). In this study, Cone numbers were negatively correlated with all the other traits. This is in contrast to Jiang et al. (2019) with *P. koraiensis* who concluded that cone number was significantly and positively correlated with the number of seeds per cone and thousand seed weight. The results may vary from ours due to differences in investigative methods and trees that were younger than ours. In this study, cone number was recorded for five consecutive years and our results are more representative and consistent with the logic of energy conservation, indicating that the evaluation and selection of the materials were more stable and effective. The number of cone layers and seed number are two important traits that can impact seed yield though both traits have rarely been investigated. In this study, both traits were significantly correlated with other cone traits but weakly correlated with other seed or nut traits which indicate that both traits have little effect on seed or nut traits. Seed traits showed a significant positive correlation with nut traits, corroborating similar results reported for *P. sylvestris* (Sivacioglu and Ayan 2008; Sevik and Topacoglu 2015). Therefore, we suggest that seed size can affect the size of nuts and that both traits can be evaluated and selected together.

Principal component analysis (PCA)

PCA is an important tool because it reduces a large number of variables into a small number of common or major components that effectively summarize a large portion of variation in a complex dataset (Ait-Sahalia and Xiu 2019). The method is very helpful for the selection of multiple traits, determining the most suitable that mostly contribute to the variation of the selected materials (Nardo et al. 2005). In this study, the accumulated contribution of the first two principal components captured 68.7% of the total variance, slightly higher than the results of Salvatore et al. (2010) and Turna (2004) for the morphological traits of cones and seeds of *P. halepensis* and *Picea orientalis* (L.) Link, respectively. This may be caused by the significant correlation between

these traits. PC I represents seed and nut traits, PC II and PC III represents cone traits and cone number. This finding confirms the correlation results in this study, and similar results have been found for growth traits and wood properties (Wang et al. 2018), indicating that principal component analysis is effective.

Comprehensive assessment and genetic gain

Breeding targets determine the research methods used in tree genetics. *P. koraiensis* is an important economic tree species, therefore, seed yield is a critical trait when evaluating clones (Jiang et al. 2018). Although comprehensive assessment is an important method for selecting elite material when analyzing multiple traits, this may reduce genetic gain when many weak correlated traits are considered (Pollak et al. 1987). When traits are negatively correlated, they should be separated in the evaluation of elite materials (Yin et al. 2017). In this study, cone number was significantly negatively correlated with the other traits; therefore, cone number should be evaluated alone according to its significance among the investigated traits. Using cone number and other traits (according to PC I results) as selection criteria, with a selection rate of 10%, 22 clones were selected from cone number and Qi data, respectively. Based on cone number, the genetic gain was 7.2%, lower than that found by Jiang et al. (2018) in the same seed orchard. This may possibly be due to the use of single year data and periodic hardening in this species used by Jiang et al. (2018). Based on the other traits, genetic gains ranged from 6.2 to 24.3%, which are lower than the results for *Pinus pinea* L. (Mutke et al. 2005) and higher than *Pinus gerardiana* (Singh 1992). This difference in genetic gain may be due to tree species, selection rates, and environments, but in general, the genetic gain was the expected selection effect in the *P. koraiensis* genetic improvement program.

Economic benefit analysis

The most important purpose for tree breeding is to promote ecological and economic benefits (Dyjakon 2019). The economic gains that can be generated from improved breeding material are of considerable significance in evaluating whether the investment can be recovered (Ivković et al. 2010). For *P. koraiensis*, aside from ecological benefits, improved wood and seed yields are two important economic benefits. With the ban on cutting natural *P. koraiensis* stands, seed yields of improved clones have even greater importance. Considering that the average genetic gains for cone number (7.2%), cone weight (17.9%), seed weight (24.3%) and weight per thousand seed (11.4%) were greater than 11.0%, a value of 10.0% may be assumed for the estimated genetic gain of the selected clones for calculating economic benefit.

Currently, the price of *P. koraiensis* seeds in Northeastern China is approximately US \$14/kg. With an average seed yield of 1.1 kg/tree/year (average cone number \times seed number per cone \times thousand seed weight/1000), the estimated economic gain is US \$1.76/tree. If elite trees are planted at a density of 7 m \times 7 m on 10,000 hectares, the estimated economic gain after 40 years will be US \$ 3.52 million per year.

Conclusions

The selection of elite or plus trees for a breeding program is a critical step to increase genetic quality and thus reach improvement goals. The range of consumptive uses of *P. koraiensis* seeds constitutes a strong motive for tree improvement. In this study, 14 traits of cones, seeds and nuts were examined in 110 *P. koraiensis* clones to evaluate clone variability and productivity. The results show moderate to high significant differences in cone, seed and nut traits, which supports the selection method. Positive and significant correlations were observed among cone, seed and nut dimensions, but cone number was negatively weakly correlated to cone, seed and nut dimensions, suggesting independent selection of elite clones, and that correlated traits together support an efficient, indirect clone selection. Phenotypic coefficients of variance were slightly higher for cone characteristics than for those of seeds and nuts. Higher repeatability was observed for cone, seed and nut traits. These parameters of repeatability and phenotypic coefficients of variance suggest substantial genetic variation among clones, and that the selection of quality materials was feasible. Cone, seed and nut dimensions and cone numbers were separately identified as indices in the selection evaluation of clones based on correlation and principal component analysis. Thus, 22 elite clones were selected based on a 10% selection rate, corresponding with a genetic gain from 6.2 to 24.3% in all cone, seed and nut traits. These results provide beneficial information to select high quality clones and the elite clones can supply resources for planting.

Author contributions DKK, YL and DZ have contributed equally to this work.

References

- Ait-Sahalia Y, Xiu DC (2019) Principal component analysis of high-frequency data. *J Am Stat Assoc* 114:287–303. <https://doi.org/10.1080/01621459.2017>
- Bai FY, Kang N, Zhang PD, Kang XY (2019) Selection of female parents with high fertility and high combining abilities for cross-breeding *Populus tomentosa*. *J For Res* 30(2):445–450
- Brancher M, Piringer M, Franco D, Belli Filho P, Lisboa H, Schaubberger G (2019) Assessing the inter-annual variability of separation distances around odor sources to protect the residents from odor annoyance. *J Environ Sci* 79:11–24
- Davis L (1967) Investments in loblolly pine clonal seed orchards: production costs and economic potential. *J For* 65:882–887
- Destaillets F, Cruz-Hernandez C, Giuffrida F, Dionisi F (2010) Identification of the botanical origin of pine nuts found in food products by gas-liquid chromatography analysis of fatty acid profile. *J Agric Food Chem* 58:2082–2087
- Dyjakon A (2019) The influence of apple orchard management on energy performance and pruned biomass harvesting for energetic applications. *Energies* 12:632
- Feng FJ, Sui X, Chen MM, Zhao D, Han HJ, Li MH (2010) Mode of pollen spread in clonal seed orchard of *Pinus koraiensis*. *J Biophys Chem* 1:33
- Field A (2013) *Discovering statistics using IBM SPSS statistics*. Sage Publication Ltd 1, Oliver's Yard
- Gonçalves E, Graça A, Martins A (2019) Grapevine clonal selection in Portugal: a different approach. *BIO Web Conf.* <https://doi.org/10.1051/bioconf/20191201003>
- Guerra F, Richards J, Fiehn O, Famula R, Stanton B, Shuren R, Sykes R, Davis M, Neale D (2016) Analysis of the genetic variation in growth, ecophysiology, and chemical and metabolic composition of wood of *Populus trichocarpa* provenances. *Tree Genet Genomes* 12(1):6–24
- Hammer Y, Harper D (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electron* 4(1):1–9
- Hietz P, Rosner S, Hietz S, Wright S (2017) Wood traits related to size and life history of trees in a Panamanian rainforest. *New Phytol* 213:170–180
- Ivković M, Wu H, Kumar S (2010) Bio-economic modelling as a method for determining economic weights for optimal multiple-trait tree selection. *Silvae Genet* 59:77–90
- Jiang GY, Jiang LP, Song SL, Wang JY, Wang Q, Wang LF, Zhang P, Zhao XY (2018) Genetic variance analysis and excellent fruit-timber families selection of half-sib *Pinus koraiensis*. *Bull Bot Res* 38(5):775–784 **in Chinese with English abstract**
- Jiang LP, Wang JY, Zhang P, Liang DY, Zhang QH, Wang BY, Pei XN, Zhao XY (2019) Variation and selection of growth and fruit traits among 170 *Pinus koraiensis* Clones. *For Res* 32(1):58–64
- Li CZ, Löfgren KG (2000) A theory of red pine (*Pinus koraiensis*) management for both timber and commercial seeds. *For Sci* 46:284–290
- Li Y, Suontama M, Burdon RD, Dungey HS (2017) Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genet Genomes* 13:60. <https://doi.org/10.1007/s11295-017-1144-x>
- Liang DY, Wang BY, Song SL, Wang JY, Wang LF, Wang Q, Ren XB, Zhao XY (2019) Analysis of genetic effects on a complete diallel cross test of *Pinus koraiensis*. *Euphytica*. <https://doi.org/10.1007/s10681-019-2414-5>
- Man BX, Zhang L, Sun JZ, Yuan HS, Sun JJ, Wang XY, Sun YB, Qiu L, Wang YL (2012) Wang, YZ (2012) On plant growth regulator accelerate seed setting earlier of *Pinus koraiensis*. *J Beihua Univ (Nat Sci)* 13(3):329–334
- Matziris D (1998) Genetic variation in cone and seed characteristics in a clonal seed orchard of Aleppo pine grown in Greece. *Silvae Genet* 47:37–41
- Miyaki M (1987) Seed dispersal of the Korean pine, *Pinus koraiensis*, by the red squirrel, *Sciurus vulgaris*. *Ecol Res* 2:147–157
- Mutke S, Gordo J, Gil L (2005) Cone yield characterization of a stone pine (*Pinus pinea* L.) clone bank. *Silvae Genet* 54:189–197
- Nardo M, Saisana M, Saltelli A, Tarantola S (2005) Tools for composite indicators building. *Eur Commission, Ispra* 15:19–20

- Pan YY, Pei XN, Wang FW, Wang CL, Shao LL, Dong LH, Zhao XY, Qu GZ (2019) Forward, backward selection and variation analysis of growth traits in half-sib *Larix kaempferi* families. *Silvae Genet* 68:1–8
- Pollak E, Van der Werf J, Quaa R (1984) Selection bias and multiple trait evaluation. *J Dairy Sci* 67:1590–1595
- Rawat K, Bakshi M (2011) Provenance variation in cone, seed and seedling characteristics in natural populations of *Pinus wallichiana* A. B. Jacks (blue pine) in India. *Ann For Res* 54:39–55
- Ren X, He H, Zhang L, Li F, Liu M, Yu G, Zhang J (2018) Modeling and uncertainty analysis of carbon and water fluxes in a broad-leaved Korean pine mixed forest based on model-data fusion. *Ecol Model* 379:39–53
- Rosengarten J (2004) The book of edible nuts. Courier Corporation, Dover Publications Inc., p 375
- Roy SM, Thapliyal R, Phartyal S (2004) Seed source variation in cone, seed and seedling characteristic across the natural distribution of Himalayan low level pine *Pinus roxburghii* Sarg. *Silvae Genet* 53:116–123
- Salvatore R, Moya D, Pulido L, Lovreglio R, LópezSerrano F, De Las HJ, Leone V (2010) Morphological and anatomical differences in Aleppo pine seeds from serotinous and non-serotinous cones. *New For* 39:329–341
- Sevik H, Topacoglu O (2015) Variation and inheritance pattern in cone and seed characteristics of Scots pine (*Pinus sylvestris* L.) for evaluation of genetic diversity. *J Environ Biol* 36:1125–1130
- Singh N (1992) Propagation, selection and establishment of clonal seed orchard of chilgoza pine (*Pinus gerardiana* Wall.). *Indian For* 118:901–908
- Sivacioglu A (2010) Genetic variation in seed cone characteristics in a clonal seed orchard of Scots pine (*Pinus sylvestris* L.) grown in Kastamonu-Turkey. *Rom Biotechnol Lett* 15:5695–5701
- Sivacioglu A, Ayan S (2008) Evaluation of seed production of Scots pine (*Pinus sylvestris* L.) clonal seed orchard with cone analysis method. *Afr J Biotechnol* 7(24):4393–4399
- Turna I (2004) Variation of morphological characters of oriental spruce (*Picea orientalis*) in Turkey. *Biologia-Bratislava* 59:519–526
- Wang H, Hong J (2004) Genetic resources, tree improvement and gene conservation of five-needle pines in East Asia. *USDA Forest Service Proceedings RMRS-P* 32: 73–78
- Wang F, Zhang QH, Tian YG, Yang SC, Wang HW, Wang LK, Li YL, Zhang P, Zhao XY (2018) Comprehensive assessment of growth traits and wood properties in half-sib *Pinus koraiensis* families. *Euphytica*. <https://doi.org/10.1007/s10681-018-2290-4>
- Xia H, Zhao GH, Zhang LS, Sun XY, Yin SP, Liang DY, Li Y, Zheng M, Zhao XY (2016) Genetic and variation analyses of growth traits of half-sib *Larix olgensis* families in northeastern China. *Euphytica* 212:387–397
- Xiao Y, Ma WJ, Lu N, Wang Z, Wang N, Zhai WJ, Kong LS, Qu GZ, Wang QX, Wang JH (2019) Genetic variation of growth traits and genotype-by-environment interactions in clones of *Catalpa bungei* and *Catalpa fargesii* f. *duclouxii*. *Forests* 10:57. <https://doi.org/10.3390/f10010057>
- Xu ZW, Yu GR, He NP, Wang QF, Wang SZ, Xu XF, Wang RL, Zhao N (2018) Biogeographical patterns of soil microbial community as influenced by soil characteristics and climate across Chinese forest biomes. *Appl Soil Ecol* 124:298–305
- Yin SP, Xiao ZH, Zhao GH, Zhao X, Sun XY, Zhang Y, Wang FW, Li SC, Zhao XY, Qu GZ (2017) Variation analyses of growth and wood properties of *Larix olgensis* clones in China. *J For Res* 28(4):687–697
- Yuan HW, Niu SH, Zhou XQ, Du QP, Li Y, Li W (2016) Evaluation of seed production in a first-generation seed orchard of Chinese pine (*Pinus tabulaeformis*). *J For Res* 27:1003–1008
- Zhang Z, Zhang HG, Zhou Y, Liu L, Yu HY, Wang X, Feng WJ (2015) Variation of seed characters in Korean pine (*Pinus koraiensis*) multi-clonal populations. *J Beijing For Univ* 37:67–78
- Zhao XY, Xia H, Wang XW, Wang C, Liang DY, Li KL, Liu GF (2016) Variance and stability analyses of growth characters in half-sib *Betula platyphylla* families at three different sites in China. *Euphytica* 208:173–186
- Zheng HQ, Hu DH, Wang RH, Wei RP, Yan S (2015) Assessing 62 Chinese Fir (*Cunninghamia lanceolata*) breeding parents in a 12-year grafted clone test. *Forests* 6:3799–3808
- Zhu BA, Wang XP, Fang JY, Piao SL, Shen HH, Zhao SQ, Peng CH (2010) Altitudinal changes in carbon storage of temperate forests on Mt Changbai, Northeast China. *J Plant Res* 123:439–452

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.